Structure-function relationships underlying the dual N-acetylmuramic and N-acetylglucosamine specificities of the bacterial peptidoglycan deacetylase PdaC

Laia Griffol-Romero1, María Angela Sainz-Polo1, David Albesa-Jove1,6, Marcelo E. Guerin5,6, Xevi Biarnés1, and Antoni Planas1,3

From the 1Laboratory of Biochemistry, Institut Químic de Sarrià, University Ramon Llull, 08017 Barcelona, Spain, the 2Structural Biology Unit, Center for Cooperative Research in Biosciences (CIC bioGUNE), Bizkaia Technology Park, Ed. 801A, 48160 Derio, Spain, and the 3Basque Foundation for Science (IKERBASQUE), 48011 Bilbao, Spain

Edited by Gerald W. Hart

Bacillus subtilis PdaC (BsPdaC) is a membrane-bound, multidomain peptidoglycan N-deacetylase acting on N-acetylmuramic acid (MurNAc) residues and conferring lysozyme resistance to modified cell wall peptidoglycans. BsPdaC contains a C-terminal family 4 carbohydrate esterase (CE4) catalytic domain, but unlike other MurNAc deacetylases, BsPdaC also has GlcNAc deacetylase activity on chitooligosaccharides (COSs), characteristic of chitin deacetylases. To uncover the molecular basis of this dual activity, here we determined the X-ray structure of the BsPdaC CE4 domain at 1.54 Å resolution and analyzed its mode of action on COS substrates. We found that the minimal substrate is GlcNAc3 and that activity increases with the degree of glycan polymerization. COS deacetylation kinetics revealed that BsPdaC operates by a multiple-chain mechanism starting at the internal GlcNAc units and leading to deacetylation of all but the reducing-end GlcNAc residues. Interestingly, BsPdaC shares higher sequence similarity with the peptidoglycan GlcNAc deacetylase SpPgdA than with other MurNAc deacetylases. Therefore, we used ligand docking simulations to analyze the dual GlcNAc- and MurNAc-binding specificities of BsPdaC and compared them with those of SpPgdA and BsPdaA, representing peptidoglycan deacetylases highly specific for GlcNAc or MurNAc residues, respectively. BsPdaC retains the conserved Asp-His-His metal-binding triad characteristic of CE4 enzymes acting on GlcNAc residues, differing from MurNAc deacetylases that lack the metal-coordinating Asp residue. BsPdaC contains short loops similar to those in SpPgdA, resulting in an open binding cleft that can accommodate polymeric substrates. We propose that PdaC is the first member of a new subclass of peptidoglycan MurNAc deacetylases.

The bacterial cell wall peptidoglycan (PGN) is an elaborate polymeric mesh composed of a glycan chain of alternating β1,4-linked N-acetylmuramic (GlcNAc) and N-acetylmuramic acid (MurNAc) units cross-linked via peptidyl bridges attached to the 3-O-lactoyl group of MurNAc residues. Pathogenic bacteria utilize acetylation (6-O-acetylation of MurNAc) and deacetylation (2-N-deacetylation of GlcNAc and/or MurNAc residues) of their cell wall PGN to evade detection by the innate immune system. N-Deacetylation was first identified in 1971 in lysozyme-resistant Bacillus cereus strains, which contain high proportions of nonacetylated glucosamine (GlcN) residues in the cell wall PGN (1). Later, nonacetylated muramic acid residues (MurN) were identified in the PGN of Bacillus anthracis (2). The presence of deacetylated sugars in PGN strongly reduces the activity of the muramidase lysozyme, which hydrolyzes the MurNacβ(1→4)GlcNAc glycosidic bonds of the glycan strands. Lysozyme is an important factor of the innate immune system in humans, it is present in many tissues and body fluids, and it is secreted in large amounts by cells of the immune system at the site of infection (3, 4). The innate immune system senses intact PGN and PGN fragments originated by lysozyme and other lytic enzymes using diverse recognition mechanisms, which include peptidoglycan recognition proteins, Toll-like receptors, and Nod-like receptors (3). Efficient PGN degradation promotes increased duration and intensity of inflammatory signals and instructs immune responses to infection. Bacteria that are able to modify the structure of PGN to make it more resistant to lysozyme both evade the antibacterial activity of lysozyme and delay or suppress pro-inflammatory immune responses (3).

The abbreviations used are: PGN, peptidoglycan; MurNAc, N-acetylmuramic acid; GlcN, nonacetylated glucosamine; 3D, three-dimensional; aa, amino acids; RMSD, root mean square deviation; MDP, N-acetylmuramyl-dipeptide; FL, full-length; CD, catalytic domain; AoOMU, 4-methylumbelliferyl acetate; MU, 4-methylumbelliferyl; DMF, N,N-dimethylformamide; PDB, Protein Data Bank; COS, chitooligosaccharide; CE4, carbohydrate esterase family 4.

This is an open access article under the CC BY license.
Deacetylation of MurNAc residues of PGN is also involved in endospore formation and germination. Spores of *Bacillus* species have a thick peptidoglycan (spore cortex) that contains high abundance of spore-specific muramic acid δ-lactam. Two enzymes are required for its synthesis, the amidase CwlD (muramoyl-L-alanine amidase) that removes the peptide side chain from the 3-O-lactoyl group of MurNAc residues and a MurNAc deacetylase to generate the free amino group for lactam formation (4, 5). *Bacillus subtilis* has six identified or predicted polysaccharide deacetylases: PdaA and PdaB are MurNAc deacetylases shown to be involved in sporulation (6, 7). PdaC was reported to be also a MurNAc deacetylase but not associated with sporulation and germination (8). The other gene products, YlxY, YxkH, and YheN, have not been characterized (it is not known whether they deacetylate GlcNAc or MurNAc residues), but they are not associated with sporulation (7, 8).

Peptidoglycan GlcNAc deacetylases (EC 3.5.1.104) and MurNAc deacetylases (EC 3.5.1.1.) are classified in carbohydrate esterase family 4 (CE4) in the CAZy (Carbohydrate Active Enzyme) database (www.cazy.org)5 (9) together with chitin deacetylases (EC 3.5.1.41), poly-B1,6-GlcNAc deacetylases (EC 3.5.1.1.), and some acetylxylan esterases (EC 3.1.1.72). CE4 enzymes share a conserved region known as the NodB homologous domain due to its similarity to the NodB oligosaccharide deacetylase, one of the first deacetylases of this family to be characterized (10). They operate by metal-assisted general acid/base catalysis, as first proposed for the *Streptococcus pneumoniae* peptidoglycan GlcNAc deacetylase SpPgdA when its X-ray structure was solved (11). The family is characterized by the conservation of five sequence motifs (named MT1 to -5), which include the Asp-His-His metal-binding triad (second Asp (D) of the MT1 sequence FTDDG and the two histidines (H) of the MT2 sequence H(S/T)XXH), the general base Asp (first D in MT1), and the general acid His in MT5 (I(V/I)LXHD) (11, 12).

A number of peptidoglycan GlcNAc deacetylases have been characterized, and their 3D structures determined by X-ray crystallography (Table S1) (12). They are specific for GlcNAc residues in the peptidoglycan chain, but they have also been shown to deacetylate COS, homooligomers of GlcNAc residues, but with lower efficiency. In contrast, few peptidoglycan MurNAc deacetylases have been biochemically characterized, although this activity has been annotated to a number of putative polysaccharide deacetylases based on genetic knockout studies (13). The *B. subtilis* PdaA deacetylates MurNAc residues of peptidoglycan devoid of the peptide linked to the muramic acid 3-O-lactoyl group of MurNAc residues, which is consistent with its function during sporulation to form muramic acid δ-lactam residues in the spore cortex peptidoglycan (6, 14). It is specific for MurNAc residues, and it is not active on COS. The crystal structure of BsPdaA revealed a modified (β/α)5 helix and/or one β-strand of the β/α barrel) (11, 12, 16). A second structure for a putative MurNAc deacetylase,

Table 1

**Deacetylation activity of BsPdaC-CD on PGN and chitooligosaccharides**

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Specific activity (v) [EF]</th>
</tr>
</thead>
<tbody>
<tr>
<td>(GlcNAc)_3</td>
<td>No activity</td>
</tr>
<tr>
<td>(GlcNAc)_4</td>
<td>ND</td>
</tr>
<tr>
<td>(GlcNAc)_5</td>
<td>3.72 ± 0.07</td>
</tr>
<tr>
<td>(GlcNAc)_6</td>
<td>2.22 ± 0.06</td>
</tr>
<tr>
<td>(GlcNAc)_7</td>
<td>1.08 ± 0.04</td>
</tr>
<tr>
<td>(GlcNAc)_8</td>
<td>0.75 ± 0.03</td>
</tr>
<tr>
<td>(GlcNAc)_9</td>
<td>0.50 ± 0.02</td>
</tr>
</tbody>
</table>

5 Please note that the JBC is not responsible for the long-term archiving and maintenance of this site or any other third party hosted site.

that of the BaCE4 from *B. anthracis*, has been determined (17). Although not biochemically characterized, it is likely to be a MurNAc deacetylase due to its close similarity to BsPdaA. Both enzymes lack the otherwise conserved Asp residue of the metal-binding catalytic triad, having an Asn residue that points away from the active metal ion into the core of the protein (17).

More recently, Kobayashi et al. (8) identified a novel peptidoglycan MurNAc deacetylase, BsPdaC, whose expression is regulated by the essential YycFG two-component system that controls cell wall metabolism in *B. subtilis* (18, 19). Interestingly, BsPdaC deacetylates MurNAc residues of intact peptidoglycan, or at least with the 1-Ala–D-Glu dipeptide attached to the muramic 3-O-lactoyl group, being inactive on the glycan backbone devoid of peptidyl substitutions. More striking is that, as opposed to BsPdaA, BsPdaC also deacetylates COS, an activity that was thought to be restricted to GlcNAc deacetylases.

To understand the molecular bases of such dual activity, we report the X-ray 3D structure of the CE4 catalytic domain of BsPdaC and the mode of action on COS substrates. The dual GlcNAc- and MurNAc-binding specificity at the catalytic site is analyzed by using a combination of X-ray crystallography, enzyme kinetics, ligand-docking simulations, and protein biochemistry/biophysics. We propose that BsPdaC represents the first member of a novel subclass of MurNAc deacetylases.

**Results**

**Enzyme kinetics and deacetylation pattern**

BsPdaC is a multidomain protein composed of a transmembrane region (aa 7–26), two domains with unknown function (aa 54–138 and 156–235), and a CE4 catalytic domain (Fig. S1). The full-length BsPdaC (residues 27–467, lacking the transmembrane region, here referred to as BsPdaC-FL) and the isolated catalytic domain (BsPdaC-CD, aa 270–467) were recombinantly expressed and purified to apparent homogeneity (Fig. S2).

BsPdaC-CD and BsPdaC-FL displayed similar specific activity with PGN (63.1 and 81.8 min⁻¹, respectively) (Table 1) as well as similar thermal stability (Fig. S3). The isolated catalytic domain is expected to have the same MurNAc specificity as the full-length enzyme reported previously (8). To assess it, PGN was incubated with BsPdaC-FL for an extended period, and the
Structure and specificity of BsPdaC

released acetate was quantified. After dialysis to remove the free acetate, the product was incubated with BsPdaC-CD, and it was observed that essentially no further deacetylation took place upon a long incubation time. A parallel experiment where the order of the enzymes was inverted (first incubation with BcPdaC-CD, then with BcPdaC-FL) gave the same result (Table S2). This indicates that both enzymes deacetylate the same residues (MurNAc) of PGN. Moreover, when the final deacetylated PGN product from both experiments was treated with SpPgda (a deacetylase specific for GlcNAc residues), an additional release of acetate was observed, with a specific activity of 70–80% that of SpPgda on intact PGN (Table S2). It is then concluded that the isolated catalytic domain has the same specificity and retains the same MurNAc specificity as the full-length enzyme. The isolated catalytic domain was used for further studies.

Enzyme activity on COS substrates was determined by two methods: by quantification of the free amino groups generated from deacetylation of (GlcNAc)₅ substrates using the fluorescamine reagent (11), and by HPLC-MS analysis of the deacetylation reaction in which the sequence of GlcNAc deacetylation events could be monitored (16). The pH and temperature profile of the enzyme activity with (GlcNAc)₅ substrate showed a pH optimum of 8.5 at 37 °C and an optimum temperature of 55 °C at pH 7 (Fig. S3).

BsPdaC is a metal-dependent deacetylase as most of the CE4 enzymes (12). Treatment with 20 mM EDTA only reduced the activity by 10% relative to untreated and freshly purified enzyme, indicating that the metal cation is strongly bound to the active site. Conditions to generate the apo-form of the enzyme for reconstitution experiments included increasing concentrations of urea in the presence of EDTA. After treatment with 200 mM EDTA in 2 M urea, the recovered protein after dialysis still retained 65% activity, and full activity could be restored upon the addition of Zn²⁺ cation. Treatment with 7 M urea and 200 mM EDTA was required to dissociate the metal cation. Protein refolding was quantitative, resulting in an inactive protein that recovered 100% activity upon the addition of Zn²⁺ cation (Fig. S3C).

Enzyme activity was evaluated with COS substrates (GlcNAc)ₙ, n = 2–5. The enzyme was inactive with the disaccharide substrate, the trisaccharide was the minimal COS substrate, and the activity increased with the degree of polymerization (Table 1). Although Kobayashi et al. (8) reported single monodeacetylation activity of BsPdaC on (GlcNAc)₅, we have found a more complex mode of action. HPLC-MS time course monitoring of the deacetylation reaction revealed that the enzyme sequentially deacylates different GlcNAc units, leading to final products in which all but one GlcNAc unit were deacetylated (A1D1, A2D2, and A3D3). For the chitotetraose substrate, mono-, di-, and trideacetylated products were sequentially formed (A3D1, A2D2, and A1D1, where A represents GlcNAc and D is GlcN). Likewise, with the chitopentaose substrate, A4D1, A3D2, A2D3, and A1D4 were generated along the reaction. Fig. 1 (B and C) summarizes the evolution of the different products. Whereas the final A1D3 from tetracetylchitotetraose was detected at rather short reaction times, the final A1D4 product from pentaacetylchitopentaose was only detected after a long reaction time. The structure of the different products with regard to their pattern of acetylation was determined by the MALDI-TOF-MS/MS method according to Cord-Landwer et al. (20), consisting in ¹⁸O labeling of the reducing end and re-N-acetylation with [²H₆]acetic anhydride, which allows the identification and quantification of the different products in the reaction mixtures (Table S3). BsPdaC follows a multiple-chain mechanism (or distributive mechanism) as presented in Fig. 2A. Interestingly, with the tetrascarbohydrate substrate, the first deacetylation occurs at the two central GlcNAc units with equal efficiencies, whereas with the pentaosaccharide substrate, any of the three internal GlcNAc units are also deacetylated but with preference for the second residue from the reducing end (51% as compared with 14 and 35% for the other two internal residues). Deacetylation proceeds in time, leading to final products in which all but the reducing-end GlcNAc units are deacylated (DDDA and DDDDA, respectively).

Initial rates corresponding to the first deacetylation event leading to monodeacetylated products were determined by HPLC-MS. Likewise, the same values were obtained by the fluorescamine method at short reaction times, corresponding to monodeacetylation, as verified by HPLC-MS. The specific activity on (GlcNAc)₅ was twice that on (GlcNAc)₄ and 10-fold higher than the activity with the shortest substrate (GlcNAc)₃ (Table 1). Michaelis–Menten kinetic parameters were determined for the tetrascarbohydrate substrate (Fig. S4), with apparent $k_{cat}$ and $K_m$ values of 0.93 ± 0.03 s⁻¹ and 6.25 ± 0.47 mM, respectively, and apparent $k_{cat}/K_m$ of 149 M⁻¹ s⁻¹, which are of the same order as those reported for a peptidoglycan substrate (8).

The increase in activity from a chitotriose to a chitopentaose substrate suggests that the binding site cleft has at least six subsites, from -3 to +2, to account for the different activities when considering the first deacetylation event (Fig. 2B). For the chitotetraose substrate, three subsites (−1, 0, +1) explain the observed monodeacetylation pattern, but additional subsites −2 and +2 may exist, with similar affinities to agree with the fact that both binding modes are alike (50% deacetylation at each central GlcNAc unit). With the chitopentaose substrate, if subsites −2 and +2 have similar affinities, a new subsite −3 might be required to account for the preferred binding mode 1 (BM1’, Fig. 2B) (51%) over BM2’ (35%), because BM1’ does not occupy subsite +2. A potential subsite +3 is not required, because BM3’ is the least preferred binding mode (14%) that may be accounted for by the loss of interactions with subsite −2. Binding to subsite +1 is required in any case, because it is always occupied not only in the different binding modes for the first monodeacetylation reactions, but also in the subsequent deacetylations leading to the final products in which the reducing-end GlcNAc unit is not deacetylated.

The crystal structure of BsPdaC-CD

The 3D structure of BsPdaC-CD was obtained by X-ray crystallography at 1.54 Å resolution (Table S4 and Fig. 3). BsPdaC-CD crystallized as a homodimer. The protein monomer adopts a (β/α)₇ barrel topology. The central core comprises seven parallel β-strands that form a distorted β-barrel
Figure 1. MS time monitoring of the BsPdaC-CD deacetyase reactions with substrates (GlcNAc)$_4$ and (GlcNAc)$_5$. A, MS spectra at increasing reaction time. B, evolution of peak areas for [M + H]$^+$ ions of the deacetylated products. C, magnification of B at short reaction time. Relative areas from the MS spectra are given.
surrounded by α-helices. The putative ligand-binding site is located in the central region of the β-barrel and is flanked by a series of connecting loops.

The structure contains a Zn$^{2+}$ ion coordinated with Asp$^{286}$, His$^{336}$, and His$^{340}$, which form the conserved metal-binding triad of CE4 enzymes (with few exceptions; see below), and two oxygens of a tartaric acid molecule from the crystallization buffer (Fig. 3C). The conserved catalytic residues, general base Asp$^{285}$ and general acid His$^{427}$, have the right orientation in the active site for catalysis as in other CE4 crystal structures. CE4 enzymes differ in a series of surface loops that shape the binding site cleft and are structural elements involved in substrate specificity, as proposed in the “subsite-capping model” (16, 21). Loops 1–6 (loop 1, residues 313–320, β2-α2), loop 2 (341–348, β3-α3), loop 3 (376–382, β4-α4), loop 4 (401–410, β5-α5), loop 5 (419–421, α5-β6), and loop 6 (428–437, β6-α6).

Figure 2. BsPdaC pattern of deacetylation. A, deacetylation sequence of COS (GlcNAc)$_4$ and (GlcNAc)$_5$. Symbolic representation of glycans is according to Ref. 47, with the reducing end on the right. Dashed arrow, slower deacetylations at longer reaction times (see Table S2). B, binding of COS ligands to BsPdaC. Shown are binding modes for the first deacetylation event for substrates (GlcNAc)$_4$ and (GlcNAc)$_5$. Negative subsites are on the nonreducing end, and positive subsites are on the reducing end. Subsite 0 is the catalytic site.

Figure 3. The crystal structure of BsPdaC-CD. A, cartoon representation of the overall structure of BsPdaC-CD showing the loops (L1–L6) surrounding the active site of the CE4 domain. B, surface representation of BsPdaC-CD showing the COS-binding site. C, active site of BsPdaC-CD showing the catalytic residues (Asp$^{285}$, general base; His$^{336}$, general acid) and metal-binding triad (Asp$^{286}$-His$^{336}$-His$^{340}$), a Zn$^{2+}$ ion, and carboxylate oxygens of a tartaric acid molecule (TLA) from the crystallization buffer. Loops according to the subsite capping model are as in Fig. S11 and flanking secondary structures are as follows: loop 1 (residues 313–320, β2-α2), loop 2 (341–348, β3-α3), loop 3 (376–382, β4-α4), loop 4 (401–410, β5-α5), loop 5 (419–421, α5-β6), and loop 6 (428–437, β6-α6).
6 C-terminal residues of the Strep-tag are disordered and not seen in the electron density map). The C-terminal α-helix of one monomer runs parallel to the substrate binding cleft of the other monomer. The side chain of Arg\textsuperscript{477} protrudes perpendicularly from the helix axis, and the guanidyl group interacts with His\textsuperscript{427} (general acid), Asp\textsuperscript{286} (metal coordination), and a tartaric acid molecule (from the crystallization mother liquor) in the active site of the other monomer. The orientation of this terminal helix is probably a consequence of the crystal packing because the enzyme in solution is fully active, and this dimerization contact in the X-ray structure would block access of the substrate to the active site.

Despite much effort, attempts to co-crystallize BsPdaC-CD-D285S (inactive mutant at the general base residue) with the tetrasaccharide ligand (GlcNAc\textsubscript{4}) were unsuccessful, probably due to the presence of the extended C-terminal α-helix. The structure of the mutant superimposes with that of the WT enzyme with an RMSD of 0.27 Å. A new construct moving the Strep-tag to the N terminus and leaving an intact C terminus with the original native sequence was prepared (BsPdac-CD-NtStrep and purified following the same protocol. The WT enzyme had the same specific activity on COS substrates as the initial BsPdaC-CD enzyme. Upon mutation of the general base residue (Asp\textsuperscript{285}), this new construct again failed to co-crystallize with the (GlcNAc\textsubscript{4}) substrate. It is worth noting that ligand binding was assessed by monitoring the thermal unfolding by CD in the absence and presence of COS substrates (Fig. S7), observing a $T_m$ increase of 3.3 and 3.7 °C upon binding of (GlcNAc\textsubscript{4})\textsubscript{1} and (GlcNAc\textsubscript{4})\textsubscript{3}, respectively.

**Enzyme-ligand complexes by ligand-docking simulations**

Computational docking of (GlcNAc\textsubscript{4})\textsubscript{1} on the X-ray structure of BsPdaC-CD showed two preferred binding modes with low energy (Fig. 4) in which either of the two internal GlcNAc residues are located in the catalytic site (subsite 0) and would render the AADA and ADAA products. Both binding modes are of similar energy, indicating 50% population each, which is in agreement with the observed enzyme specificity for the first deacetylation event on this substrate as shown in Fig. 2. The substrate expands subsites $-2$ to $+2$, and all GlcNAc units interact with protein residues through hydrogen bonding and stacking interactions (Fig. 4C). A subsite $-3$ is not seen in these models with a tetrasaccharide ligand, but additional surface interactions may occur with a longer and more reactive pentasaccharide substrate.

The enzyme-ligand models (Fig. 4C) predict some key interactions that were probed by site-directed mutagenesis. Tyr\textsuperscript{430}, located in loop 6, interacts with the substrate through a stacking interaction with the GlcNAc unit at subsite $-2$ and hydrogen-bonds with the GlcNAc unit at subsite $-1$. Mutation to Ala (Y430A) reduced the specific activity to 6% of the WT activity. Likewise, Trp\textsuperscript{402} is located in loop 4 and establishes a tight hydrophobic (stacking) interaction with the GlcNAc unit in subsite 0, on the opposite side (relative to the sugar plane) to the metal binding triad (Aps\textsuperscript{286}-His\textsuperscript{336}-His\textsuperscript{340}). Mutation to Ala (W402A) has a drastic effect, with a residual activity <0.1% than that of the WT enzyme.

To obtain further insights into the substrate specificity of BsPdaC, a series of fragment-based ligand-docking simulations were conducted with monosaccharide (MurNAc and GlcNAc) and disaccharide (GlcNAc\textsubscript{β}(1→4)MurNAc and MurNAc\textsubscript{β}(1→4)GlcNAc) probes. Similar docking experiments on the 3D structure of BsPdaA and SpPgdA, as representative peptidoglycan deacetylases specific for MurNAc and GlcNAc units, respectively, were also analyzed for comparison (Figs. S8–S10). These docking experiments illuminate significant differences between GlcNAc- and MurNAc-specific enzymes and the here-reported BsPdaC with dual specificity, depending on the substrate.

The monosaccharide GlcNAc probe is able to fit into the active site of BsPdaC with a proper orientation for catalysis (Fig. S8A); the C=O of the acetamido group is coordinated with the Zn\textsuperscript{2+} cation (distance 2.1 Å), and the NH is at hydrogen-bonding distance to the general acid His-427 (3.4 Å). No productive binding was observed for the MurNAc probe (Fig. S8B). However, with the disaccharide GlcNAc\textsubscript{β}(1→4)MurNAc probe, productive binding placed the MurNAc unit in subsite 0 and the GlcNAc unit in subsite $-1$ (Fig. S8C). In addition, as shown before, docking of the (GlcNAc\textsubscript{4})\textsubscript{1} substrate also gave productive complexes for catalysis that were in agreement with the experimentally observed specificity of the enzyme. The active site (subsite 0) is thus able to accommodate both GlcNAc and MurNAc residues in a productive binding mode for catalysis. Because the PGN glycan chain is composed of alternating GlcNAc and MurNAc units, it seems that subsites $-1$ and/or +1 are unable to accommodate a MurNAc residue, even in the absence of peptidyl substitutions, because PGN digested with l-alanine amidase is not deacylated (8).

Docking of (GlcNAc\textsubscript{4})\textsubscript{1} showed that the 3-OH of the GlcNAc unit in subsite 0 does not establish any interaction with the protein other than coordination with the Zn\textsuperscript{2+} cation (Fig. 4C), thus leaving room for the 3-lactoyl substitution (as also seen in the ligand-docking simulations with the PGN disaccharide (Fig. S8C)), but the 3-OH of the GlcNAc unit in subsite $-1$ hydrogen-bonds with Tyr\textsuperscript{430} in both binding modes, leaving no room for 3-OH substitutions. Subsite $+1$, on the other side, shows a weaker binding of the GlcNAc unit, where only the 6-OH hydrogen-bonds with protein residues (Asp\textsuperscript{599} in both binding modes and Trp\textsuperscript{402} in binding mode 2). Therefore, the simulations suggest that subsite $-1$ is specific for GlcNAc residues and dictates the specificity shown by the enzyme. BsPdaC requires at least a dipeptidyl substitution at the 3-O-lactoyl group of MurNAc residues, not included in the docking simulations with PGN disaccharide, indicating that additional interactions with the peptide substituent are important for activity on PGN substrates.

Similar fragment-based ligand-docking simulations were performed on BsPdaA (canonical PGN MurNAc deacetylase) with oligosaccharide probes (Fig. S9). The GlcNAc monosaccharide probe did not get into the active site, but mainly went to subsite $+2$, whereas the MurNAc probe gave an ensemble of docked structures compatible with productive binding in subsite 0. With the disaccharide ligands, only the MurNAc unit was found in subsite 0 in a productive orientation, consistent with the MurNAc deacetylase specificity of this enzyme that is inac-
ative on COS substrates. The opposite behavior was observed for SpPgdA, a specific GlcNAc deacetylase (Fig. S10). Only the GlcNAc monosaccharide probe was able to productively bind insubsite 0, and the disaccharide ligands always placed the GlcNAc unit into subsite 0, with disaccharide binding to subsites 0 and 1 for GlcNAc(1→4)MurNAc and to subsites −1 and 0 for MurNAc(1→4)GlcNAc, again consistent with the deacetylation specificity on GlcNAc residues of peptidoglycan and COS substrates.

BsPdaC sequence and phylogeny of family CE4 enzymes

The sequence of the catalytic domain was added to the multiple sequence alignment of CE4 family members (16), which includes chitin deacetylases, peptidoglycan deacetylases, some acetylxylan esterases, and poly-β1,6-GlcNAc deacetylases (Table S1 and Fig. S11), allowing the location of the catalytic residues within the conserved motifs of this family of enzymes (MT1 to -5) (11, 21). Motif 1 (TFDDG) contains the catalytic base (first aspartate, Asp285 in BsPdaC) and one residue of the metal coordination triad (second aspartate, Asp286 in BsPdaC), motif 2 (H(S/T)XXH) contains two histidines (His336 and His340 in BsPdaC) involved in metal coordination along with the aspartate in MT1, and motif 5 (I(V/I)LXHD) contains the catalytic acid (His427 in BsPdaC). The sequence alignment also identified the loops that are characteristic of CE4 enzymes and have been related to substrate specificity (16), which are short in BsPdaC as compared with other family members (loops 1–6, Fig. S11).

A phylogenetic analysis based on the multiple-sequence alignment showed that the peptidoglycan deacetylase PgdA from S. pneumoniae (SpPgdA) is the closest CE4 enzyme to BsPdaC (Fig. 5). Both proteins share 30% overall sequence iden-

---

**Figure 4. Docking simulation of BsPdaC-CD(GlcNAc)x complexes.** A, binding mode 1 (AAAA). B, binding mode 2 (AAAA), where A is the GlcNAc residue in subsite 0 (catalytic site). Shown are structures of minimum energy from the ensemble of docked structures in each minimum. The C terminal α-helix is not represented. Loops are colored as in Fig. 3. C, protein-ligand interaction from docking simulations. BM, binding mode.
tity with the same domain organization and 44% sequence identity for the CE4 catalytic domain. Although they are both peptidoglycan deacetylases, **SpPgdA** acts on GlcNAc residues (11, 22), whereas **BsPdaC** deacetylates MurNAc residues of peptidoglycan substrates (8). Interestingly, the other two known peptidoglycan MurNAc deacetylases with solved 3D structure, **BsPdaA** (biochemically characterized) (15) and **BaCE4** (annotated) (17), are more distantly clustered in the phylogenetic tree.

**Discussion**

**PdaC is a metalloenzyme that follows a multiple-chain mechanism on COS substrates**

BsPdaC is a metalloenzyme as expected for a CE4 enzyme and here proven by the fact that removal of the metal cation leads to an inactive enzyme whose activity is restored by reconstitution with Zn$^{2+}$/H$^{1001}$ cation. This is in contrast with a previous report (8), which suggested that BsPdaC might not be a metalloenzyme because protein unfolding and refolding was necessary to dissociate the metal cation that is strongly bound to the active site. The enzyme is highly active on intact PGN and active on COS with degree of polymerization $\geq 3$. It follows a multiple-attack mechanism starting with deacetylation of the internal GlcNAc residues and proceeding to deacetylate all but the reducing-end unit of the substrate. This behavior is unique among characterized MurNAc deacetylases, which do not deacetylate COS and prefer a modified PGN substrate in which the peptidyl substitutions have been previously removed (14).

**PdaC is a new subclass of MurNAc deacetylases with different sequence and structural features**

BsPdaC has a different substrate specificity as compared with canonical MurNAc deacetylases, represented by BsPdaA (14, 15). Comparison of the 3D structures unravels significant differences at their active sites. First, the metal coordination differs between these two types of MurNAc deacetylases. BsPdaC maintains the conserved metal-binding triad (Asp-His-His) found in all characterized enzymes, whereas BsPdaA has a different arrangement of metal-binding residues. This difference in metal coordination could account for the observed differences in substrate specificity and catalytic activity between the two enzymes.
chitin deacetylases and peptidoglycan GlcNAc deacetylases (Fig. S11). In contrast, MurNAc deacetylases (structural information currently limited to BsPdaC and BaCE4 (BA0424), Table S1) lack the conserved Asp, which is replaced by an Asn that has a different side-chain orientation pointing away from the metal ion. BsPdaC was initially crystallized as the apoenzyme, but the metal coordination could be solved by soaking experiments with Cd\(^{2+}\), showing that it was only coordinated with two active-site residues, His\(^{124}\) and His\(^{128}\), and the remaining ligands were water molecules (15). BaCE4 was solved with Zn\(^{2+}\) cation, where only the two conserved His coordinated the metal cation, and coordination was completed with an acetate and cacodylate anions from the crystallization mother liquor (17). A structural comparison of the metal coordination in the X-ray structures of BsPdaC, BsPdaA, and SpPgda (as representative of a GlcNAc deacetylase) is shown in Fig. 6.

Second, the structural and sequence alignment of BsPdaC with CE4 family members identifies the loops that shape the binding site cleft and that have been proposed to define the substrate specificity of CDAs (16). BsPdaC reveals short loops, similar in shape and length to other PGN deacetylases, that result in an open binding cleft able to accommodate polymeric substrates and exhibit a multiple-chain mode of action. As a consequence of the nonconserved nature and structural arrangement of these loops surrounding the active site, both MurNAc deacetylase subclasses display different surface electrostatic potential at the active site. BsPdaC has a positively charged binding cleft, as opposed to GlcNAc deacetylases, such as SpPgda, in which it is negatively charged. BsPdaC lies in between, with positive and negative patches in the binding cleft (Fig. S12).

Third, substrate-binding cavities of BsPdaC, SpPgda, and BsPdaA are compared in Fig. 7. The conservation of BsPdaC with SpPgda is notably higher than with BsPdaA, with very similar 3D structures (RMSD between catalytic domain CA atoms of 1.15 Å). The structure and sequence composition of subsite 0 is identical in BsPdaC and SpPgda. This reveals how BsPdaC, similarly to SpPgda, has the capacity to accommodate GlcNAc residues at the catalytic site for deacetylation (subsite 0), confirmed by ligand-docking simulations and deacetylase activity on COS. However, BsPdaC does not deacetylate GlcNAc residues from PGN. Such a pattern of deacetylation would require the binding of MurNAc units at subsite \(-1\), similarly to SpPgda. This has not been detected by docking. This different behavior can be explained by the sequence differences at subsite \(-1\) of both structures (Fig. 7A): BsPdaC Arg\(^{405}\) and Tyr\(^{430}\) are replaced by Lys\(^{395}\) and His\(^{420}\) in SpPgda. Bulkier side chains of BsPdaC hamper the binding of the 3-lactoyl substitution of a MurNAc residue in subsite \(-1\), in the same location, where it is accommodated in SpPgda (Fig. S13). Subsite \(+1\) is also identical in BsPdaC and SpPgda. On the other hand, major structural and sequence differences are observed between BsPdaC and SpPgda at both sides of the active site (Fig. 7B). Remarkably, a negative charge is introduced in negative subsites (Tyr\(^{318}\) to Asp\(^{106}\), respectively) and a positive charge is removed (Arg\(^{315}\) to Phe\(^{103}\), respectively) next to it. BsPdaC and BsPdaA, although both deacetylate MurNAc residues of PGN, differ in the C3 substitution of their substrates, a negatively charged lactoyl group in the PGN substrate of BsPdaA and a peptide-linked lactoyl group in the PGN substrate of BsPdaC. In subsite 0, BsPdaC has Tyr\(^{377}\) as part of the conserved MT3 motif of CE4 enzymes (RXPY). This Tyr is strictly conserved in chitin deacetylases, highly conserved in PGN GlcNAc deacetylases (with few exceptions where it is replaced by Gly), but differs in BsPdaA, which has an Arg residue instead (Fig. 5). The side chain of Arg\(^{166}\) in BsPdaA is oriented toward the active site, and the positive charge of the guanidinyl group may contribute to the stabilization of the free 3-lactoyl group of MurNAc of its PGN substrate (Fig. 7A and Fig. S13), as opposed to BsPdaC, in which the substrate is uncharged.

**Biological functions of MurNAc deacetylation**

Genes encoding for PGN deacetylases are predominantly expressed in Gram-positive bacteria, especially in *Bacillus* species, *Listeria monocytogenes*, and *Streptococcus pneumoniae* (3, 4, 23). Pathogenic *B. cereus* and *B. anthracis* contain 10 polysaccharide deacetylase homologues, from which at least six are PGN deacetylases, involved in virulence and PGN biogenesis during elongation and cell division (13, 24, 25). PGN GlcNAc deacetylases are best characterized for their function in conferring lysozyme resistance to the bacterial cell wall PGN to evade the host innate immune system (3, 4). These deacetylases have been proposed as novel antibacterial targets, but screening studies in the search of selective inhibitors are still scarce (26–28).

MurNAc deacetylation of PGN has been involved in sporulation and germination in different *Bacillus* species participating in muramic \(\delta\)-lactam synthesis (29). PdaA-deficient spores have no muramic \(\delta\)-lactam structure in the spore cortex PGN and cannot germinate (6). These MurNAc deacetylases require previous release of the peptide substitution at the 3-\(\delta\)-lactoyl group of MurNAc units by the action of an 1-\(\delta\)-alanine amidase (CwlD). Further evidence was provided by the fact that introduction of the *pdaA* and *cwlD* genes into *Escherichia coli* cells led to lactam formation in its cell wall peptidoglycan (5). However, BsPdaC is not associated with sporulation, because a *pdaC*-deficient strain was not affected in sporulation and germination, and the only reported phenotype was lysozyme sensitivity (8). The function of PdaC in its native organism is unknown, other than conferring lysozyme resistance to the cell wall that may not be relevant for a soil bacteria such as *B. subtilis*. BsPdaC is encoded by the *yjeA* gene, which is part of the YycFG regulon involved in regulation of cell wall metabolism, but *yjeA* does not appear to play a prominent role in cell autolysis (18).
PGN hydrolysis by host lysozymes is an initial event for triggering the innate immune responses. Besides deacetylation of polymeric PGN to confer lysozyme resistance, other modifications of PGN fragments (muropeptides) may impair the detection by host immune receptors. N-Acetylmuramyl dipeptide (MDP) was shown to be the minimal PGN fragment to directly interact with Nod2 receptors, ultimately resulting in activation of inflammatory response via the NF-κB and mitogen-activated protein kinase signaling pathways (30). Recently, it has been reported that N-deacetylated MDP did not activate NF-κB, thus suggesting that the MurNAc acetyl group is an important feature for recognition by the Nod2 signaling cascade (31). Whether MurNAc deacetylases act only on polymeric PGN or also on PGN fragments after the action of lytic enzymes, including lysozyme, is unknown. BsPdaC is currently the only characterized MurNAc deacetylase active on peptide-substituted PGN chains, yet the minimal size of its PGN substrates is unknown. We hypothesize that PdaC-type MurNAc deacetylase activity might be present in pathogenic bacteria, but not yet identified because of the lack of a clear sequence signature to distinguish GlcNAc and MurNAc deacetylases. This activity may still be hidden in the large number of annotated putative polysaccharide deacetylases, some proposed to be GlcNAc deacetylases but not biochemically characterized. Identification and characterization of homologues of this new subclass of MurNAc deacetylases in other bacteria, especially in pathogenic bacteria, will certainly contribute to unravel the complex degradation-modification–signaling mechanisms in bacteria-host interactions.

**Conclusion**

BsPdaC is a specific MurNAc deacetylase on peptidoglycan, but it also has GlcNAc deacetylase activity on COS. We have determined the X-ray 3D structure of the CE4 catalytic domain of BsPdaC at 1.54 Å resolution and analyzed the mode of action on COS substrates. The catalytic site retains the conserved Asp-His-His metal-binding triad characteristic of CE4 enzymes acting on GlcNAc residues, which differs from other MurNAc deacetylases that lack the metal-coordinating Asp. On COS, the enzyme follows a multiple-chain (or distributive) mechanism that starts at the internal GlcNAc units and further leads to deacetylation of all but the reducing-end GlcNAc residues. BsPdaC is not associated with sporulation, and its biological function, other than conferring lysozyme resistance, remains to be elucidated. We propose that PdaC is the first member of a new subclass of peptidoglycan MurNAc deacetylases based on these differential functional and structural characteristics.

**Materials and methods**

**Cloning of BsPdaC full-length (FL) and catalytic domain (CD) constructs**

A synthetic and codon-optimized gene coding for BsPdaC (amino acid residues 27–467, without signal peptide and Nt-
transmembrane region) (GeneArt® gene synthesis service, Thermo Fisher Scientific) was cloned into a pET22b (+) vector containing a C-terminal Strep-tag sequence (WSHPQFEK). Both the synthetic gene and the vector were digested with the restriction enzymes NdeI and SacI (New England Biolabs) and ligated with T4 DNA ligase (Bio-Rad). From the generated construct (pET22b-BsPdaC-FL), the catalytic domain (from Glu270 to Lys467) was amplified by PCR with primers CD fwd and CD_rvs and cloned into the pET22b (+) vector following the same strategy (list of primers in Table S5). An N terminal Strep-tagged variant of the catalytic domain (BsPdaC-CD-NtStrep) was also prepared. The same sequence corresponding to the catalytic domain (from Glu270 to Lys467) was amplified by PCR with primers CD-Nt_fwd (containing a Kasl restriction site and the tobacco etch virus proteolysis site) and CD-Nt_rvs (containing a Psfl restriction site) and cloned into a pPR-IBA2 vector containing an N-terminal Strep-tag sequence. Both the amplified gene and the vector were digested with the restriction enzymes KasI and Psfl (New England Biolabs) and ligated with T4 DNA ligase (Bio-Rad).

The inactive mutants D285S of all constructs and BsPdaC-CD-W402A and Y430A were generated by site-directed mutagenesis (QuickChange SDM, Agilent) with the mutagenic primers listed in Table S5. The final plasmids were verified by gene sequencing. E. coli BL21(DE3) cells were transformed with each plasmid for protein expression.

Protein expression and purification

E. coli BL21(DE3) cells harboring the pET expression plasmid pET22b-BsPdaC-FL (full-length protein), pET22b-BsPdaC-CD (catalytic domain with Strep-tag at the C terminus), pPR-IBA2-BsPdaC-CD-NtStrep (catalytic domain with Strep-tag at the N terminus) and the inactive mutants pET22b-BsPdaC-CD-D285S and pPR-IBA2-BsPdaC-CD-NtStrep-D285S were grown in an autoinduction medium (32) containing 100 μg/ml ampicillin at 25 °C, 170 rpm for 48 h. The cells were harvested by centrifugation, suspended in PBS buffer (50 mM phosphate, 300 mM NaCl, pH 7) with 1 mM serine protease inhibitor phenylmethylsulfonyl fluoride, and disrupted by sonication. After centrifugation, the protein in the supernatant was purified by affinity chromatography with a Strep Trap column (5 ml; GE Healthcare). The protein was eluted with 2.5 mM d-desethylibiotin in PBS, and, after concentration with an Amicon Ultra-15 centrifugal filter (Millipore), it was loaded onto a Superdex 200 size-exclusion chromatography column (GE Healthcare) and eluted with PBS, displaying different oligomeric forms (Fig. S2). Fraction F1 is composed of soluble oligomers, fraction F2 corresponds to dimers, and fraction F3 is the monomeric enzyme. The relative intensity of the different fractions was slightly different for both BsPdaC-FL and BsPdaC-CD as well as between different batches of the same protein. The monomeric fraction (F3) resulted in being the most active for both BsPdaC-FL and BsPdaC-CD proteins, and it was the one used for kinetic characterization and crystal screening. Analysis of the eluted proteins by SDS-PAGE showed the expected apparent molecular masses, which were assessed by MALDI-TOF MS (52.7 kDa for the FL enzyme and 24.4 kDa for the CD protein). The thermal stability of both the FL and CD proteins were similar, with a melting temperature (T_m) of about 60 °C, as determined by a thermal shift assay (Fig. S4). Protein concentration was determined with the BCA protein assay kit (Thermo Fisher Scientific).

Thermal stability

Purified protein (BsPdaC-FL or BsPdaC-CD; 4 μM) in PBS (50 mM phosphate, 300 mM NaCl, pH 7) was mixed with the commercial dye Sypro Orange (Thermo Fisher Scientific; 1:5000 dilution according to the manufacturer’s protocol) in a final volume of 25 μl. The samples were subjected to a thermal gradient in a thermocycler (Rotogene 3000, Corbett Research), consisting of 1 min at 25 °C followed by 1 °C increments (30 s at each temperature) up to 95 °C. The fluorescence was measured (λ_ex = 483 nm, λ_em = 560 nm) and T_m was determined by fitting the data to a Boltzmann sigmoidal equation using Prism (GraphPad Software, La Jolla, CA).

Enzyme activity of gel filtration fractions

All fractions collected from the size-exclusion chromatography were evaluated with 4-methylumbelliferyl acetate (AcOMU) as nonspecific esterase substrate. 10 μM acetate/min/μM enzyme was determined by monitoring the fluorescence increase (λ_ex = 340 nm, λ_em = 460 nm) due to 4-methylumbelliferone (MU) release and expressed as initial rates (μM/s) using MU standards in the same buffer.

Enzyme activity on peptidoglycan

Enzyme activity toward B. subtilis peptidoglycan (from Sigma-Aldrich) was determined by monitoring the release of acetic acid with a commercial acetate determination kit (Acetic Acid Assay Kit, Megazyme). Reactions were performed by incubating the peptidoglycan substrate (1 mg/ml) with the active enzymes (BsPdaC-FL or BsPdaC-CD; 0.25–5 μM) in PBS (50 mM phosphate, 300 mM NaCl) buffer, pH 7.0. Enzyme activity was determined by monitoring the fluorescence increase (λ_ex = 340 nm, λ_em = 460 nm) due to 4-methylumbelliferone (MU) release and expressed as initial rates (μM/s) using MU standards in the same buffer.

To assess that BsPdaC-CD maintains the same MurNAc deacetylase activity as the full-length enzyme, end-point reactions with both BsPdaC-CD and BsPdaC-FL (0.5 μM enzyme, 1 mg/ml PGN in PBS buffer, pH 7.0, 37 °C for 48 h with agitation) were set up, and the release of acetic acid was determined as described above. After dialysis against PBS buffer to remove the free acetate from the samples, BsPdaC-CD was added to the sample of the initial BsPdaC-FL reaction, and BsPdaC-FL was added to the sample of the initial BsPdaC-CD reaction. After a 48-h incubation at 37 °C, the released acetate was determined. Finally, the GlcNAc peptidoglycan deacetylase SpPgdA (0.8 μM) was added to the final reactions to monitor acetate release. As reference, the activity of SpPgdA on B. subtilis PGN was determined.
**Enzyme kinetics by the fluorescamine assay**

Enzyme activity was determined by fluorescamine labeling of the amino groups generated in the deacetylation reaction using COS substrates (11). All reactions were performed by incubating (GlcNAc)ₙ, n = 4 or 5, substrate (2 mM), and purified BsPdaC-CD (0.15–1.3 μM) in PBS (50 mM phosphate, 300 mM NaCl), pH 7, at 37 °C in a final volume of 200 μl in microtiter plates. For determination of the pH optimum, the pH was varied from 6.0 to 10.6 at 37 °C. For optimum temperature determination, the temperature was varied from 25 to 75 °C at pH 7.0 (Fig. S3). At different time intervals, aliquots (20 μl) were withdrawn and added to 90 μl of water/1-propanol (1:1) to stop the reaction. Then 20 μl of 2 mg/ml fluorescamine in DMF were added, and the reaction was incubated for 10 min at room temperature. The labeling reaction was terminated by adding 150 μl of DMF/water (1:1), and fluorescence was quantified (λₑx = 340 nm, λₑm = 460 nm). A set of glucosamine standards were treated following the same protocol.

**Enzyme kinetics by HPLC-MS**

The formation of products with different degrees of acetylation from (GlcNAc)ₙ, n = 3–5, substrates was monitored by HPLC-MS (16) (Agilent 1260 HPLC-MS, electrospray ionization (ESI+), single-quadrupole MS detector) using an XBridge BEH Amide 2.5-μm, 3.0 × 100-mm XP column (Waters) in combination with an XBridge BEH Amide Guard Cartridge (2PK) precolumn (2.5 μm, 4.6 × 20 mm; Waters). Reactions were performed by incubating the (GlcNAc)ₙ substrate (2 mM) and purified BsPdaC-CD (5 μM) in PBS, pH 7.0, at 37 °C and stopped at different time intervals as described above. 5-μl samples were injected into the system and eluted at 60 °C with acetonitrile/water (65:35, v/v), 1% formic acid at a flow rate of 0.4 μl/min. MS detection was performed on both SIM mode (for monitoring [M + H]⁺ of substrate and deacetylation products) and SCAN mode (for total ion monitoring. 250–1100 m/z scan range). The m/z values monitored in SIM mode were as follows (A for GlcNAc, D for GlcNH₂): (a) reactions with tetraacetylchitotetraose (A4) substrate: 831.4 (A4), 789.4 (A3D1), 747.4 (A2D2), and 353.3 (A1D3); (b) reactions with pentaacetylchitopentaose (A5) substrate: 1035.4 (A5), 992.4 (A4D1), 950.4 (A3D2), 908.4 (A2D3), and 867.3 (A1D4). Specific activities for the initial monodeacetylation reactions with (GlcNAc)ₙ substrates were performed as above, using the corresponding monodeacetylated product standards for quantification (prepared by reaction of the corresponding (GlcNAc)ₙ substrate with Vibrio cholera CDA as reported (16)). Michaelis–Menten parameters were determined for the initial monodeacetylation reaction with (GlcNAc)₅, varying the substrate concentration from 0.5 to 17.5 mM with 3.7 μM enzyme in PBS buffer, pH 7.0, 37 °C.

**Determination of the deacetylation pattern**

To determine the pattern of acetylation of the different products formed during the time course monitoring, preparative reactions with 2 mM (GlcNAc)₄ and (GlcNAc)₅ substrates and 0.1 mg of purified BsPdaC-CD in PBS, pH 7.0, at 37 °C, in a final volume of 1 ml, were incubated for different reaction times. The freeze-dried samples were analyzed for pattern of acetylation by MS (MALDI-TOF-MS/MS analysis) after previous ¹⁸O labeling and re-N-acetylation with [¹⁸H₆]acetic anhydride according to the protocol by Cord-Landwehr et al. (20), and the relative amount of each sequence in the mixtures was calculated based on the peak intensities compared for each ion type.

**Effect of metal ions and EDTA**

Purified protein (BsPdaC-CD) in PBS (50 mM phosphate, 300 mM NaCl, pH 7.0) was dialyzed against PBS with EDTA (200 mM) and urea (2 or 7 M). During the refolding step, EDTA and urea were removed by dialysis (first with PBS and then with water). Activity of the treated proteins and the native enzyme with and without the addition of metal ions (Zn²⁺) was determined by HPLC-MS. The metal content of the apoenzyme (after the refolding step from 7 M urea) was [Zn²⁺] (μM)/[E (μM)] = 0.1 by inductively coupled plasma MS (NexION 300XX, PerkinElmer Life Sciences), meaning that ~90% of the bound metal was removed following this procedure.

**Effect of ligands on thermal stability**

A J-810 CD spectropolarimeter (Jasco Corp., Tokyo, Japan) equipped with a Peltier thermal device was used. Measurements were carried out in a range from 20 to 90 °C at 222 nm by using quartz cuvettes with a 1-mm optical path length. Temperature was increased stepwise at 1 °C/min. Samples were 10 μM BsPdaC-CD-NtStrep-tag in 50 mM Tris buffer, 150 mM NaF, pH 7.5. The effect of substrate on the thermal stability was tested under the same conditions, using 1 mM (GlcNAc)₂, (GlcNAc)₄, and (GlcNAc)₅. Transitions were fitted according to a Boltzmann sigmoidal equation using the GraphPad software.

**Crystallographic structure**

BsPdaC-CD crystallization and data collection—Crystals of BsPdaC-CD (WT and D285S mutant) were obtained by mixing 0.25 μl of the protein in 50 mM Tris, pH 7.0, and 150 mM NaCl at 5.6 or 10 mg ml⁻¹, respectively, with 0.25 μl of a mother liquor containing 0.2 M ammonium tartrate dibasic, pH 6.6, 20% (w/v) PEG 3.350 or 0.4 M ammonium phosphate monobasic by sitting-drop vapor-diffusion crystallization. Crystals appeared after 8–10 days and grew as prisms, reaching 0.1–0.8 × 0.5–0.4 × 0.05–0.07 mm. Single crystals of BsPdaC-CD and BsPdaC-CD-D285S were cryocooled in liquid nitrogen by using a cryoprotectant solution of the mother liquor supplemented with 15% (v/v) ethylene glycerol or 30% (v/v) glycerol, respectively. X-ray diffraction data were collected on a PILATUS 6M F pixel array detector at the microfocus I02 beamline (λ = 0.9796 Å, Diamond Light Source) and integrated with XDS (33) following standard procedures.

Structure determination and refinement—Structures of BsPdaC-CD and BsPdaC-CD-D285S were solved using as a template the previously reported S. pneumoniae peptidoglycan deacetylase SpPgdA (PDB code 2C1G) and molecular replacement methods implemented in Phaser (34) and the PHENIX suite (35). Model rebuilding was carried out with Buccaneer (36) and the CCP4 suite (37). The final manual building was performed with Coot (38) and refinement with phenix.refine (35). The structures were validated by MolProbity (39). Data
Structure and specificity of BsPdaC

collection and refinement statistics are presented in Table S4. Atomic coordinates and structure factors have been deposited with the Protein Data Bank, accession codes 6H8L and 6H8N corresponding to BsPdaC-CD and BsPdaC-CD-D285S crystal structures, respectively. Molecular graphics and structural analyses were performed with the UCSF Chimera package (40).

Modeling of BsPdaC, BsPdaA, and SpPgdA structures in complex with substrates

The 3D structures of the enzyme-ligand complexes of SpPgdA, BsPdaA, and BsPdaC were modeled with AUTODOCK 4.2 and AUTODOCK VINA (41, 42). Protein structures were directly taken from the Protein Data Bank with PDB accession codes 2C1G (SpPgdA), 1W1B (BsPdaA), and 6H8L (BsPdaC, this work). Ligand structures covering the different range of accepted substrates were extracted from other protein structures with PDB accession codes 6AVE (GlcNAc), ITWQ (MurNAc), 9LYZ (GlcNAcβ(1→4)MurNAc and MurNAc β(1→4)GlcNAc), and 1LZC (tetraacetylchitotetraose). Both the protein and ligand structures were first parametrized with AutoDockTools (41): polar hydrogens were added, AutoDock4.2 atom typing was used, and Gaisteger partial charges were computed, with randomized initial position and orientation coordinates and randomized conformations of the substrate flexible bond. The genetic algorithm implemented in AUTODOCK 4.2 were performed. For each round, an initial population of 150 members was considered, with randomized initial position and orientation coordinates and randomized conformations of the substrate flexible bond. The genetic algorithm was extended up to 27,000 off-spring generations, with a maximum of 2,500,000 energy evaluations. Only low-energy and repetitive binding poses were considered for analysis.

Sequence alignment and phylogenetic analysis of CE4 enzymes

The sequence of BsPdaC catalytic domain (CE4 domain) was retrieved from Uniprot (accession code O34798, aa 278 – 456) and incorporated into the multiple-sequence alignment guided by the structural superimposition of characterized CE4 enzymes (12, 16) by hidden Markov model comparisons using HMMER (44). BsPdaC was analyzed together with the following sequences (shown as name (organism, Uniprot accession code)) (Fig. S11): (a) chitin deacetylases: MrCDA (Mucor rouxii, P50325), CicCDA (Colletotrichum lindenmuthianum, Q6DWK3), AnrCDA (Aspergillus nidulans, Q5AQ0Q), PesCDA (Pestalotiopsis sp., A0A1L3TH9R9), PaCDA (Podospora anserine, B2AAQ0), PgtCDA (Puccinia graminis, E3K3D7), VcCDA (V. cholera, Q9KSH6), VpCDA (Vibrio parahaemolyticus, A6P415) RmNodB (Rhizobium meliloti, P02963), ArCE4 (Artrobacter sp., A0A2C8C11T7); (b) peptidoglycan deacyt- lases: SpPgdA (S. pneumoniae, Q8D6P3), SmPgd (Streptococcus mutants, Q8DV82), BcPgd (B. cereus, Q81EK9), ErPgd (Eubacterium rectale, C4ZEZ9), BsPdaA (B. subtilis, Q34928), BaCE4 (B. anthracis, Q81Z49); (c) unknown: BC0361 (B. cereus, Q81N13), BA0330 (B. anthracis, Q81ZDM9), BA0150 (B. anthracis, Q81VP2), ECU11_0510 (Encephalitozoon cuniculi, Q85US6), (d) acetyblycan esterases: SfAxeA (Streptomyces lividans, Q54413), CtxAxeA (Clostridium thermocellum, O87119); and (e) poly-β-1,6-GlcNAc deacytelases: EcPgaB (E. coli, a P75906), AdlcB (Ammonifex degensii, C9RCK9), BbPsbB (Bordetella bronchiopneumotica, A0A0C6P1R7), AarPgaB (Agrupagibacter actinomycetomatis, A5HJW8).

The phylogenetic relationships were inferred by using the maximum likelihood method based on the JTT matrix-based model (45) computed from the multiple-sequence alignment. Bootstrap analysis consisted of 500 replicates. These evolutionary analyses were conducted with MEGA7 (46), and the output dendrogram is shown in Fig. 5.


Acknowledgments—We gratefully acknowledge Diamond Light Source (proposal mx15304), ALBA, and iNEXT (proposal numbers 1618/2538) for granting access to synchrotron radiation facilities. We thank Dr. Hugo Aragunde for the sequencing of deacetylated COS and Prof Bruno Moerschbacher and Dr. Stefan Cord-Landwehr (University of Münster) for advice on MS/MS sequencing analysis. Montserrat Murcia is acknowledged for performing some of the docking simulations.

References
Structure and specificity of BsPdaC


