The Structure- and Metal-dependent Activity of *Escherichia coli* PgaB Provides Insight into the Partial De-*N*-acetylation of Poly-*β*-1,6-*N*-acetyl-D-glucosamine*[S]

Dustin J. Little[‡§1], **Joanna Poloczek**[¶1], **John C. Whitney**[‡§2], **Howard Robinson**[∥], **Mark Nitz**[¶3], **and P. Lynne Howell**[‡§4]

From the ‡*Program in Molecular Structure & Function, Research Institute, The Hospital for Sick Children, Toronto, Ontario M5G 1X8, Canada,* §*Department of Biochemistry, University of Toronto, Toronto, Ontario M5S 1A8, Canada,* ¶*Department of Chemistry, University of Toronto, Toronto, Ontario M5S 3H6, Canada, and* ∥*Biology Department, Brookhaven National Laboratory, Upton, New York 11973-5000*

Background: Polysaccharide intercellular adhesin-dependent biofilm formation in *E. coli* requires the de-*N*-acetylation of poly-*β*-1,6-*N*-acetyl-D-glucosamine by PgaB.
Results: Nickel- and iron-bound structures of PgaB have been determined, and the metal-dependent de-*N*-acetylase activity of the enzyme has been characterized.
Conclusion: PgaB has low catalytic efficiency and shows preference for Co^{2+}, Ni^{2+}, and Fe^{2+} ions.
Significance: The structure of PgaB will guide inhibitor design to combat biofilm formation.

Exopolysaccharides are required for the development and integrity of biofilms produced by a wide variety of bacteria. In *Escherichia coli*, partial de-*N*-acetylation of the exopolysaccharide poly-*β*-1,6-*N*-acetyl-D-glucosamine (PNAG) by the periplasmic protein PgaB is required for polysaccharide intercellular adhesin-dependent biofilm formation. To understand the molecular basis for PNAG de-*N*-acetylation, the structure of PgaB in complex with Ni^{2+} and Fe^{3+} have been determined to 1.9 and 2.1 Å resolution, respectively, and its activity on *β*-1,6-GlcNAc oligomers has been characterized. The structure of PgaB reveals two $(β/α)_x$ barrel domains: a metal-binding de-*N*-acetylase that is a member of the family 4 carbohydrate esterases (CE4s) and a domain structurally similar to glycoside hydrolases. PgaB displays de-*N*-acetylase activity on *β*-1,6-GlcNAc oligomers but not on the *β*-1,4-(GlcNAc)$_4$ oligomer chitotetraose and is the first CE4 member to exhibit this substrate specificity. De-*N*-acetylation occurs in a length-dependent manor, and specificity is observed for the position of de-*N*-acetylation. A key aspartic acid involved in de-*N*-acetylation, normally seen in other CE4s, is missing in PgaB, suggesting that the activity of PgaB is attenuated to maintain the low levels of de-*N*-acetylation of PNAG observed *in vivo*. The metal dependence of PgaB is different from most CE4s, because PgaB shows increased rates of de-*N*-acetylation with Co^{2+} and Ni^{2+} under aerobic conditions, and Co^{2+}, Ni^{2+} and Fe^{2+} under anaerobic conditions, but decreased activity with Zn^{2+}. The work presented herein will guide inhibitor design to combat biofilm formation by *E. coli* and potentially a wide range of medically relevant bacteria producing polysaccharide intercellular adhesin-dependent biofilms.

Bacterial biofilms are complex surface attached microbial communities that are a significant medical problem because once established they are difficult to eradicate (1, 2). A key feature of bacterial biofilms is that cellular aggregates grow within a self-produced extracellular matrix (2, 3). The extracellular matrix facilitates the adherence between bacteria and the surfaces they colonize while also providing protection against environmental stresses, a diffusion barrier against antibiotics, and isolation from the innate immune system (4, 5). The composition of the extracellular matrix can be quite diverse, but exopolysaccharides are a key component for the development and structural integrity of most biofilms (6, 7). A wide variety of medically important biofilm-forming bacteria produce partially de-*N*-acetylated poly-*β*-1,6-*N*-acetyl-D-glucosamine (dPNAG)[5] exopolysaccharides, also referred to as polysaccharide intercellular adhesin. dPNAG was first described in *Staphylococcus epidermidis* (8) but has now been determined to be a component of the biofilm matrices of *Staphylococcus aureus* (9), *Escherichia coli* (10), *Acinetobacter baumannii* (11),

---

Bordetella bronchiseptica (12), *Bordetella pertussis* (13), *Actinobacillus pleuropneumoniae* (14), *Yersinia pestis* (15), and *Burkholderia* species (16).

In *E. coli*, dPNAG production is dependent on the four-gene operon, *pgaABCD* (10, 17). PgaC, which is predicted to contain multiple transmembrane domains and a large cytosolic domain homologous to family 2 glycosyltransferases, is thought to be responsible for the production of PNAG and its transportation across the inner membrane (17). PgaD is a small integral inner membrane protein required for PNAG biosynthesis (17). The exact function of PgaD is unknown, but it has been predicted to assist PgaC in polymerization and transport of PNAG across the inner membrane (17) and in the regulation of PNAG production, because levels of PgaD are controlled post-transcriptionally by cellular concentrations of the bacterial second messenger bis-(3′-5′)-cyclic dimeric GMP (18). PgaB is a two-domain outer membrane protein that contains a putative lipidation site and is responsible for the partial de-*N*-acetylation of PNAG (17). Partial de-*N*-acetylated PNAG, or dPNAG, is the functionally relevant form of the exopolysaccharide because deletion of *pgaB* results in the retention of the fully acetylated polymer in the periplasm and abolishment of biofilm formation (10, 17). The N-terminal domain of PgaB has been classified as belonging to the family 4 carbohydrate esterases (CE4s) (19). CE4s are known for their metal-dependent deacetylation of *O*- and *N*-acetylated polysaccharides such as chitin, peptidoglycan, and acetylxylan (20). Characterization of several CE4s has shown that deacetylation is dependent on a variety of divalent metal ions (21–26), and it is proposed that the deacetylation reaction proceeds through a metal stabilized tetrahedral oxyanion intermediate (23–25). The exact function of the C-terminal domain of PgaB is unknown; however, it is required for efficient de-*N*-acetylation of PNAG *in vivo* (17). The domain shows sequence homology to COG1649 (27), a domain common to TIM barrel glycoside hydrolases (28), and has been suggested to play a role in PNAG binding and/or hydrolysis (17). PgaA is a two-domain outer membrane protein that is predicted to contain a C-terminal 16-stranded β-barrel domain that likely acts as the porin for exporting dPNAG, and an N-terminal periplasmic domain that carries multiple copies of the tetratricopeptide repeat motif that has been implicated in protein-protein and protein-ligand interactions (17, 29–35). PgaA has been suggested to play a role as a periplasmic scaffold that forms a complex with PgaB to facilitate transport of dPNAG across the outer membrane (17, 30). This is a hypothesis that is supported by the interaction between HmsF and HmsH, the PgaA and PgaB homologues in *Y. pestis*, which has been observed *in vivo* (36).

Here we present the first structural and biochemical characterization of a PNAG de-*N*-acetylase from the CE4 family. The structure of *E. coli* PgaB reveals a unique arrangement of two $(\beta/\alpha)_x$ domains. The N-terminal de-*N*-acetylase domain contains an Asp-His-His metal coordination site and exhibits different metal-dependent activity compared with other Asp-His-His containing CE4s. The active site metal of PgaB is tightly coordinated in an octahedral arrangement, and the identity of the metal depends on expression conditions. Structures of PgaB have been solved in complex with both Ni^{2+} and Fe^{3+}. Structural comparison with other CE4 members reveals that PgaB contains a unique arrangement of the conserved CE4 motifs, is missing a key conserved aspartate residue involved in de-*N*-acetylation, and has a deep active site binding pocket. These data may explain its specificity but poor catalytic efficiency toward β-1,6-GlcNAc oligomers. Together the data provide insight into the molecular basis for de-*N*-acetylation of PNAG and will guide the development of inhibitors against PgaB to prevent biofilm formation by *E. coli*.

EXPERIMENTAL PROCEDURES

Cloning, Expression, and Purification of MBP-PgaB—To generate the maltose-binding protein (MBP) fusion expression vector, a codon-optimized sequence for PgaB (synthesized by Bio Basics Inc.) lacking its predicted signal sequence was ligated into the pET16b vector (Novagen) between its NdeI and BamHI restriction sites to yield plasmid UT031. The gene was subsequently excised from UT031 and inserted into the pIADL plasmid (37) to generate pIADL-PgaB. The resulting plasmid encodes an N-terminal maltose-binding protein fused to PgaB[22–672] via a 25-amino acid linker and will be referred to as MBP-PgaB. *E. coli* BL21 (DE3) cells transformed with pIADL-PgaB were grown in l liter of LB media with 100 µg/ml kanamycin at 37 °C to an A_{600} of ~0.5–0.6, and protein expression was induced by the addition of isopropyl-D-1-thiogalactopyranoside to a final concentration of 0.5 mM. The cells were incubated post-induction overnight at 10 °C and then harvested by centrifugation at $3750 \times g$ for 75 min. Nickel- or iron-loaded enzyme was prepared by adding $NiSO_4$ or Fe(III) citrate to a final concentration of 1 mM to the culture 30 min before induction. Cell pellets were subsequently resuspended in 15 ml of 50 mM HEPES, pH 7.5, supplemented with 300 mM NaCl and one complete mini protease inhibitor mixture tablet (Roche Applied Science). Resuspended cells were lysed by sonication, and cell debris was removed by centrifugation for 1 h at $16,000 \times g$. The resulting supernatant was passed over a column containing 1.5 ml of amylose resin (Sigma-Aldrich) per liter of culture that was pre-equilibrated with 50 mM HEPES, pH 7.5. The column was washed with 10 column volumes of 50 mM HEPES, pH 7.5, and eluted with a 2–8 mM maltose gradient. The fractions containing MBP-PgaB were concentrated using an ultrafiltration device (Millipore) and stored at 4 °C. The purified MBP-PgaB was >95% pure as judged by SDS-PAGE and stable for ~2 weeks. Anerobically isolated MBP-PgaB was purified as described above except in a glove box under a N_2 atmosphere using degassed buffers.

Expression and Purification of His6-PgaB—Native and selenomethionyl (SeMet)-incorporated PgaB[42–655] for structural studies were expressed and purified as described previously (38) with the following modifications: for iron- and nickel-loaded PgaB, the bacterial cultures were grown in 1 liter of M9 minimal media (Bio Basics Inc.), pH 7.4, supplemented with 0.4% (w/v) glucose, 0.1 mM $CaCl_2$, 2 mM $MgSO_4$, and 1 mM $NiSO_4$ or $FeSO_4$ was added to the culture 30 min before induction with isopropyl-D-1-thiogalactopyranoside.

Crystallization, Data Collection, and Structure Solution—For crystallization, purified PgaB[42–655] was concentrated to 15 mg/ml at $3000 \times g$ in an ultrafiltration device (Millipore). Crys-

tals were grown using hanging drop vapor diffusion in 48-well VDX plates (Hampton Research) by streak seeding a 4-μl drop with equal amounts of protein and precipitant equilibrated against 200 μl of precipitant solution (10–16% (w/v) PEG 8000, 0.2 M calcium acetate, 0.1 M MES, pH 6.0). The crystals were cryoprotected for 10 s in reservoir solution supplemented with 10% (v/v) glycerol and 10% (v/v) ethylene glycol prior to vitrification in liquid nitrogen. Native and SeMet single-wavelength anomalous diffraction data were collected on nickel- and iron-loaded crystals, respectively, on Beamline X29 of the National Synchrotron Light Source (see Table 1). The data were indexed and scaled using HKL2000 (39). The SeMet single-wavelength anomalous diffraction data were used in conjunction with HKL2MAP (40) to locate 24 (of 26) selenium sites, and density modified phases were calculated using SOLVE/RESOLVE (41). The resulting electron density map was of high quality and enabled PHENIX AutoBuild (42) to build ~90% of the protein. The remaining residues were built manually in COOT (43). Refinement was carried out using PHENIX.REFINE (42). The structure of the nickel-loaded protein was determined using molecular replacement with PHENIX AutoMR (42). Manual model building and refinement was carried out as described above for the iron-loaded SeMet protein. Structure figures were generated using the PyMOL molecular graphics system (DeLano Scientific), quantitative electrostatics were calculated using the PDB2PQR (44, 45) and APBS (46) software packages, and the topology diagram was constructed using TOPDRAW (47).

Preparation of Substrate—Oligomers of β-1,6-GlcNAc were prepared and purified as outlined previously (48). Oligosaccharide identities were confirmed by MALDI-MS and ^{1}H NMR spectroscopy. Oligosaccharide mass was analyzed on a Waters MALDI micro MX in positive reflectron mode. The samples were deposited on a stainless steel plate with a matrix containing a combination of α-cyano-4-hydroxycinnamic acid and 3-aminoquinoline. Oligosaccharides were stored as lyophilized powders at room temperature and dissolved with deionized water or buffer as required.

[1]*H NMR Analysis of Enzyme Activity*—Spectra were acquired on a 400-MHz Varian spectrometer. NMR samples were deuterium-exchanged by lyophilizing β-1,6-(GlcNAc)$_5$ and resuspending in D$_2$O three times. Similarly, 500 μl of 50 mM sodium phosphate buffer, pH 8.0, was deuterium-exchanged and was used to dissolve the deuterium-exchanged β-1,6-(GlcNAc)$_5$ to give a final concentration of ~10 mM. MBP-PgaB in 50 mM phosphate buffer, pH 8.0, was then added to the substrate solution to a final concentration of 10 μM. The ^{1}H NMR spectrum was acquired immediately after the addition of the enzyme and again after incubation at 37 °C for 24 h.

Fluorescamine Enzyme Activity Assays—Rates of de-*N*-acetylation were determined using a fluorescamine-based assay (49). Briefly, each sample was prepared in duplicate to a final volume of 10 μl in 50 mM HEPES, pH 7.5, with 10 μM MBP-PgaB and 10 mM substrate (unless otherwise specified). To correct for the presence of free amines on the enzyme, MBP-PgaB was added to control samples at a final concentration of 10 μM just prior to reaction with fluorescamine. The reaction was performed by adding to each 10-μl sample, 20 μl of 0.5 M borate

buffer, pH 9.0, and 10 μl of a freshly prepared 20 mg/ml fluorescamine solution in dimethylformamide and mixed by pipetting. The reaction was then allowed to stand at room temperature for ~10 min before adding 80 μl of deionized water. An 80-μl aliquot of the solution was removed from each sample and transferred to a Corning 3686 half-area 96-well plate for measurement in a TECAN Safire2 plate reader (360-nm excitation, 460-nm emission, 5-nm slit widths). Rate of substrate de-*N*-acetylation was determined at 37 °C, and a minimum of four time points were taken at intervals so that the de-*N*-acetylation did not exceed 30% of the substrate. Glucosamine solutions were used as standards to calculate amine concentration. Metal-dependent PgaB activity assays were conducted with molar equivalents of metal chloride or sulfate salt solutions, added to the samples prior to substrate addition. PgaB length-dependent β-1,6-GlcNAc oligomer activity assays contained 10 μM nickel-loaded MBP-PgaB, 200 mM NaCl, and 50 mM HEPES buffer, pH 7.5, and were incubated at 37 °C for 20 h. Less than 30% substrate conversion occurred over the incubation. β-1,6-GlcNAc oligosaccharide concentrations were determined using ^{1}H NMR and dimethylformamide internal standards.

Metal Analysis—Inductively coupled plasma-atomic emission spectroscopy (ICP-AES) was carried out at the Analytical Lab for Environmental Research and Training at the University of Toronto. ICP-AES samples were prepared by diluting concentrated enzyme samples to 50 μM in 5 mM phosphate buffer, pH 7.5. Control samples were prepared the same way but using 50 mM HEPES, pH 7.5, and were subtracted from the signals acquired from the enzyme samples during ICP-AES analysis.

Determination of De-N-acetylation Position—The position of de-*N*-acetylation within β-1,6-GlcNAc oligomers was determined using a protocol adapted from Withers and co-workers (50). β-1,6-(GlcNAc)$_4$ (4.2 mg) and β-1,6-(GlcNAc)$_5$ (5.7 mg) were treated with 10 μM nickel-loaded PgaB in 5 mM HEPES buffer, pH 7.5, at 37 °C for 48 h. The samples were centrifuged at 13,000 × g for 10 min, before adding 25 mM phosphate and citrate buffer, pH 4.5, giving a substrate concentration of 5 mM. Purified *Streptomyces plicatus* hexosaminidase (*Sp*Hex) was added to a final concentration of 10 μM, and the solutions were incubated for 24 h at 37 °C. The resulting solutions were lyophilized, and β-1,6-GlcNAc oligosaccharides were treated with Ac$_2$O in MeOH for 2 h to reacetylate glucosamine residues before being separated by length using HPLC on a Grace Prevail Carbohydrate ES column (4.6 × 250 mm, 5 μm) using an acetonitrile/H$_2$O gradient from 80/20 to 50/50 over 50 min at 0.5 ml/min and were identified by retention time standards and MALDI-MS. For analysis by electrospray ionization MS, to aid in fragment deconvolution of multiply de-*N*-acetylated species, the PgaB-treated substrates were reduced with NaBH$_4$ followed by neutralization with acetic acid and subsequent dialysis against double distilled H$_2$O using a 100-Da molecular mass cutoff membrane (Spectrapor). Tandem MS was performed on the samples in positive mode using an AB Sciex QSTAR instrument with a TurboIonSpray source.

TABLE 1
Summary of data collection and refinement statistics
The values in parentheses correspond to the highest resolution shell.

	SeMet-Iron complex	Native-Nickel complex
Data collection		
Beamline	NSLS X29	NSLS X29
Wavelength (Å)	0.979	1.075
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters (Å)	$a = 91.1$, $b = 102.4$, $c = 150.9$	$a = 90.7$, $b = 102.8$, $c = 150.9$
Resolution (Å)	50.00–2.10 (2.18–2.10)	50.00–1.90 (1.93–1.90)
Total no. of reflections	1,108,638	804,755
No. of unique reflections	82,493	111,639
Redundancy	13.5 (12.6)	7.2 (7.1)
Completeness (%)	99.9 (99.3)	100.0 (100.0)
Average $I/\sigma(I)$	7.5 (2.8)	24.1 (4.2)
R_{merge} (%)[a]	11.7 (59.7)	7.8 (49.1)
Refinement		
R_{work}/R_{free}[b]	16.5/19.8	16.8/19.3
No. of atoms		
Protein	9528	9520
Acetate	8	8
MES	24	24
Calcium	2	2
Iron	2	
Nickel		2
Water	661	1009
Average B-factors (Å²)[c]		
Protein	37.9	37.6
Acetate	26.6	25.2
MES	43.1	38.7
Calcium	30.3	27.7
Iron	21.2	
Nickel		21.0
Water	35.9	38.4
Root mean square deviations		
Bond lengths (Å)	0.007	0.007
Bond angles (°)	1.04	1.03
Ramachandran plot[c]		
Total favored (%)	98.0	98.2
Total allowed (%)	99.8	99.9
Coordinate error (Å)[d]	0.20	0.17
Protein Data Bank code	4F9J	4F9D

[a] $R_{merge} = \sum\sum |I(k) - <I>|/\sum I(k)$, where $I(k)$ and $<I>$ represent the diffraction intensity values of the individual measurements and the corresponding mean values. The summation is over all unique measurements.

[b] $R_{work} = \sum ||F_{obs}| - k|F_{calc}||/|F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively. R_{free} is the sum extended over a subset of reflections (5%) excluded from all stages of the refinement.

[c] As calculated using MolProbity (73).

[d] Maximum likelihood-based coordinate error, as determined by PHENIX (42).

RESULTS

PgaB Has a Unique Two-domain Structure—A stable core of PgaB, encompassing residues 42–655 was expressed, purified, and crystallized (38). Diffraction data were collected to 1.9 and 2.1 Å on native and SeMet protein crystals, respectively, and the structure was solved using the SeMet single-wavelength anomalous diffraction technique (51). Refinement produced final models with good geometry and R factors of 16.8% (R_{free} = 19.3%) and 16.5% (R_{free} = 19.8%) for the native and SeMet derivative, respectively (Table 1). Residues 42, 62, 610–620, and 647–655 for the native structure, and 42, 61–63, 612–616, and 650–655 for the SeMet structure, were not included in the final models because there was no interpretable electron density present for these residues. PgaB crystallized in the orthorhombic space group $P2_12_12_1$ with two molecules in the asymmetric unit (asu). The two molecules are very similar, with root mean square deviations (RMSDs) of 0.22 Å over 584 equivalent Cα atoms. Although two molecules of PgaB are present in the asu, the protein elutes from a calibrated analytical size exclusion column (data not shown) at a molecular weight that is consistent with PgaB being a monomer in solution.

The structure of PgaB consists of two domains that both adopt $(\beta/\alpha)_x$ barrel folds (Fig. 1, *A* and *B*). The domains are arranged in tandem with the C-terminal domain tilted ~45° away from the N-terminal domain. The N-terminal domain (residues 42–310) has an even distribution of surface charge and has a large pocket at the top of the distorted $(\beta/\alpha)_7$ barrel (Fig. 1C). Structural alignment of the N-terminal domain with the CE4 peptidoglycan deacetylase PgdA from *Streptococcus pneumoniae* (*Sp*PgdA) (25) shows strong conservation of both the canonical $(\beta/\alpha)_7$ fold and the active site, with RMSDs of 2.6 Å (over 141 equivalent Cα atoms) and 0.6 Å (over 11 equivalent Cα atoms), respectively (Fig. 2, *A* and *B*). However, topological differences occur between PgaB and other CE4 structures. Although not described originally in the NodB homology domain (20), CE4s often contain a "capping" helix that seals the bottom of the $(\beta/\alpha)_7$ barrel (52) and are generally encoded at the beginning or end of the sequence. In PgaB, this helix (α2 colored *purple* in Fig. 1, *A* and *B*) occurs internally in the sequence when compared with previously reported structures (21–26, 52). Another difference for PgaB is the presence of four β-hairpins, β2–3, β6–7, β9–10, and β11–12 (Figs. 1, *A* and *B*, and 2C). These structural elements appear to be unique to PgaB because they are not present in any of the other CE4 structures determined to date (Fig. 2C). Hairpins β6–7, β9–10, and β11–12 carry a number of surface-exposed hydrophobic residues (Phe-154, Tyr-190, Ile-192, Leu-200, Tyr-209, and Tyr-216) that lie along α7 (supplemental Fig. S1). These residues along with Trp-387 (from α12) form a hydrophobic groove leading into the de-*N*-acetylase active site (supplemental Fig. S1).

The C-terminal domain (residues 311–655) has a $(\beta/\alpha)_8$ TIM barrel fold. Because the C terminus of PgaB was truncated to facilitate crystallization, the eighth helix of this fold, α26, is missing. Analysis of the crystal packing reveals that α1 from molecule two in the asu packs tightly against β23 from molecule one, where α26 would be located. This suggests the PgaB dimer seen in the asu is most likely an artifact of crystallization because of the truncation of PgaB. Likewise, this packing also occurs for the second PgaB molecule in the asu, resulting in a head-to-toe orientation (supplemental Fig. S2). In addition to the canonical $(\beta/\alpha)_8$ fold characteristic of Tiny TIM (53), eight short α-helices are present on one face of the barrel, creating a pronounced electronegative groove ~45 Å long and 10 Å wide (Fig. 1C). A structural similarity search using the DALI fold recognition server (54) indicates that the C-terminal domain of PgaB has structural similarity to a large number of glycoside hydrolases. The top structural hits have Z-scores ranging from ~15 to 18 and include a variety of glycoside hydrolase (GH) families. The most predominantly occurring GH families were GH5, GH14, GH18, and GH20 (19). Families GH5, GH18, and GH20 contain proteins that bind or hydrolyze GlcNAc substrates, such as gangliosides, chitin, or PNAG (19). The top hit, a member of GH14, was β-amylase from *Bacillus cereus* (Protein Data Bank entry 1J10) (55), with an RMSD of 3.7 Å over 278 equivalent Cαs (Fig. 3A). The pronounced C-terminal groove (Fig. 1C) contains predominantly negatively charged and aromatic residues that are well conserved among PgaB homologues (Fig. 3B) and are common to glycoside hydrolase active
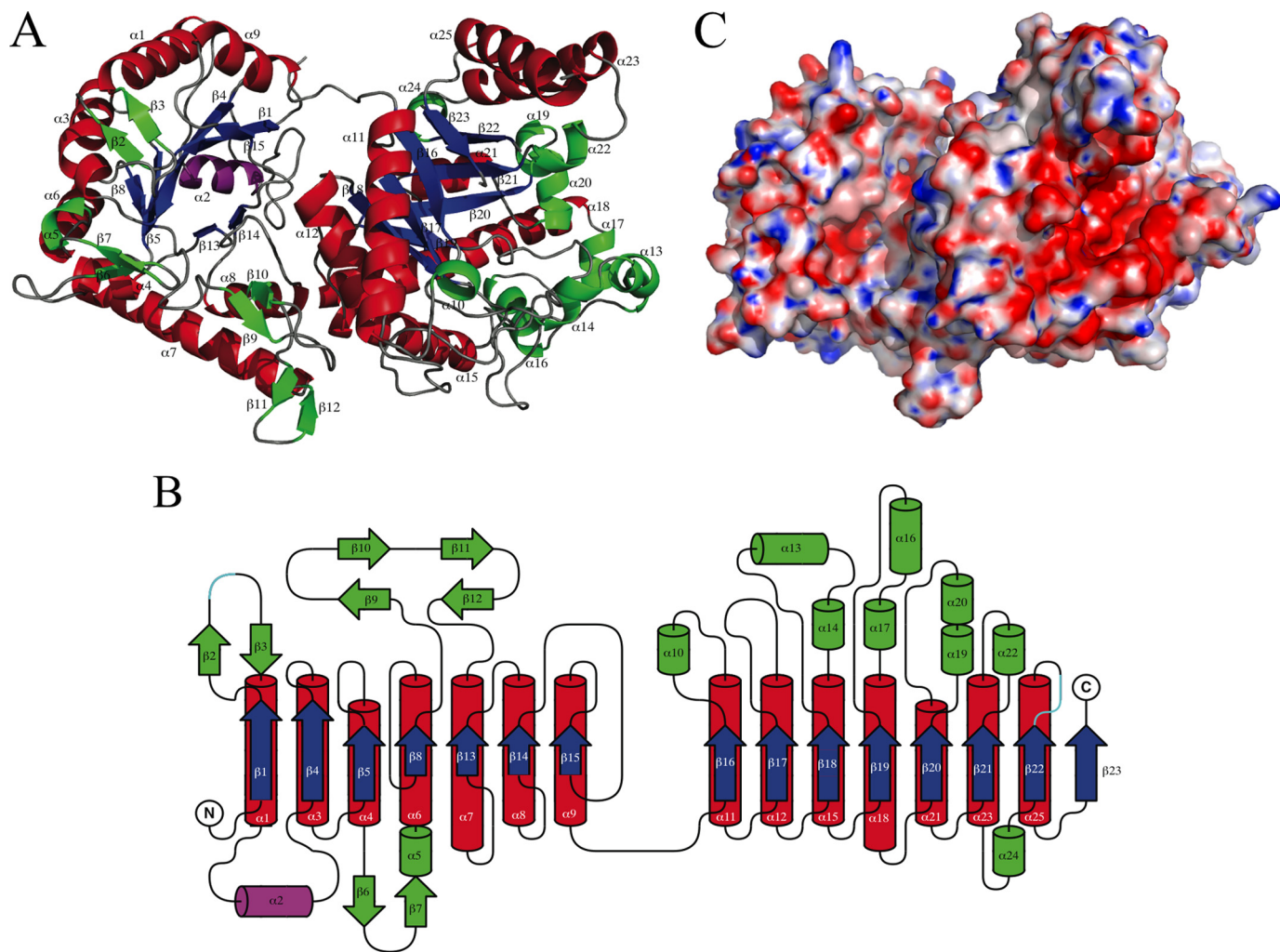
FIGURE 1. PgaB structure. *A* and *B*, overall structure of PgaB shown in cartoon (*A*) and topology (*B*) representations with the secondary structural elements labeled. The canonical $(\beta/\alpha)_x$ folds are colored *blue* (β-strands) and *red* (α-helices), respectively. Additional secondary structure elements are colored *green*, except the CE4 capping helix, which is shown in *purple*. Residues not modeled in the topology diagram are colored *cyan*. *C*, electrostatic surface representation of PgaB (generated in PyMOL using APBS) shown in the same orientation as in *A*. The face of the C-terminal domain contains a pronounced electronegative groove that extends toward the N-terminal domain. Quantitative electrostatics are colored from *red* (-20 kT) to *blue* ($+20$ kT).

sites (56). This groove is large enough to accommodate a PNAG octamer in an extended conformation. Interestingly, examination of the $|F_o - F_c|$ electron density maps revealed a large positive peak ($4\,\sigma$) in the groove that could be confidently modeled as a MES molecule (Fig. 3*B*). The presence of a MES molecule in the active site groove has also been observed for the chitinase Chia1b from *Oryza sativa* (Protein Data Bank entry 3IWR) (57).

The De-N-acetylase Domain Has a Unique Arrangement of the Canonical CE4 Motifs—CE4s generally contain five conserved motifs, numbered MT1 through MT5 as they occur in the sequence. These motifs form the active site and are required for enzymatic activity (21, 23–25). Superposition of the PgaB and *Sp*PgdA active sites (Fig. 2*B*) reveal that PgaB contains elements from each of the five conserved motifs with six significant differences. PgaB has the canonical MT1 and MT2 motifs. MT1 contains the catalytic base, Asp-114, and the first metal coordinating residue, Asp-115, and MT2 contains the metal coordinating residues, His-184 and His-189. Interestingly though, PgaB has four residues separating His-184 and His-189

in MT2, as opposed to a three-residue spacer as seen in previous CE4 structures (21, 23–25). The MT3 motif is well conserved; however, the buried arginine, Arg-289, responsible for coordinating the catalytic base, Asp-114, occurs later in the sequence in MT5, and spans the base of the active site in a different orientation (Fig. 2*B*). MT4 has the most significant differences. Most notably, PgaB has a water molecule in place of the conserved aspartic acid Asp-391 in *Sp*PgdA. This aspartic acid is believed to be responsible for activating the catalytic acid His-417 in *Sp*PgdA. We hypothesize that this water molecule would not be able to activate His-55 in the same manner that an aspartic acid residue would, thus reducing the catalytic efficiency of PgaB. Next, Trp-392 in *Sp*PgdA, which is responsible for forming the active site groove along with Tyr-367 from MT3 (25), is replaced with Met-67 in PgaB (Fig. 2*B*). Furthermore, Leu-274 in PgaB replaces Trp-385 in *Sp*PgdA, and these residues form one side of the *N*-acetyl hydrophobic binding pocket. The MT5 motif is conserved in PgaB; however, the catalytic histidine, His-55, is found at the beginning of the sequence after $\beta 1$, instead of occurring two residues after the second *N*-acetyl

A

B

C
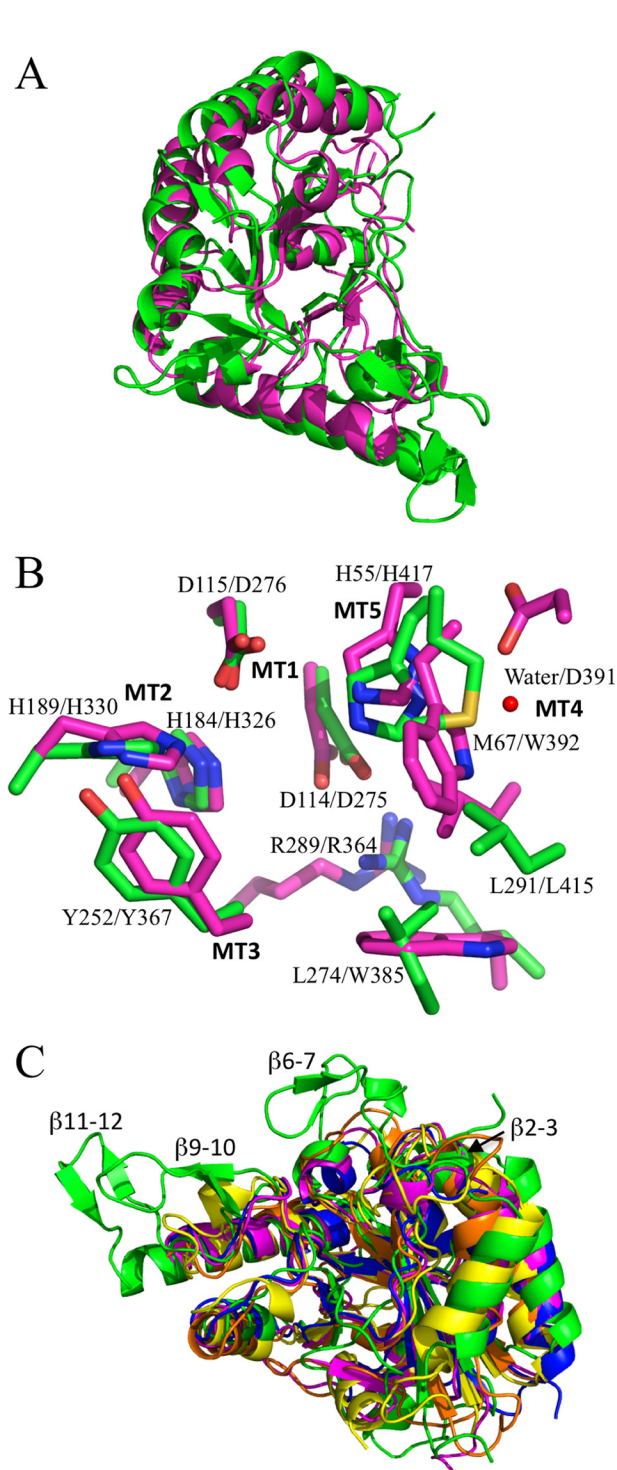
FIGURE 2. **Conservation of the CE4 fold.** *A*, superposition of the PgaB N-terminal domain (*green*) with *Sp*PgdA C-terminal domain (*magenta*) in cartoon representation shows the structural similarities between the de-*N*-acetylase domains. *B*, active site residues of PgaB and *Sp*PgdA shown in stick representation and colored as in *A* show differences in the canonical CE4 motifs. *C*, superposition of the PgaB N-terminal domain with other CE4 structures shown in cartoon representation reveals four β-hairpins, β2–3, β6–7, β9–10, and β11–12, that are unique to PgaB. The structures and corresponding RMSDs to the PgaB N-terminal domain (*green*) are: *Sp*PgdA C-terminal domain (*magenta*), 2.6 Å over 141 equivalent Cα atoms; CE4 from *Streptomyces lividans* (*blue*), 2.8 Å over 141 equivalent Cα atoms; CDA from *C. lindemuthianum* (*orange*), 2.9 Å over 146 equivalent Cα atoms; and PgdA from *S. mutans* (*yellow*), 3.3 Å over 141 equivalent Cα atoms.

A

B
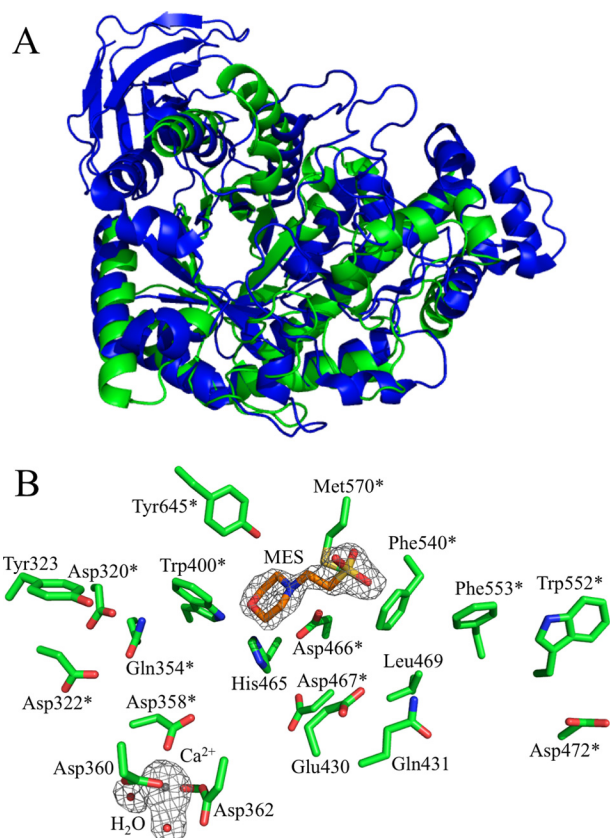
FIGURE 3. **The C-terminal domain of PgaB has structural similarity to glycoside hydrolases.** *A*, superposition of the C-terminal domain of PgaB (*green*) and β-amylase from *B. cereus* (*blue*) shown in cartoon representation. *B*, residues lining the electronegative groove with a bound MES molecule (*orange*) shown in stick representation, and calcium ion (*gray*) with the $|F_o - F_c|$ electron density map contoured at 3.0 σ. The residues indicated with an *asterisk* are ≥90% conserved among PgaB homologues for bacteria that have been shown to produce dPNAG in their biofilms.

hydrophobic binding pocket residue Leu-291. Together, the differences in MT4 and MT5 create a substrate-binding pocket that is deeper and wider in PgaB as compared with *Sp*PgdA (supplemental Fig. S3).

PgaB Exhibits Length-dependent De-N-acetylation of β-1,6-GlcNAc Oligomers—Previous *in vivo* characterization of the *pgaABCD* operon showed that deletion of *pgaB* or D115A or H184A mutations in a Δ*csrA* hyper-biofilm-forming strain produced fully acetylated PNAG that was retained in the periplasm (17). To confirm and characterize this activity *in vitro*, suitable substrates were needed. Although the natural polymer has poor solubility, short synthetic oligomers of defined length have improved solubility characteristics (48) and were used to examine the activity of PgaB. The de-*N*-acetylation activity was first confirmed by ^{1}H NMR and MALDI-MS. MBP-PgaB-treated β-1,6-(GlcNAc)$_5$ developed a triplet at 2.88 ppm in the ^{1}H NMR spectrum (supplemental Fig. S4, *A* and *B*). This signal is consistent with a proton at C2 of a glucosamine residue and agrees well with the signal observed in ^{1}H NMR from isolated dPNAG (8, 10). Further confirmation of de-*N*-acetylation was obtained through MALDI-MS, which indicated the presence of a compound with a molecular weight 42 *m/z* lower than that of β-1,6-(GlcNAc)$_5$, corresponding to the loss of an acetate group (supplemental Fig. S4C). To characterize

A

B
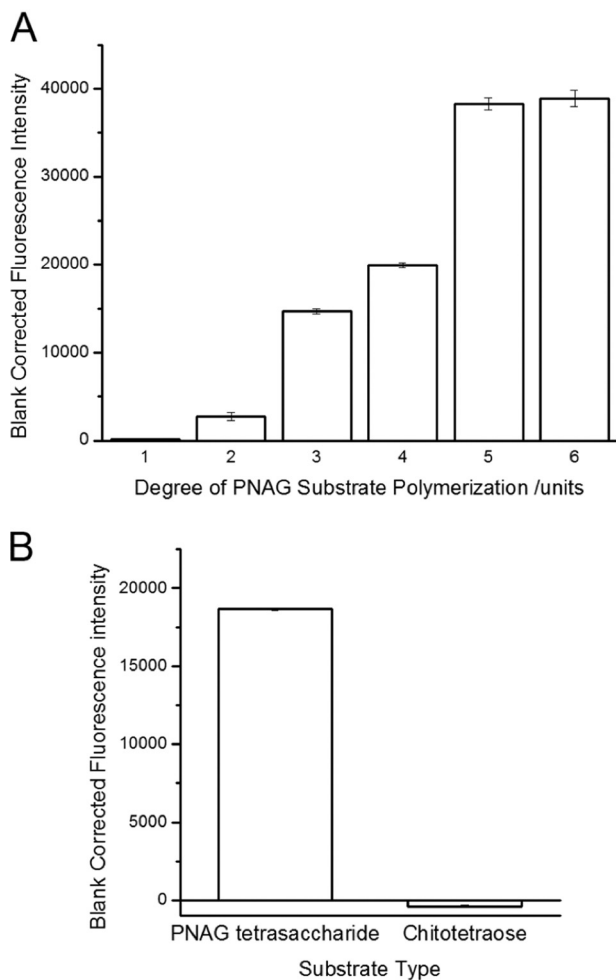
FIGURE 4. **Substrate dependence of PgaB.** Fluorescamine assay for Ni-loaded MBP-PgaB incubated with varying length β-1,6-GlcNAc oligomers showing increased rates of de-N-acetylation with increasing oligomer lengths (*A*) and β-1,6-(GlcNAc)$_4$ oligomer or β-1,4-(GlcNAc)$_4$ chitotetraose showing de-N-acetylation specificity for the β-1,6-(GlcNAc)$_4$ oligomer (*B*). The *bars* represent duplicate experiments with standard deviations.

de-N-acetylation on specific-length β-1,6-GlcNAc oligomers, a discontinuous fluorescamine assay was used to quantify the production of free amines (49). Using this assay, we observed a positive correlation between substrate length and de-N-acetylase activity. The de-N-acetylation rate increased as the length of the β-1,6-GlcNAc oligomer increased up to β-1,6-(GlcNAc)$_5$, with no further increase in de-N-acetylation rate with β-1,6-(GlcNAc)$_6$ (Fig. 4*A*). The observed k_{cat}/K_m for β-1,6-(GlcNAc)$_4$ and β-1,6-(GlcNAc)$_5$ are 0.057 $\pm$ 0.01 and 0.25 $\pm$ 0.01 M^{-1} s^{-1}, respectively (supplemental Fig. S5). Michaelis-Menten kinetic parameters could not be determined because of the limited solubility of the substrates, which prevented saturation of the de-N-acetylase activity. However, based on the solubility-limited kinetic profiles of PgaB, the K_m values for β-1,6-(GlcNAc)$_4$ and β-1,6-(GlcNAc)$_5$ are estimated to be greater than 100 and 25 mM, respectively. To determine whether PgaB was specific for β-1,6-linked GlcNAc oligomers, substrate specificity was evaluated using chitotetraose, a β-1,4-(GlcNAc)$_4$ oligomer. Fluorescamine assays showed no increase in free amine concentration for chitotetraose over a

24-h period, suggesting that PgaB has specificity for β-1,6-GlcNAc oligomers (Fig. 4*B*).

PgaB Activity Is Metal-dependent—Previously characterized CE4s that contain an Asp-His-His metal coordination site exhibit metal-dependent de-N-acetylase activity with a preference for Co^{2+} and Zn^{2+} (21, 23–25). To probe the metal-dependent activity of PgaB, the fluorescamine assay was used to conduct an activity-based metal screen. MBP-PgaB was used to eliminate any bias that may occur from scavenging metals from immobilized metal affinity chromatography. MBP-PgaB (as isolated) showed ~3- and 6-fold higher activity with Ni^{2+} and Co^{2+}, respectively (Fig. 5*A*). The fold increase in activity varied slightly between protein purifications, but Co^{2+} and Ni^{2+} consistently increased activity. In contrast, the addition of Zn^{2+} consistently resulted in a reduction in activity. As observed previously with the chitin deacetylase CDA from *Collectotrichum lindemuthianum*, lengthy incubations of PgaB with 100-fold excess of the metal chelator EDTA did not completely abolish activity (24). Complete inhibition of PgaB was achieved with 50 mM of the metal chelator dipicolinic acid (data not shown). Dipicolinic acid is smaller then EDTA and has been shown to inhibit enzymes by forming a ternary complex with the metal ion, whereas EDTA inhibition is dependent on the diffusion of the metal from the metal-enzyme complex (58–60).

PgaB Expression Conditions Determine the Identity of the Active Site Metal—Examination of SeMet PgaB difference electron density maps revealed two large 21 and 33 σ peaks for each molecule in the asu, suggesting the presence of two metal ions per PgaB molecule. The 21 σ peak, found in the C-terminal domain, is hepta-coordinate with contributions from the side chains of Asp-358, Asp-360 (bidentate), and Asp-362, a backbone carbonyl contribution from Leu-362, and two water molecules (Fig. 3*B*). This metal is likely Ca^{2+} because 0.2 M calcium acetate was present in the crystallization solution. The N-terminal de-N-acetylase domain contains the 33 σ peak, which is coordinated in an octahedral manor by the side chains of Asp-115, His-184, and His-189, an acetate molecule (bidentate), and a water molecule (Fig. 5*B*). The arrangement of the Asp-His-His residues is characteristic of the metal coordination found in other CE4s (21, 23–25).

To determine the identity of the metal, anomalous diffraction data were collected on the SeMet PgaB$^{42-655}$ crystal above and below the absorption edge for the transition metals iron through zinc. A significant decrease in anomalous signal was observed for the 33 σ peak for the data collected below the iron absorption edge (1.77 Å) (supplemental Fig. S6*A*), suggesting the presence of significant amounts of iron. This is the first structure of a CE4 family member that has been determined with iron coordinated in the active site. To confirm the metal bound was not the result of the cytosolic expression of SeMet PgaB in iron-rich (0.5 mM) minimal media, native sulfur-containing protein was prepared using LB media, which has ~0.1 mM iron and zinc (61). Anomalous diffraction data were subsequently collected as described previously on crystals grown from this protein and again suggested that PgaB was predominantly loaded with iron. These data suggest that PgaB binds iron under these expression conditions because its availability
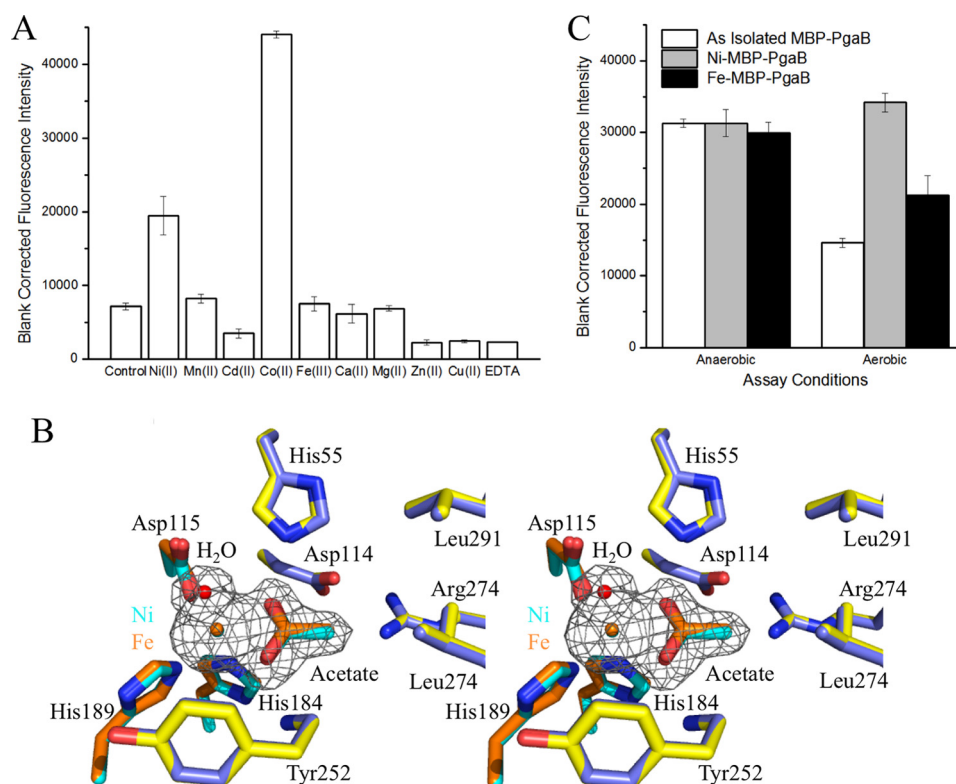
FIGURE 5. **Metal dependence of PgaB.** *A*, fluorescamine activity assay of MBP-PgaB (as isolated) in the presence of various metals. Increased rates were observed with the addition of Co^{2+} and Ni^{2+}. The *bars* represent duplicate experiments with standard deviations. *B*, superposition of the nickel- and iron-complexed PgaB de-*N*-acetylase active sites shown in stick representation and wall-eyed stereo view. Nickel and coordinating residues are shown in *cyan*, whereas iron and coordinating residues are shown in *orange*. Additional active site residues conserved among CE4s are colored *blue* (nickel complex) and *yellow* (iron complex). The $|F_o - F_c|$ electron density map for the nickel-complexed structure is shown in *gray* and contoured at 3.0 σ. *C*, fluorescamine activity assay comparison of MBP-PgaB (as isolated), nickel-loaded MBP-PgaB, and iron-loaded MBP-PgaB under aerobic and anaerobic conditions. The *bars* represent duplicate experiments with standard deviations.

is higher than that of nickel or cobalt and that PgaB maintains metal ion coordination throughout purification.

Although our initial metal-based activity screen had shown no increase in de-*N*-acetylation activity in the presence of additional Fe^{3+} (Fig. 5*A*), we wanted to test whether Fe^{2+} could increase activity. To determine whether Fe^{2+}-loaded PgaB may have increased catalytic activity, MBP-PgaB was expressed and purified under aerobic and anaerobic conditions in LB media with the addition of 1 mM iron citrate or nickel sulfate prior to induction to increase cellular levels of the respective metal. Cobalt could not be assessed using this procedure because of toxicity effects of the metal on *E. coli* growth. ICP-AES analysis of the purified proteins revealed ~70% iron or nickel loading (data not shown), respectively. The activity of the iron-loaded enzyme under anaerobic conditions was similar to that of nickel-loaded enzyme (Fig. 5*C*). However, in the aerobically prepared and incubated samples, the activity of the iron-loaded enzyme was significantly lower, suggesting that Fe^{2+}-loaded PgaB is the more active species (Fig. 5*C*).

To determine whether there are any structural differences that occur when the protein is loaded with different metals, nickel-loaded PgaB[42–655] was expressed, purified, and crystallized. Diffraction data were collected to 1.9 Å, and the structure was solved using the molecular replacement technique. Analysis of the difference electron density maps revealed a 43 σ peak coordinated by Asp-115, His-184, His-189, an acetate, and a

water molecule (Fig. 5*B*). Anomalous diffraction data were collected above (1.47 Å) and below (1.51 Å) the nickel absorption edge, and a significant decrease in the anomalous signal was observed below the nickel absorption edge (supplemental Fig. S6*B*), indicating the presence of significant amounts of nickel. No significant differences were observed in the nickel- and iron-complexed PgaB structures (RMSD of 0.12 Å over 587 equivalent Cα) or their active sites, where both metals were observed to have octahedral coordination geometries (Fig. 5*B*).

Determination of De-N-acetylation Position—The position of PgaB-catalyzed de-*N*-acetylation within the β-1,6-(GlcNAc)$_5$ and β-1,6-(GlcNAc)$_4$ oligomers was determined by treatment of the partially de-*N*-acetylated oligomers with *Sp*Hex prior to reacetylation and HPLC analysis (Fig. 6). *Sp*Hex is an exoglycosidase that cleaves *N*-acetyl-β-glucosaminyl oligosaccharides from the nonreducing terminus but does not cleave glucosaminyl residues (50). Analysis revealed a combination of GlcNAc, (GlcNAc)$_2$, and (GlcNAc)$_3$ cleavage products for the β-1,6-(GlcNAc)$_4$ oligomer, implying that de-*N*-acetylation occurred at the second or third residue from the nonreducing terminus (Fig. 6*B*). For the β-1,6-(GlcNAc)$_5$ oligomer, only two major cleavage products were observed, GlcNAc and (GlcNAc)$_3$, indicating that de-*N*-acetylation occurred predominantly at the third or central residue (Fig. 6*C*).

Using electrospray ionization MS analysis of the de-*N*-acetylated substrates, it was possible to observe di-de-*N*-acetylated
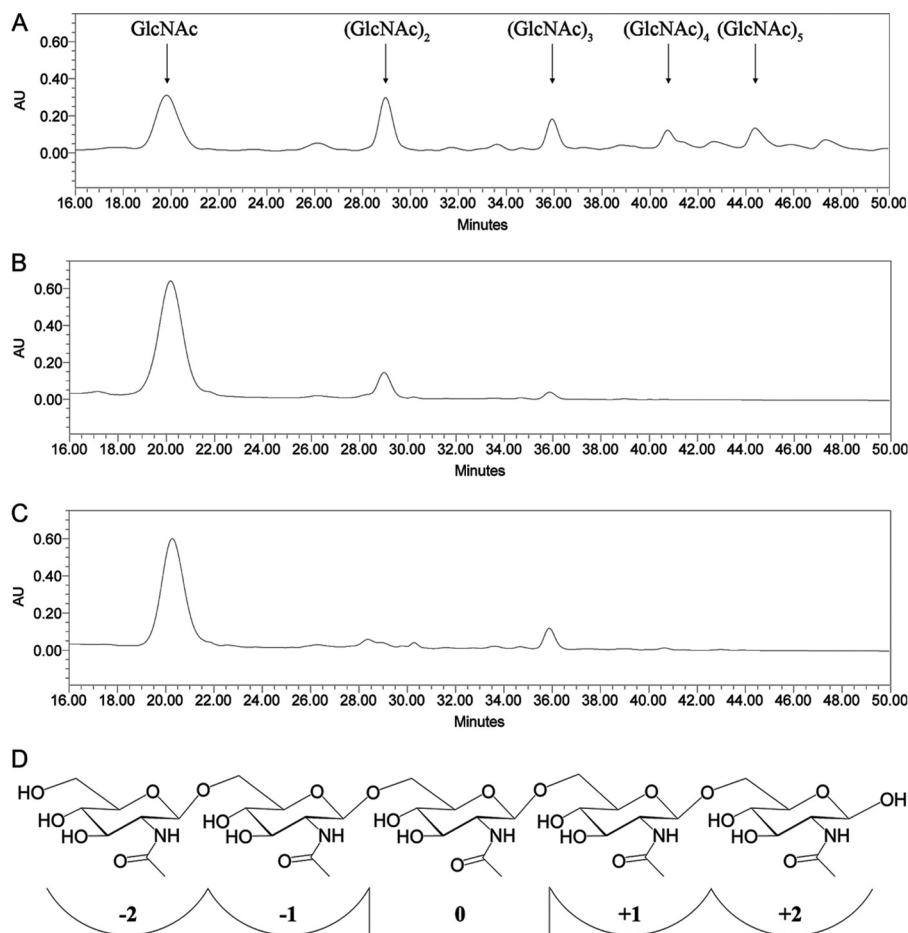
FIGURE 6. **HPLC analysis of the de-N-acetylation position.** *A*, analysis of β-1,6-GlcNAc standards with *arrows* indicating elution time of GlcNAc through (GlcNAc)$_5$. *AU*, absorbance units. *B* and *C*, analysis of re-N-acetylated *Sp*Hex degradation products of MBP-PgaB treated (*B*) β-1,6-(GlcNAc)$_4$ and (*C*) β-1,6-(GlcNAc)$_5$. The peaks were identified by comparing elution time to the β-1,6-(GlcNAc)$_n$ standards and confirmed by MALDI-MS. *D*, representation of productive binding mode of β-1,6-(GlcNAc)$_5$ to PgaB, showing de-N-acetylation preference for the central or *O*-subsite GlcNAc.

β-1,6-(GlcNAc)$_5$. Electrospray ionization MS-MS analysis was performed to determine the location of the de-N-acetylated residues. Analysis of the single de-N-acetylated β-1,6-(GlcNAc)$_5$ oligomer indicated that de-N-acetylation occurred on the central residue (supplemental Figs. S7A and S8A), in agreement with the results of the *Sp*Hex digestion assay. Analysis of the di-de-N-acetylated β-1,6-(GlcNAc)$_5$ oligomer revealed the presence of more than one species, corresponding to de-N-acetylation of the central residue and a reducing terminal residue (supplemental Figs. S7B and S8B).

DISCUSSION

Previous *in vivo* studies have shown that PgaB is required for de-N-acetylation of PNAG in *E. coli*, and impairment of this function prevents polymer export and abolishes biofilm formation (17). PgaB belongs to the CE4 family, members of which are responsible for the de-*N*- and *O*-acetylation of various polysaccharides. However, limited structural or mechanistic data are available for this family of enzymes, and no PNAG de-N-acetylases have been characterized to date. To further understand the mechanism of PNAG de-N-acetylation, we have determined the structure of PgaB and characterized its metal-dependent activity on short β-1,6-GlcNAc oligomers.

The N-terminal domain of PgaB shows structural features common to other CE4s, including a distorted $(\beta/\alpha)_7$ barrel fold and the conserved Asp-His-His metal coordinating triad. The similarity of the PgaB and *Sp*PgdA active sites (Fig. 2B) suggests that PgaB may utilize a catalytic mechanism comparable with that proposed for *Sp*PgdA (25). Despite the similarity of the active site residues, PgaB shows no de-N-acetylase activity on chitotetraose (Fig. 3B), a pseudosubstrate for *Sp*PgdA. The preference of PgaB for β-1,6-linked GlcNAc oligomers is likely due to its active site architecture. The active site of PgaB is situated in a deep bowl-shaped binding pocket. In contrast, *Sp*PgdA has a shallow elongated pocket that is narrower on one end (supplemental Fig. S3). The deep active site pocket found in PgaB is likely better suited to accommodate the PNAG polymer because of the inherent flexibility of the β-1,6-linkage, which allows for compact oligosaccharide conformations (62) that could extend into the deep pocket, unlike the more rigid linear β-1,4-linked polymers found in chitin and peptidoglycan.

PgaB demonstrates activity against β-1,6-GlcNAc oligomers with increasing rates up to β-1,6-(GlcNAc)$_5$, with catalytic efficiency ~4-fold greater for β-1,6-(GlcNAc)$_5$ than for β-1,6-(GlcNAc)$_4$ and dramatically reduced activity with shorter oligomers (Fig. 3A). The flanking subsites that recog-

nize the longer β-1,6-GlcNAc oligomers are likely formed by residues in β-hairpins $\beta6-7$, $\beta9-10$, and $\beta11-12$, which are not present in other CE4 structures (Fig. 2*C*). However, the catalytic efficiency (k_{cat}/K_m) of PgaB relative to other amide bond cleaving CE4s is markedly lower. PgaB has a 350,000–1,000,000-fold reduction in k_{cat}/K_m relative to CDA from *C. lindemuthianum*, which has rates of 56,000 and 89,000 $\text{M}^{-1}\text{s}^{-1}$ for β-1,4-(GlcNAc)$_4$ and β-1,4-(GlcNAc)$_5$, respectively (50). Similarly, PgaB has a 2000-fold reduction in rates relative to *Sp*PgdA, which has a rate of 150 $\text{M}^{-1}\text{s}^{-1}$ for β-1,4-(GlcNAc)$_3$ (25). The low catalytic efficiency may be related to the *in vivo* function of PgaB. The structure suggests that the reduced efficiency of the enzyme may be the consequence of the water molecule that is found in place of the conserved aspartic acid residue in the MT4 motif of other CE4 enzymes. This aspartic acid is believed to be responsible for tuning the pK_a of the catalytic histidine in MT5 (Fig. 2*A*), because the mutation D391N in *Sp*PgdA completely inactivates the enzyme (25). The reduced activity correlates with the biological process of PNAG biosynthesis. Only low levels of de-*N*-acetylation (22%) are observed in the secreted dPNAG polymer (17), whereas the level of de-*N*-acetylation of peptidoglycan can reach 80% (63) and can by even higher for chitin (64), suggesting that PgaB may be designed for low catalytic efficiency. Despite the low levels of de-*N*-acetylation, comparable activity was observed between β-1,6-(GlcNAc)$_5$ and β-1,6-(GlcNAc)$_6$, suggesting that there are four subsites flanking the active site (Fig. 6*D*). The enzyme also shows a de-*N*-acetylation preference for the central sugar of β-1,6-(GlcNAc)$_5$ (Fig. 6*C*). Di-de-*N*-acetylation of β-1,6-(GlcNAc)$_5$ was detected showing subsequent de-*N*-acetylations occurring to the reducing terminal side of the first de-*N*-acetylation site (supplemental Figs. S7 and S8). However, considering that the native substrate for PgaB is PNAG, other factors including the C-terminal domain may play role in catalytic efficiency, and increased rates may be seen with much longer β-1,6-GlcNAc oligomers that could span both domains. This hypothesis is supported by the observation that the C-terminal domain is required for the efficient de-*N*-acetylation of PNAG *in vivo*, because the rate of de-*N*-acetylation decreases from ~22% to ~1% when the domain is removed (17). However, the exact role this domain plays in de-*N*-acetylation remains unclear. Structurally, the domain is similar to a vast number of glycoside hydrolases including DspB, a well studied glucosaminidase that cleaves PNAG polymers and can disperse biofilms produced by a wide range of bacteria (65–67). Although glycoside hydrolases typically show activity with artificial substrates such as *para*-nitrophenyl glycosides (68), no hydrolase activity could be observed for PgaB with *para*-nitrophenyl GlcNAc or any of the synthetic β-1,6-GlcNAc oligomers used in the de-*N*-acetylation assays (data not shown). Together these data suggest that this domain may play a role in PNAG binding and may guide the polymer to the de-*N*-acetylase domain; this hypothesis is supported by the large number of conserved residues that form a negatively charged groove that would be able to accommodate a PNAG octamer in an extended conformation (Figs. 1*C* and 3*B*). Similarly, the domain could also bind the polymer after de-*N*-acetylation to initiate an interaction with PgaA to facilitate export of dPNAG across the membrane. Further studies involving this domain will be required to understand the exact role it plays in the biosynthesis and export of dPNAG.

The majority of CE4s characterized to date catalyze deacetylation using a divalent metal cofactor, predominantly Co^{2+} or Zn^{2+} (21, 23–25). The studies presented herein indicate PgaB as a metal-dependent de-*N*-acetylase with a preference for Co^{2+}, Ni^{2+}, and Fe^{2+} (Fig. 5, *A* and *C*). The preference of Ni^{2+} rather than Zn^{2+} for a CE4 enzyme has only been previously documented for the xylan de-*O*-acetylase CE4 from *Clostridium thermocellum* (*Ct*CE4) (23). Interestingly, *Ct*CE4 is also missing important CE4 motifs, including a metal coordinating histidine from MT2. The preference of Ni^{2+} over Zn^{2+} for PgaB and *Ct*CE4 does not appear to be from altered coordination chemistry, because Zn^{2+}-bound CE4 structures also have octahedral arrangements. In addition to PgaB and *Ct*CE4, two other CE4s, PdaA from *Bacillus subtilis* and ECU11_510 from *Encephalitozoon cuniculi*, also contain significant deviations to the canonical CE4 motifs (22, 52). PdaA has modifications in MT1 and MT3, whereas ECU11_510 has modification in MT1–MT4 (22, 52). This comparison suggests that modifications to CE4 motifs may provide a means to alter substrate specificity, metal preference, or the reaction mechanism. Despite the variety of metal ions CE4s use for catalysis, the ability of PgaB to use Fe^{2+} as the metal cofactor is a first for the CE4 family. However, it is possible that other CE4s are also active with Fe^{2+}, because a number of zinc-dependent deacetylases, such as LpxC and histone deacetylase 8, have recently been reclassified as non-heme Fe^{2+} hydrolases (69–72). The oxidation of Fe^{2+} under aerobic conditions during purification and enzymatic assays may have prevented the iron dependence of other CE4s from being observed. The identity of the endogenous metal that PgaB binds *in vivo* is difficult to assess. However, the promiscuity that has been observed for PgaB suggests that the bound metal ion will most likely change depending on cellular availability and environmental conditions. It is perhaps not surprising that proteins involved in biofilm formation are able to adapt to perform their function, because biofilms are an important process for bacterial survival under stress or low nutrient conditions.

The structures of PgaB reported here will help facilitate further studies into the biosynthesis of this important biofilm exopolysaccharide, dPNAG. Further characterization of both the de-*N*-acetylase and C-terminal domains of PgaB will help guide the development of novel inhibitors to combat biofilm formation in *E. coli* and the broad range of medically relevant bacteria that produce the same conserved dPNAG polysaccharide.

REFERENCES

1. Costerton, J. W., Stewart, P. S., and Greenberg, E. P. (1999) Bacterial biofilms. A common cause of persistent infections. *Science* **284,** 1318–1322

2. Hall-Stoodley, L., Costerton, J. W., and Stoodley, P. (2004) Bacterial biofilms. From the natural environment to infectious diseases. *Nat. Rev. Microbiol.* **2,** 95–108

3. López, D., Vlamakis, H., and Kolter, R. (2010) Biofilms. *Cold Spring Harb. Perspect. Biol.* **2,** a000398

4. Donlan, R. M., and Costerton, J. W. (2002) Biofilms. Survival mechanisms of clinically relevant microorganisms. *Clin. Microbiol. Rev.* **15,** 167–193

5. Mah, T. F., and O'Toole, G. A. (2001) Mechanisms of biofilm resistance to antimicrobial agents. *Trends Microbiol.* **9,** 34–39

6. Sutherland, I. (2001) Biofilm exopolysaccharides. A strong and sticky framework. *Microbiology* **147,** 3–9

7. Vu, B., Chen, M., Crawford, R. J., and Ivanova, E. P. (2009) Bacterial extracellular polysaccharides involved in biofilm formation. *Molecules* **14,** 2535–2554

8. Mack, D., Fischer, W., Krokotsch, A., Leopold, K., Hartmann, R., Egge, H., and Laufs, R. (1996) The intercellular adhesin involved in biofilm accumulation of *Staphylococcus epidermidis* is a linear β-1,6-linked glucosaminoglycan. Purification and structural analysis. *J. Bacteriol.* **178,** 175–183

9. McKenney, D., Pouliot, K. L., Wang, Y., Murthy, V., Ulrich, M., Döring, G., Lee, J. C., Goldmann, D. A., and Pier, G. B. (1999) Broadly protective vaccine for *Staphylococcus aureus* based on an *in vivo*-expressed antigen. *Science* **284,** 1523–1527

10. Wang, X., Preston, J. F., 3rd, and Romeo, T. (2004) The pgaABCD locus of *Escherichia coli* promotes the synthesis of a polysaccharide adhesin required for biofilm formation. *J. Bacteriol.* **186,** 2724–2734

11. Choi, A. H., Slamti, L., Avci, F. Y., Pier, G. B., and Maira-Litrán, T. (2009) The pgaABCD locus of *Acinetobacter baumannii* encodes the production of poly-β-1–6-*N*-acetylglucosamine, which is critical for biofilm formation. *J. Bacteriol.* **191,** 5953–5963

12. Sloan, G. P., Love, C. F., Sukumar, N., Mishra, M., and Deora, R. (2007) The Bordetella Bps polysaccharide is critical for biofilm development in the mouse respiratory tract. *J. Bacteriol.* **189,** 8270–8276

13. Conover, M. S., Sloan, G. P., Love, C. F., Sukumar, N., and Deora, R. (2010) The Bps polysaccharide of *Bordetella pertussis* promotes colonization and biofilm formation in the nose by functioning as an adhesin. *Mol. Microbiol.* **77,** 1439–1455

14. Izano, E. A., Sadovskaya, I., Vinogradov, E., Mulks, M. H., Velliyagounder, K., Ragunath, C., Kher, W. B., Ramasubbu, N., Jabbouri, S., Perry, M. B., and Kaplan, J. B. (2007) Poly-*N*-acetylglucosamine mediates biofilm formation and antibiotic resistance in *Actinobacillus pleuropneumoniae*. *Microb. Pathog.* **43,** 1–9

15. Jarrett, C. O., Deak, E., Isherwood, K. E., Oyston, P. C., Fischer, E. R., Whitney, A. R., Kobayashi, S. D., DeLeo, F. R., and Hinnebusch, B. J. (2004) Transmission of *Yersinia pestis* from an infectious biofilm in the flea vector. *J. Infect. Dis.* **190,** 783–792

16. Yakandawala, N., Gawande, P. V., LoVetri, K., Cardona, S. T., Romeo, T., Nitz, M., and Madhyastha, S. (2011) Characterization of the poly-β-1,6-*N*-acetylglucosamine polysaccharide component of *Burkholderia* biofilms. *Appl. Environ. Microbiol.* **77,** 8303–8309

17. Itoh, Y., Rice, J. D., Goller, C., Pannuri, A., Taylor, J., Meisner, J., Beveridge, T. J., Preston, J. F., 3rd, and Romeo, T. (2008) Roles of pgaABCD genes in synthesis, modification, and export of the *Escherichia coli* biofilm adhesin poly-β-1,6-*N*-acetyl-D-glucosamine. *J. Bacteriol.* **190,** 3670–3680

18. Boehm, A., Steiner, S., Zaehringer, F., Casanova, A., Hamburger, F., Ritz, D., Keck, W., Ackermann, M., Schirmer, T., and Jenal, U. (2009) Second messenger signalling governs *Escherichia coli* biofilm induction upon ribosomal stress. *Mol. Microbiol.* **72,** 1500–1516

19. Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) The Carbohydrate-Active EnZymes database (CAZy). An expert resource for Glycogenomics. *Nucleic Acids Res.* **37,** D233–D238

20. Caufrier, F., Martinou, A., Dupont, C., and Bouriotis, V. (2003) Carbohydrate esterase family 4 enzymes. Substrate specificity. *Carbohydr. Res.* **338,** 687–692

21. Deng, D. M., Urch, J. E., ten Cate, J. M., Rao, V. A., van Aalten, D. M., and Crielaard, W. (2009) *Streptococcus mutans* SMU.623c codes for a functional, metal-dependent polysaccharide deacetylase that modulates interactions with salivary agglutinin. *J. Bacteriol.* **191,** 394–402

22. Urch, J. E., Hurtado-Guerrero, R., Brosson, D., Liu, Z., Eijsink, V. G., Texier, C., and van Aalten, D. M. (2009) Structural and functional characterization of a putative polysaccharide deacetylase of the human parasite *Encephalitozoon cuniculi*. *Protein Sci.* **18,** 1197–1209

23. Taylor, E. J., Gloster, T. M., Turkenburg, J. P., Vincent, F., Brzozowski, A. M., Dupont, C., Shareck, F., Centeno, M. S., Prates, J. A., Puchart, V., Ferreira, L. M., Fontes, C. M., Biely, P., and Davies, G. J. (2006) Structure and activity of two metal ion-dependent acetylxylan esterases involved in plant cell wall degradation reveals a close similarity to peptidoglycan deacetylases. *J. Biol. Chem.* **281,** 10968–10975

24. Blair, D. E., Hekmat, O., Schüttelkopf, A. W., Shrestha, B., Tokuyasu, K., Withers, S. G., and van Aalten, D. M. (2006) Structure and mechanism of chitin deacetylase from the fungal pathogen *Colletotrichum lindemuthianum*. *Biochemistry* **45,** 9416–9426

25. Blair, D. E., Schüttelkopf, A. W., MacRae, J. I., and van Aalten, D. M. (2005) Structure and metal-dependent mechanism of peptidoglycan deacetylase, a streptococcal virulence factor. *Proc. Natl. Acad. Sci. U.S.A.* **102,** 15429–15434

26. Oberbarnscheidt, L., Taylor, E. J., Davies, G. J., and Gloster, T. M. (2007) Structure of a carbohydrate esterase from *Bacillus anthracis*. *Proteins* **66,** 250–252

27. Forman, S., Bobrov, A. G., Kirillina, O., Craig, S. K., Abney, J., Fetherston, J. D., and Perry, R. D. (2006) Identification of critical amino acid residues in the plague biofilm Hms proteins. *Microbiology* **152,** 3399–3410

28. Bateman, A., Coin, L., Durbin, R., Finn, R. D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E. L., Studholme, D. J., Yeats, C., and Eddy, S. R. (2004) The Pfam protein families database. *Nucleic Acids Res.* **32,** D138–D141

29. Whitney, J. C., Hay, I. D., Li, C., Eckford, P. D., Robinson, H., Amaya, M. F., Wood, L. F., Ohman, D. E., Bear, C. E., Rehm, B. H., and Howell, P. L. (2011) Structural basis for alginate secretion across the bacterial outer membrane. *Proc. Natl. Acad. Sci. U.S.A.* **108,** 13083–13088

30. Keiski, C. L., Harwich, M., Jain, S., Neculai, A. M., Yip, P., Robinson, H., Whitney, J. C., Riley, L., Burrows, L. L., Ohman, D. E., and Howell, P. L. (2010) AlgK is a TPR-containing protein and the periplasmic component of a novel exopolysaccharide secretin. *Structure* **18,** 265–273

31. Gatto, G. J., Jr., Geisbrecht, B. V., Gould, S. J., and Berg, J. M. (2000) Peroxisomal targeting signal-1 recognition by the TPR domains of human PEX5. *Nat. Struct. Biol.* **7,** 1091–1095

32. Jínek, M., Rehwinkel, J., Lazarus, B. D., Izaurralde, E., Hanover, J. A., and Conti, E. (2004) The superhelical TPR-repeat domain of *O*-linked GlcNAc transferase exhibits structural similarities to importin alpha. *Nat. Struct. Mol. Biol.* **11,** 1001–1007

33. Wu, Y., and Sha, B. (2006) Crystal structure of yeast mitochondrial outer membrane translocon member Tom70p. *Nat. Struct. Mol. Biol.* **13,** 589–593

34. Koo, J., Tammam, S., Ku, S. Y., Sampaleanu, L. M., Burrows, L. L., and Howell, P. L. (2008) PilF is an outer membrane lipoprotein required for multimerization and localization of the *Pseudomonas aeruginosa* Type IV pilus secretin. *J. Bacteriol.* **190,** 6961–6969

35. Zeytuni, N., and Zarivach, R. (2012) Structural and functional discussion of the tetra-trico-peptide repeat, a protein interaction module. *Structure* **20,** 397–405

36. Abu Khweek, A., Fetherston, J. D., and Perry, R. D. (2010) Analysis of HmsH and its role in plague biofilm formation. *Microbiology* **156,** 1424–1438

37. McCafferty, D. G., Lessard, I. A., and Walsh, C. T. (1997) Mutational analysis of potential zinc-binding residues in the active site of the enterococcal D-Ala-D-Ala dipeptidase VanX. *Biochemistry* **36,** 10498–10505

38. Little, D. J., Whitney, J. C., Robinson, H., Yip, P., Nitz, M., and Howell, P. L. (2012) Combining *in situ* proteolysis and mass spectrometry to crystallize *Escherichia coli* PgaB. *Acta Crystallogr Sect. F Struct. Biol. Cryst. Commun.* **68,** 842–845

39. Otwinowski, Z., and Minor, W. (1997) Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276,** 307–326

40. Pape, T., and Schneider, T. R. (2004) HKL2MAP. A graphical user interface for phasing with SHELX programs. *J. Appl. Crystallogr.* **37,** 843–844

41. Terwilliger, T. C., and Berendzen, J. (1999) Automated MAD and MIR

structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **55,** 849–861

42. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., Mc-Coy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX. A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 213–221

43. Emsley, P., and Cowtan, K. (2004) Coot. Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2126–2132

44. Dolinsky, T. J., Czodrowski, P., Li, H., Nielsen, J. E., Jensen, J. H., Klebe, G., and Baker, N. A. (2007) PDB 2PQR. Expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res.* **35,** W522–W525

45. Dolinsky, T. J., Nielsen, J. E., McCammon, J. A., and Baker, N. A. (2004) PDB 2PQR. An automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* **32,** W665–W667

46. Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) Electrostatics of nanosystems. Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A.* **98,** 10037–10041

47. Bond, C. S. (2003) TopDraw. A sketchpad for protein structure topology cartoons. *Bioinformatics* **19,** 311–312

48. Leung, C., Chibba, A., Gomez-Biagi, R. F., and Nitz, M. (2009) Efficient synthesis and protein conjugation of β-(1→6)-D-N-acetylglucosamine oligosaccharides from the polysaccharide intercellular adhesin. *Carbohydr. Res.* **344,** 570–575

49. Udenfriend, S., Stein, S., Böhlen, P., Dairman, W., Leimgruber, W., and Weigele, M. (1972) Fluorescamine. A reagent for assay of amino acids, peptides, proteins, and primary amines in the picomole range. *Science* **178,** 871–872

50. Hekmat, O., Tokuyasu, K., and Withers, S. G. (2003) Subsite structure of the endo-type chitin deacetylase from a deuteromycete, *Colletotrichum lindemuthianum*. An investigation using steady-state kinetic analysis and MS. *Biochem. J.* **374,** 369–380

51. Hendrickson, W. A. (1991) Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. *Science* **254,** 51–58

52. Blair, D. E., and van Aalten, D. M. (2004) Structures of *Bacillus subtilis* PdaA, a family 4 carbohydrate esterase, and a complex with *N*-acetyl-glucosamine. *FEBS Lett.* **570,** 13–19

53. Walden, H., Bell, G. S., Russell, R. J., Siebers, B., Hensel, R., and Taylor, G. L. (2001) Tiny TIM. A small, tetrameric, hyperthermostable triose-phosphate isomerase. *J. Mol. Biol.* **306,** 745–757

54. Holm, L., and Sander, C. (1993) Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233,** 123–138

55. Oyama, T., Miyake, H., Kusunoki, M., and Nitta, Y. (2003) Crystal structures of β-amylase from *Bacillus cereus* var. mycoides in complexes with substrate analogs and affinity-labeling reagents. *J. Biochem.* **133,** 467–474

56. Ketudat Cairns, J. R., and Esen, A. (2010) β-Glucosidases. *Cell Mol. Life Sci.* **67,** 3389–3405

57. Kezuka, Y., Kojima, M., Mizuno, R., Suzuki, K., Watanabe, T., and Nonaka, T. (2010) Structure of full-length class I chitinase from rice revealed by x-ray crystallography and small-angle X-ray scattering. *Proteins* **78,** 2295–2305

58. Kidani, Y., Hirose, J., and Koike, H. (1976) Coordination chemical studies on metalloenzymes. Kinetics and mechanism of the Zn(II) exchange re-

action between chelating agent and apo-bovine carbonic anhydrase. *J. Biochem.* **79,** 43–51

59. Kidani, Y., and Hirose, J. (1977) Coordination chemical studies on metalloenzymes. II. Kinetic behavior of various types of chelating agents towards bovine carbonic anhydrase. *J. Biochem.* **81,** 1383–1391

60. Labbé, G., Krismanich, A. P., de Groot, S., Rasmusson, T., Shang, M., Brown, M. D., Dmitrienko, G. I., and Guillemette, J. G. (2012) Development of metal-chelating inhibitors for the Class II fructose 1,6-bisphosphate (FBP) aldolase. *J. Inorg. Biochem.* **112,** 49–58

61. Outten, C. E., and O'Halloran, T. V. (2001) Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis. *Science* **292,** 2488–2492

62. Grachev, A. A., Gerbst, A. G., Gening, M. L., Titov, D. V., Yudina, O. N., Tsvetkov, Y. E., Shashkov, A. S., Pier, G. B., and Nifantiev, N. E. (2011) NMR and conformational studies of linear and cyclic oligo-(1→6)-β-D-glucosamines. *Carbohydr. Res.* **346,** 2499–2510

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

64. Nwe, N., Furuike, T., and Tamura, H. (2009) The mechanical and biological properties of chitosan scaffolds for tissue regeneration templates are significantly enhanced by chitosan from *Gongronella butleri*. *Materials* **2,** 374–398

65. Itoh, Y., Wang, X., Hinnebusch, B. J., Preston, J. F., and Romeo, T. (2005) Depolymerization of β-1,6-*N*-acetyl-D-glucosamine disrupts the integrity of diverse bacterial biofilms. *J. Bacteriol.* **187,** 382–387

66. Parise, G., Mishra, M., Itoh, Y., Romeo, T., and Deora, R. (2007) Role of a putative polysaccharide locus in *Bordetella* biofilm development. *J. Bacteriol.* **189,** 750–760

67. Kaplan, J. B., Ragunath, C., Ramasubbu, N., and Fine, D. H. (2003) Detachment of *Actinobacillus actinomycetemcomitans* biofilm cells by an endogenous β-hexosaminidase activity. *J. Bacteriol.* **185,** 4693–4698

68. Stütz, A. E., and Wrodnigg, T. M. (2011) Imino sugars and glycosyl hydrolases. Historical context, current aspects, emerging trends. *Adv. Carbohydr Chem. Biochem.* **66,** 187–298

69. Gattis, S. G., Hernick, M., and Fierke, C. A. (2010) Active site metal ion in UDP-3-*O*-((*R*)-3-hydroxymyristoyl)-*N*-acetylglucosamine deacetylase (LpxC) switches between Fe(II) and Zn(II) depending on cellular conditions. *J. Biol. Chem.* **285,** 33788–33796

70. Hernick, M., Gattis, S. G., Penner-Hahn, J. E., and Fierke, C. A. (2010) Activation of *Escherichia coli* UDP-3-O-[(*R*)-3-hydroxymyristoyl]-*N*-acetylglucosamine deacetylase by Fe^{2+} yields a more efficient enzyme with altered ligand affinity. *Biochemistry* **49,** 2246–2255

71. Gantt, S. L., Gattis, S. G., and Fierke, C. A. (2006) Catalytic activity and inhibition of human histone deacetylase 8 is dependent on the identity of the active site metal ion. *Biochemistry* **45,** 6170–6178

72. Gantt, S. L., Joseph, C. G., and Fierke, C. A. (2010) Activation and inhibition of histone deacetylase 8 by monovalent cations. *J. Biol. Chem.* **285,** 6036–6043

73. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity. All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 12–21