Published in final edited form as:

Biochemistry. 2016 August 16; 55(32): 4495–4508. doi:10.1021/acs.biochem.6b00536.

Structural Studies on a Glucosamine/Glucosaminide *N*-Acetyltransferase

Brandon J. Dopkins[†], Peter A. Tipton^{*‡}, James B. Thoden[†], and Hazel M. Holden^{*†}

[†]Department of Biochemistry, University of Wisconsin, Madison, Wisconsin 53706, United States

[‡]Department of Biochemistry, University of Missouri, Columbia, Missouri 65211, United States

Abstract

Glucosamine/glucosaminide *N*-acetyltransferase or GlmA catalyzes the transfer of an acetyl group from acetyl CoA to the primary amino group of glucosamine. The enzyme from *Clostridium acetobutylicum* is thought to be involved in cell wall rescue. In addition to glucosamine, GlmA has been shown to function on di- and trisaccharides of glucosamine as well. Here we present a structural and kinetic analysis of the enzyme. For this investigation, eight structures were determined to resolutions of 2.0 Å or better. The overall three-dimensional fold of GlmA places it into the tandem GNAT superfamily. Each subunit of the dimer folds into two distinct domains which exhibit high three-dimensional structural similarity. Whereas both domains bind acetyl CoA, it is the C-terminal domain that is catalytically competent. On the basis of the various structures determined in this investigation, two amino acid residues were targeted for further study: Asp 287 and Tyr 297. Although their positions in the active site suggested that they may play key roles in catalysis by functioning as active site bases and acids, respectively, this was not borne out by characterization of the D287N and Y297F variants. The kinetic properties revealed that both residues were important for substrate binding but had no critical roles as acid/base catalysts. Kinetic analyses also indicated that GlmA follows an ordered mechanism with acetyl CoA binding first followed by glucosamine. The product *N*-acetylglucosamine is then released prior to CoA. The investigation described herein provides significantly new information on enzymes belonging to the tandem GNAT superfamily.

Graphical abstract

^{*}Corresponding Authors (H.M.H.) Hazel_Holden@biochem.wisc.edu. Fax: 608-262-1319. Phone: 608-262-4988. ^{*}(P.A.T.) TiptonP@missouri.edu.

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.6b00536. Five additional figures (PDF)

Accession Codes

X-ray coordinates have been deposited in the Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (accession nos. 5kf1, 5kf2, 5kf8, 5kf9, 5kga, 5kgh, 5kgj, 5kgp).

Notes

The authors declare no competing financial interest.

Most bacteria contain on the outside of their cytoplasmic membranes a complex glycoconjugate referred to as the peptidoglycan or murein.¹ This large structural polymer not only protects the organism from degradation but also plays a key role in cell shape maintenance. Additionally, it serves as an anchoring point for various proteins and teichoic acids.^{2,3} Whereas the exact molecular architecture of the peptidoglycan varies among bacterial species, common features typically include linear glycan chains composed of alternating *N*-acetylmuramic acid and *N*-acetylglucosamine (GlcNAc) residues attached via β -1 $\rightarrow$ 4 linkages (Scheme 1).

Interestingly, some bacteria contain a high proportion of deacetylated glucosamine residues in their peptidoglycans. It has been estimated that the peptidoglycan of *Streptococcus pneumoniae*, for example, contains between 40 and 80% glucosamine rather than *N*-acetylglucosamine residues.⁴ As first reported for *Bacillus cereus* in 1971, the presence of these deacetylated sugars confers resistance to degradation by lysozyme.⁵ Indeed, it has been recently shown that the lack of a peptidoglycan *N*-acetylglucosamine deacetylase in *S. pneumoniae* results in its hypersensitivity to exogenous lysozyme.⁶ Likewise it has been demonstrated that peptidoglycan deacetylation in *Helicobacter pylori* moderates the immune response of the host.⁷

The peptidoglycan is, by no means, a static structure, however. It is constantly undergoing modifications and remodeling with the release of peptidoglycan fragments and the incorporation of new ones.⁸ Up to 50% of the cell wall is thought to be turned over in one generation.⁹ The sugar fragments produced are most likely recovered and recycled.¹⁰ One enzyme involved in this recycling process is NagZ, a β -*N*-acetylglucosaminidase that in *Bacillus subtilis* is highly specific for the *N*-acetyl group of the substrate.¹¹ Given that many of the sugar fragments from the peptidoglycan are deacetylated, it has been suggested that an enzyme is required to acetylate the glucosamine moieties of these fragments, which can then be subsequently processed by NagZ. Recently, an enzyme from *Clostridium acetobutylicum*, referred to as GlmA, has been characterized and has been proposed to play such a role in cell wall rescue.¹² Importantly, GlmA shows high specificity for both glucosamine and chitosan fragments (β -1 $\rightarrow$ 4-linked glucosamine residues).

Here we describe the high-resolution structure of *C. acetobutylicum* GlmA and show that it belongs to the tandem GNAT superfamily. For this investigation, eight structures were

determined to resolutions of 2.0 Å or better. In addition, a detailed analysis of the kinetic properties of GlmA was conducted. Taken together the data presented here shed new light on this fascinating glucosamine/glucosaminide *N*-acetyltransferase.

MATERIALS AND METHODS

Protein Expression and Purification

The gene encoding GlmA (*C. acetobutylicum* ATCC 824) was synthesized by DNA2.0 using optimized *Escherichia coli* codons. It was placed into the pJ411 expression plasmid, which ultimately leads to a protein with a C-terminal histidine tag of the sequence GGGHHHHH.

The pJ411-*glmA* plasmid was used to transform Rosetta2-(DE3) *E. coli* cells (Novagen). The cultures were grown with shaking in lysogeny broth supplemented with 50 mg/L kanamycin and 50 mg/L chloramphenicol at 37 °C until an optical density of 0.8 was reached at 600 nm. The flasks were cooled in an ice bath, and the cells were induced with 1 mM isopropyl β -D-1-thiogalactopyranoside and allowed to express protein at 16 °C for 24 h.

The cells were harvested by centrifugation and disrupted by sonication on ice. The lysate was cleared by centrifugation, and GlmA was purified utilizing Ni-NTA resin (Qiagen) according to the manufacturer's instructions. The protein was dialyzed against 10 mM Tris-HCl (pH 8.0) and 200 mM NaCl and concentrated to 8.2 mg/mL based on an extinction coefficient of $0.64 \text{ (mg/mL)}^{-1} \text{ cm}^{-1}$.

Site-Directed Mutagenesis

The site-directed mutant variants were generated using the QuikChange method of Stratagene. These protein variants were expressed and purified as described above for the wild-type enzyme.

Kinetic Analyses

The GlmA reaction was monitored spectrophotometrically at 412 nm by following the production of CoA through its reaction with DTNB. Reactions were conducted in 200 mM sodium phosphate, pH 7.3, in a total volume of 1 mL at 25 °C. Stock solutions of DTNB and GlcN were prepared fresh daily. Reactions were initiated by the addition of GlmA, and the rate of product formation was obtained from the rate of change in absorbance using a value of $14 \text{ 100 M}^{-1} \text{ cm}^{-1}$ for the extinction coefficient of 2-nitro-5-thiobenzoic acid at 412 nm. Data obtained in experiments in which only one substrate concentration was varied were fitted to the Michaelis–Menten equation; when both substrate concentrations were varied the data were fitted to eq 1, which describes a sequential mechanism. Initial velocity data obtained using GlcNAc as an inhibitor were fitted to eq 2 or 3, which describe noncompetitive and uncompetitive inhibition, respectively. In eq 1 *A* and *B* are substrate concentrations, V_{max} is the maximal velocity, *K* is the Michaelis constant for the substrate indicated, and K_i is the dissociation constant for the substrate indicated. The same definitions apply in eqs 2 and 3, and K_{is} and K_{ii} are the inhibition constants on the slope and intercept terms, respectively.

$$\nu = \frac{V_{\max}AB}{K_{iA}K_B + K_BA + K_AB + AB} \quad (1)$$

$$\nu = \frac{V_{\max}A}{K_A \left(1 + \frac{I}{K_{is}}\right) + A \left(1 + \frac{I}{K_{ii}}\right)} \quad (2)$$

$$\nu = \frac{V_{\max}A}{K_A + A \left(1 + \frac{I}{K_{ii}}\right)} \quad (3)$$

Relevant kinetic plots are provided in Supporting Information.

Crystallizations

Crystallization conditions for GlmA were surveyed by the hanging drop method of vapor diffusion using a laboratory-based sparse matrix screen. The enzyme was initially tested either in the presence of 5 mM CoA and 20 mM sugar or in the absence of any ligands. Two well diffracting crystal forms were identified from the crystallization experiments set up at 20 °C: one at pH values between 5–7 and one at pH values between 8–9. Both crystal forms appeared in the absence or presence of added ligands.

X-ray diffraction quality crystals of the protein in the absence of ligands were subsequently grown from precipitant solutions composed of 8–11% poly(ethylene glycol) 8000 and 100 mM Homo-PIPES (pH 5.0). These crystals belonged to the monoclinic space group $P2_1$ with unit cell dimensions of $a = 63.2 \text{ \AA}$, $b = 65.5 \text{ \AA}$, $c = 90.5 \text{ \AA}$, $\beta = 107.1^\circ$, and one dimer in the asymmetric unit. The crystals were prepared for X-ray data collection by serially transferring them to a cryoprotectant solution composed of 25% poly(ethylene glycol) 8000, 300 mM NaCl, 17% ethylene glycol, and 100 mM Homo-PIPES (pH 5.0).

The second crystal form of the protein in the absence of ligands was obtained from precipitant solutions composed of 7–10% poly(ethylene glycol) 5000 and 100 mM HEPPS (pH 8.0). These crystals belonged to the monoclinic space group C2 with unit cell dimensions of $a = 118.3 \text{ \AA}$, $b = 44.6 \text{ \AA}$, $c = 74.2 \text{ \AA}$, $\beta = 120.9^\circ$ and one subunit in the asymmetric unit. For X-ray data collection the crystals were serially transferred to a cryoprotectant solution composed of 25% poly(ethylene glycol) 5000, 300 mM NaCl, 15% ethylene glycol, and 100 mM HEPPS (pH 8.0). Selenomethionine labeled protein was prepared as previously described.¹³

Crystals of the GlmA/sugar complexes, both at pH 5.0 and 8.0, were obtained by cocrystallizing in the presence of 5 mM CoA and 20 mM glucosamine, *N*-acetylglucosamine, or galactosamine. Crystals of a GlmA complex with 5 mM CoA and 20

mM (GlcN)₂ (connected by a β -1 $\rightarrow$ 4 linkage) were obtained from poly(ethylene glycol) 8000 and 100 mM MOPS (pH 7.0).

The mutant protein variants were crystallized in a similar manner.

X-ray Data Collection, Processing and Structural Analyses

X-ray data sets were collected at either the Advanced Photon Source SBC-19BM or in-house. For in-house data collection, a Bruker AXS Platinum 135 CCD detector controlled with the APEX software suite was utilized (Bruker AXS Inc.). The X-ray source was Cu K α radiation from a Rigaku RU200 X-ray generator equipped with Montel optics and operated at 50 kV and 90 mA. The data sets were processed with SAINT and scaled with SADABS (Bruker AXS Inc.). Data collected at the synchrotron were processed and scaled with HKL3000.¹⁴ Relevant X-ray data collection statistics are listed in Table 1.

The initial structure of GlmA was determined via single wavelength anomalous dispersion (SAD) using crystals grown at pH 5.0 and X-ray data collected at the peak wavelength of 0.97919 Å. The locations of all 16 selenium atoms in the substructure were determined by SHELXD,¹⁵ which with 2-fold averaging and solvent flattening allowed for a complete tracing of the polypeptide chain. Fourier difference maps based on this model were employed to solve the other structures using crystals grown at pH 5.0. In addition, this model served as the search probe for molecular replacement with PHASER¹⁶ to determine the structures of the protein from the crystals grown at pH 8.0. All models were subjected to alternate cycles of refinement with REFMAC¹⁷ and manual model building with COOT.^{18,19}

Model refinement statistics for all structures are listed in Table 2.

RESULTS AND DISCUSSION

Overall Structure of GlmA. The initial GlmA structure determined in this investigation utilized crystals grown at pH 5.0. These crystals belonged to the space group $P2_1$ and contained a dimer in the asymmetric unit. The quality of the electron density for the polypeptide chain backbone, from Met 1 to Ala 316, was excellent. There was only one break between Ala 160 and Lys 165 in the second subunit of the dimer. Three amino acid residues, in each subunit, adopted dihedral angles outside of the allowed regions of the Ramachandran plot: Gln 144 ($\phi = \sim -118^\circ$, $\psi = \sim -140^\circ$), Lys 216 ($\phi = \sim 53^\circ$, $\psi = \sim -110^\circ$), and Lys 282 ($\phi = \sim -111^\circ$, $\psi = \sim -95^\circ$). The electron densities for the backbone atoms of these residues were unambiguous. All three are located near the protein surface. Lys 216 resides in the “nucleophile elbow” of the Ramachandran plot. This strained conformation was first observed in enzymes belonging to the α/β hydrolase fold.²⁰

The goal of this initial structural analysis was to define the architecture of the apo enzyme. Surprisingly, however, the electron density map demonstrated that each subunit contained two bound ligands. The first was acetyl CoA as evidenced by the electron density displayed in Figure 1a. The second was modeled as a pantothenic acid diphosphate moiety (Figure 1b). Clearly these ligands accompanied the protein during the purification process because they were never added to the crystallization trials. Regardless, this initial structure of GlmA

defined its overall three-dimensional architecture, albeit not that of the apo enzyme. In fact, it was never possible to prepare crystals of the apo enzyme after exhaustive attempts.

Shown in Figure 1c is a ribbon representation of the GlmA dimer. The molecule has overall dimensions of $\sim 80 \text{ \AA} \times 70 \text{ \AA} \times 70 \text{ \AA}$ and a total buried subunit/subunit interface of $2700 \text{ \AA}^2$. The two subunits are related by a rotation of 179.4° , and their α -carbons superimpose with a root-mean-square deviation of $0.3 \text{ \AA}$. Each subunit of the GlmA dimer contains 12 β -strands and eight α -helices and can be envisioned in terms of two domains that extend from Met 1 to Ser 155 and Gly 167 to Ala 316. These two domains are related by a rotation of 177° , and their α -carbons superimpose with a root-mean-square deviation of $2.5 \text{ \AA}$ for 124 target pairs. Both domains adopt the “GNAT” architectural motif of a mixed α,β -fold.^{21,22}

The superposition of the two domains, as shown in Figure 2, suggests that the present day version of GlmA most likely arose via gene duplication. The tandem GNAT fold displayed in GlmA has been previously observed, for example, in a yeast *N*-myristoyltransferase,^{23,24} in an *N*-acetyltransferase from *Mycobacterium tuberculosis*,²⁵ and in an enzyme involved in clavulanic acid biosynthesis.²⁶ Both domains in GlmA contain six β -strands and four α -helices, and in each domain the fourth β -strand contains a β -bulge resulting from the dihedral angles adopted by Asn 79 and Gly 251 in the N- and C-terminal domains, respectively. These β -strands lie near the CoA binding pockets. In addition, Pro 252 in the C-terminal domain adopts the *cis* conformation. Strikingly, however, there is less than 10% amino acid sequence identity between the N- and C-terminal domains.

Overall Structure of GlmA Determined at pH 8.0

GlmA is reportedly inactive at pH 5.0.¹² As such the next structure of the enzyme to be determined in this study was that using crystals grown at pH 8.0. These crystals belonged to the space group C2. The dimer packed in the unit cell with its 2-fold rotational axis coincident to a crystallographic dyad, thereby leading to one subunit in the asymmetric unit. Again, no exogenous ligands were added to the crystallization conditions, but as shown in Figure 3a,b there was clear electron density for acetyl CoA in the N-terminal domain and a fragment of CoA in the C-terminal domain. The α -carbons for the GlmA models determined at pH 5.0 and 8.0 superimpose with a root-mean-square deviation of $0.5 \text{ \AA}$. A perusal of the active site pockets for both models reveals no major adjustments in side chain positions. Most likely it is the protonation state of several residues lining the active site pocket and/or the ionization state of the amine sugar substrate that limits the activity of GlmA at pH 5.0. The only significant difference between the two models determined at pH 5.0 versus pH 8.0 is the orientation of the adenosine moieties of the CoA ligands bound in the N-terminal domains (Figure 3c). This conformational difference most likely arises from crystal packing effects.

Structure of GlmA in the Presence of Either Substrate or Product

The third structure determined for this investigation was that of the GlmA/CoA/GlcN abortive complex at pH 8.0. An acetyl CoA moiety was again bound in the N-terminal domain even though the crystals were grown in the presence of only CoA. In addition, only a fragment of the CoA ligand was observed binding in the C-terminal domain as indicated

by the electron density presented in Figure 4a. The electron density for the sugar moiety was unambiguous, however (Figure 4a).

A close-up view of the GlmA active site is presented in Figure 4b. The pyrophosphoryl moiety of the CoA fragment participates in hydrogen bonding interactions with the backbone amide groups of Gly 265 and Arg 266, and a water molecule. In addition, the positive end of a helix dipole moment projects toward the pyrophosphoryl group. Two additional hydrogen bonds occur between the CoA fragment and the carbonyl group of Ile 253 and the backbone amide of Ile 255. The aromatic side chains of Trp 193 and Trp 288 abut either side of the glucosamine ligand. The *cis*-proline, Pro 252, also lines the active site on the same side as Trp 193. The GlcN ligand adopts the ⁴C₁ conformation. Its C-1 hydroxyl group lies within hydrogen bonding distance to a HEPPS molecule observed binding in the active site. The C-2 amino moiety interacts with the carbonyl groups of Asp 287 and Trp 288 and a water molecule. The carboxylate of Asp 287 bridges the C-3 and C-4 hydroxyls of glucosamine. There is also an ordered water molecule that sits within 3.2 Å of the C-4 hydroxyl. Finally, the C-6 hydroxyl group lies within hydrogen bonding distance to the side chains of Trp 193 and Glu 196. The distance between the sulfur of the CoA fragment and the C-2 amino group is 3.2 Å.

To probe the structural changes that occur upon catalysis, the ternary complex of GlmA/CoA/GlcNAc was subsequently determined at pH 8.0. Shown in Figure 5 is the observed electron density corresponding to GlcNAc and the CoA fragment. Unfortunately, as can be seen, the β -mercaptoethylamine moiety of the CoA fragment was not visible in the electron density map so a complete model of the active site after catalysis could not be obtained from this structure. However, the α -carbons for the GlmA/CoA/GlcN and GlmA/CoA/GlcNAc complex models superimpose with a root-mean-square deviation of 0.2 Å, thereby indicating that no major structural changes occur upon catalysis.

Kinetic Analyses

Initial velocity kinetic studies were performed to determine the order of substrate addition and product release. An intersecting pattern was obtained when both acetyl CoA and GlcN concentrations were varied systematically, indicating that the enzyme follows a sequential kinetic mechanism. The kinetic parameters characterizing the GlmA reaction are given in Table 3. A consistent picture of the kinetic mechanism emerged from experiments using GlcNAc as a product inhibitor (Table 4). GlcNAc exhibited non-competitive inhibition versus GlcN whether acetyl CoA was saturating or subsaturating, which is the behavior expected for the second substrate to bind and the first product to be released. However, the prediction that acetyl CoA binds first and CoA leaves last could not be tested directly because the DTNB-based assay was not compatible with experiments using CoA as a product inhibitor. GlcNAc exhibited uncompetitive inhibition versus acetyl CoA when GlcN was present at subsaturating levels. Uncompetitive inhibition patterns are not predicted to occur with product inhibitors in simple two substrate sequential mechanisms and suggest that GlcNAc forms a dead-end complex with the enzyme-acetyl CoA complex. This interpretation was confirmed by testing whether GlcNAc inhibited the enzyme when GlcN was present at saturating concentration. No inhibition was observed, confirming that high

levels of GlcN prevented binding of GlcNAc. The kinetic mechanism deduced from the kinetic studies is shown in Scheme 2. The proposal that acetyl CoA is the first substrate to bind and that CoA is the last product to leave is consistent with the repeated observations that GlmA crystallized with CoA or acetyl CoA bound, even when they were not added to the crystallization solutions.

It is not clear whether general acid/base catalysis is required for acetylation of the sugar amine. A reaction mechanism for GlmA can be envisioned, however, whereby a catalytic base might be required to promote the removal of a hydrogen from the sugar amino group as it attacks the carbonyl carbon of acetyl CoA leading to a tetrahedral intermediate. Collapse of this intermediate could be assisted by the presence of a catalytic acid to protonate the sulfur group of the cofactor. On the other hand, the neutral amine could act as the nucleophile toward the carbonyl, generating a cationic intermediate that deprotonates spontaneously, and the CoA thiolate may be an adequate leaving group.²⁷ In the absence of structures with intact acetyl CoA bound at the active site, it was difficult to evaluate whether any residues were positioned appropriately to play roles as general acid/base catalysts. Given that its N-terminal domain always binds acetyl CoA, however, a putative model of the Michaelis complex was subsequently constructed by superimposing the N- and C-terminal domains to give an approximation for where acetyl CoA binds in the C-terminal catalytic domain (data not shown). From this model, the side chain of Asp 287 appears to be the only potential catalytic base near the C-2 amino group, albeit at ~4.6 Å. In addition, the side chain of Tyr 297 projects into the active site pocket in a manner that suggests it might function as the catalytic acid.

Site-directed Mutagenesis Experiments

In order to test the roles of these amino acids in GlmA catalysis, the D287N and Y297F mutant proteins were constructed, and their kinetic parameters were determined (Table 5). The most striking aspect of the kinetic data for the GlmA variants is the negligible change (2-fold) in k_{cat} . The three-dimensional structures of the D287N and Y297F variants were also determined to ensure that any observed kinetics effects were due to the mutation and not to gross conformational changes. In the case of the D287N variant, the α -carbons for the wild-type and the mutant protein superimpose with a root-mean-square deviation of 0.2 Å. Within experimental error the two structures are identical. Again, only a fragment of CoA was observed binding in the C-terminal domain. The situation with the Y297F variant was more intriguing. Again, the polypeptide chains for the mutant protein and the wild-type enzyme were nearly identical. Indeed, the α -carbons superimpose with a root-mean-square deviation of 0.2 Å. Excitingly, however, for the first time electron density was clearly visible for the acetyl group of the cofactor bound in the C-terminal domain (Figure 6a). The backbone amide group of Ile 253 lies within 3.2 Å of the carbonyl group of the acetyl moiety, and thus it might play a role in stabilization of the presumed tetrahedral intermediate.

By superimposing the Y297F structure onto the coordinates for the GlmA/CoA/GlcN complex, it was possible to put forth a model for the Michaelis complex as depicted in Figure 6b. The amino group of the sugar is ideally situated to conduct a nucleophilic attack

on the re face of the carbonyl carbon of acetyl CoA. Asp 287 is positioned with its side chain carboxyl over 4.5 Å from the amino group of the sugar, too great a distance for it to serve as a general base in the reaction, unless it moves closer during the catalytic cycle. The mutagenesis results also suggest that it does not act as a general base in the catalytic reaction, since substitution of Asp 287 with an asparagine residue diminishes k_{cat} by only 2-fold. The $\text{p}K$ of the amino group in glucosamine is 7.6, so the reaction may proceed by nucleophilic attack of the neutral amine on acetyl CoA. An attractive role for Asp 287 is in facilitating the binding of glucosamine. It is poised to form hydrogen bonds with the C-3 and C-4 hydroxyl groups of the sugar. Substitution of aspartate with asparagine places a hydrogen bond donor where a hydrogen bond acceptor is required, which presumably is the basis for the observed 90-fold increase in K_{m} .

Tyr 297 is not positioned to interact directly with glucosamine, and the k_{cat} in the Y297F protein is decreased by only 2-fold, arguing against assigning a critical role as a general acid catalyst to Tyr 297. The most important kinetic parameter for interpreting the behavior of the Y297F protein may be $k_{\text{cat}}/K_{\text{m}}$, which reflects the efficiency with which the GlmA·acetyl CoA complex combines with glucosamine to form a productive complex. The value of the $k_{\text{cat}}/K_{\text{m}}$ is decreased more than 100-fold in the Y297F protein, which suggests that productive binding of glucosamine requires proper binding of acetyl CoA, which may be perturbed by the substitution. It is unclear, however, why the substitution of Tyr 297 with a phenylalanine residue allowed the acetyl moiety of acetyl CoA to be visualized in the active site.

Structure of GlmA in the Presence of GalN

Previous investigations on GlmA have indicated that it cannot function on GalN.¹² The structure of the GlmA/CoA/GalN ternary complex was subsequently determined to address why GalN cannot serve as a substrate. Shown in Figure 7a is the electron density observed for the ligands in the C-terminal domain. The polypeptide chains for the complexes of GlmA with GlcN versus GalN are nearly identical such that their α -carbons superimpose with a root-mean-square deviation of 0.1 Å. The GalN ligand, however, adopts a quite strained chair conformation and is pushed out of the active site (Figure 7b). As a consequence, its C-2 amino group is positioned ~1 Å away from that observed for the C-2 amino group in GlcN. This difference is most likely the reason why GalN is not a substrate for GlmA. In fact, galactosamine at a concentration of 5 mM did not cause any inhibition of the GlmA reaction when glucosamine was present at 0.1 mM. In light of the apparently very low affinity that GlmA exhibits toward galactosamine, it is remarkable that it was observed in the crystal structure.

Structure of GlmA in the Presence of (GlcN)₂

GlmA has been shown to function on (GlcN)₂ with the sugar residues joined together by a β -(1–4) linkage. Indeed, the enzyme shows a reduced V_{max} for GlcN alone as compared to (GlcN)₂ or (GlcN)₃.¹² The final structure determined in this investigation was that of the GlmA/CoA/(GlcN)₂ ternary complex. Shown in Figure 8a is the observed electron density for the ligands in the C-terminal domain. The reducing end of the disaccharide adopts the α -anomeric configuration. Both sugars assume the ⁴C₁ pucker. A close-up view of the

hydrogen bonding pattern surrounding the disaccharide is presented in Figure 8b with the positions of the sugar and HEPPS molecules observed in the GlmA/CoA/GlcN complex superimposed. The disaccharide is positioned into the active site via hydrogen bonding interactions with the side chains of Glu 196, Asp 287, and Glu 290, the backbone carbonyl group of Trp 288, and five water molecules. The amino group of the terminal nonreducing GlcN ligand is shifted by ~ 1.2 Å relative to that observed for GlcN alone. The pyranosyl group at the reducing end of the disaccharide fills in the region occupied in other GlmA structures by a HEPPS molecule. The GlmA binding pocket for the sugars is noticeably wide (Figure 8c), which explains the enzyme's ability to function on GlcN-containing muropeptides and on β -1,4-linked chitosan oligomers (such as (GlcN)₃).¹²

Comparison of GlmA with MshD

A search of the Protein Data Bank utilizing PDBfold²⁸ reveals that GlmA and MshD from *M. tuberculosis* demonstrate significant three-dimensional similarity albeit GlmA functions as a dimer and MshD assumes a monomeric structure. Specifically, the α -carbons for the two enzymes superimpose with a root-mean-square deviation of 2.9 Å for 254 targeted pairs. The amino acid sequence identity and similarity between these two proteins is not obvious, however. MshD catalyzes the last step in the production of mycothiol, a key molecule involved in cell protection against oxidative stress. Mycothiol is composed of an acetylated cysteine residue, a glucosamine residue, and an inositol group. Shown in Figure 9a is a superposition of the ribbon drawings for GlmA and MshD with their associated ligands ((GlcN)₂ and desacetylmycothiol, respectively). The observed positions for the cofactors residing in the N-terminal domains of both proteins are similar until the pantetheine groups. Differences in the torsional angles within the pantetheine moieties, however, result in the acetyl groups of the cofactors being rotated by nearly 90° relative to one another. The positions of the CoA fragments in the C-terminal domains are more similar. The main difference between these two proteins lies in the orientations of the amino-containing substrates. Whereas in GlmA the disaccharide substrate projects outward toward the solvent, in MshD the ligand is buried between the two domains of the subunit. Indeed, binding of the substrate in MshD invokes a large conformational change, which is not observed in GlmA. A close-up view of the substrate binding orientations for GlmA and MshD is presented in Figure 9b. Given the lack of sequence homology between the two proteins, it is not surprising that there are numerous amino acid substitutions that preclude the two substrates from binding in similar orientations. In particular, the conformations of two loops in GlmA, namely, those connecting β -strand 5 to α -helix 4 and β -strand 9 to β -strand 10, fill in the pocket required for desacetylmycothiol binding.

CONCLUSION

Thus far only a handful of X-ray structures have been determined for members of the tandem GNAT superfamily. Furthermore, many of the proteins whose coordinates have been deposited in the Protein Data Bank have not been biochemically analyzed. The molecular analysis of GlmA described here thus provides significantly new information on this intriguing superfamily of bacterial *N*-acetyltransferases. In particular, the various ternary complexes of GlmA show that the mode of binding of the sugar substrates is quite different

to that observed for MshD. Given the size of the bacterial “acetylome,” and the fact that there is limited sequence homology between the various *N*-acetyltransferases, it is clear that additional structural and functional analyses of these enzymes will be required to more fully understand their biochemical characteristics.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

A portion of the research described in this paper was performed at Argonne National Laboratory, Structural Biology Center at the Advanced Photon Source (U.S. Department of Energy, Office of Biological and Environmental Research, under Contract DE-AC02-06CH11357). We gratefully acknowledge Dr. Norma E. C. Duke for help in the X-ray data collection.

Funding

This research was supported in part by an NIH grant (GM115921 to H.M.H.).

ABBREVIATIONS

CoA	coenzyme A
DTNB	5,5'-dithio-bis(nitrobenzoic acid)
GalN	galactosamine
GlcN	glucosamine
GlcNAc	<i>N</i> -acetyl-glucosamine
HEPPS	<i>N</i> -2-hydroxyethylpiperazine- <i>N'</i> -3-propanesulfonic acid
Homo-PIPES	homopiperazine-1,4-bis(2-ethanesulfonic acid)
MOPS	3-(<i>N</i> -morpholino)-propanesulfonic acid
Ni-NTA	nickel nitrilotriacetic acid
Tris	<i>tris</i> -(hydroxymethyl)aminomethane

REFERENCES

- (1). Vollmer W, Blanot D, de Pedro MA. Peptidoglycan structure and architecture. FEMS Microbiol Rev. 2008; 32:149–167. [PubMed: 18194336]
- (2). Dramsi S, Magnet S, Davison S, Arthur M. Covalent attachment of proteins to peptidoglycan. FEMS Microbiol Rev. 2008; 32:307–320. [PubMed: 18266854]
- (3). Neuhaus FC, Baddiley J. A continuum of anionic charge: structures and functions of D-alanyl-teichoic acids in gram-positive bacteria. Microbiol Mol. Biol. Rev. 2003; 67:686–723. [PubMed: 14665680]
- (4). Ohno N, Yadomae T, Miyazaki T. Identification of 2-amino-2-deoxyglucose residues in the peptidoglycan of *Streptococcus pneumoniae*. Carbohydr. Res. 1982; 107:152–155. [PubMed: 7139657]

- (5). Araki Y, Nakatani T, Hayashi H, Ito E. Occurrence of non-N-substituted glucosamine residues in lysozyme-resistant peptidoglycan from *Bacillus cereus* cell walls. *Biochem. Biophys. Res. Commun.* 1971; 42:691–697. [PubMed: 4993454]
- (6). Vollmer W, Tomasz A. Peptidoglycan *N*-acetylglucosamine deacetylase, a putative virulence factor in *Streptococcus pneumoniae*. *Infect. Immun.* 2002; 70:7176–7178. [PubMed: 12438406]
- (7). Wang G, Maier SE, Lo LF, Maier G, Dosi S, Maier RJ. Peptidoglycan deacetylation in *Helicobacter pylori* contributes to bacterial survival by mitigating host immune responses. *Infect. Immun.* 2010; 78:4660–4666. [PubMed: 20805339]
- (8). Fisher JF, Mobashery S. The sentinel role of peptidoglycan recycling in the beta-lactam resistance of the Gram-negative Enterobacteriaceae and *Pseudomonas aeruginosa*. *Bioorg. Chem.* 2014; 56:41–48. [PubMed: 24955547]
- (9). Park JT, Uehara T. How bacteria consume their own exoskeletons (turnover and recycling of cell wall peptidoglycan). *Microbiol Mol. Biol. Rev.* 2008; 72:211–227. [PubMed: 18535144]
- (10). Johnson JW, Fisher JF, Mobashery S. Bacterial cell-wall recycling. *Ann. N. Y. Acad. Sci.* 2013; 1277:54–75. [PubMed: 23163477]
- (11). Litzinger S, Fischer S, Polzer P, Diederichs K, Welte W, Mayer C. Structural and kinetic analysis of *Bacillus subtilis* *N*-acetylglucosaminidase reveals a unique Asp-His dyad mechanism. *J. Biol. Chem.* 2010; 285:35675–35684. [PubMed: 20826810]
- (12). Reith J, Mayer C. Characterization of a glucosamine/glucosaminide *N*-acetyltransferase of *Clostridium acetobutylicum*. *J. Bacteriol.* 2011; 193:5393–5399. [PubMed: 21784938]
- (13). Thoden JB, Holden HM. Molecular structure of galactokinase. *J. Biol. Chem.* 2003; 278:33305–33311. [PubMed: 12796487]
- (14). Minor W, Cymborowski M, Otwinowski Z, Chruszcz M. HKL-3000: the integration of data reduction and structure solution-from diffraction images to an initial model in minutes. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 2006; 62:859–866. [PubMed: 16855301]
- (15). Schneider TR, Sheldrick GM. Substructure solution with SHELXD. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 2002; 58:1772–1779. [PubMed: 12351820]
- (16). McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *J. Appl. Crystallogr.* 2007; 40:658–674. [PubMed: 19461840]
- (17). Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 1997; 53:240–255. [PubMed: 15299926]
- (18). Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 2004; 60:2126–2132. [PubMed: 15572765]
- (19). Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 2010; 66:486–501. [PubMed: 20383002]
- (20). Nardini M, Dijkstra BW. Alpha/beta hydrolase fold enzymes: the family keeps growing. *Curr. Opin. Struct. Biol.* 1999; 9:732–737. [PubMed: 10607665]
- (21). Vetting MW, de Carvalho LPS, Yu M, Hegde SS, Magnet S, Roderick SL, Blanchard JS. Structure and functions of the GNAT superfamily of acetyltransferases. *Arch. Biochem. Biophys.* 2005; 433:212–226. [PubMed: 15581578]
- (22). Favrot L, Blanchard JS, Vergnolle O. Bacterial GCN5-Related *N*-acetyltransferases: from resistance to regulation. *Biochemistry.* 2016; 55:989–1002. [PubMed: 26818562]
- (23). Weston SA, Camble R, Colls J, Rosenbrock G, Taylor I, Egerton M, Tucker AD, Tunnicliffe A, Mistry A, Mancia F, de la Fortelle E, Irwin J, Bricogne G, Pauptit RA. Crystal structure of the anti-fungal target *N*-myristoyl transferase. *Nat. Struct. Biol.* 1998; 5:213–221. [PubMed: 9501915]
- (24). Bhatnagar RS, Futterer K, Farazi TA, Korolev S, Murray CL, Jackson-Machelski E, Gokel GW, Gordon JI, Waksman G. Structure of *N*-myristoyltransferase with bound myristoylCoA and peptide substrate analogs. *Nat. Struct. Biol.* 1998; 5:1091–1097. [PubMed: 9846880]
- (25). Vetting MW, Roderick SL, Yu M, Blanchard JS. Crystal structure of mycothiol synthase (Rv0819) from *Mycobacterium tuberculosis* shows structural homology to the GNAT family of *N*-acetyltransferases. *Protein Sci.* 2003; 12:1954–1959. [PubMed: 12930994]

- (26). Iqbal A, Arunlanantham H, Brown T Jr, Chowdhury R, Clifton IJ, Kershaw NJ, Hewitson KS, McDonough MA, Schofield CJ. Crystallographic and mass spectrometric analyses of a tandem GNAT protein from the clavulanic acid biosynthesis pathway. *Proteins*. 2010; 78:1398–1407. [PubMed: 20014241]
- (27). Cordes EH, Jencks WP. General acid catalysis of semicarbazone formation. *J. Am. Chem. Soc.* 1962; 84:4319–4328.
- (28). Krissinel E, Henrick K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 2004; 60:2256–2268. [PubMed: 15572779]
- (29). Laskowski RA, Moss DS, Thornton JM. Mainchain bond lengths and bond angles in protein structures. *J. Mol. Biol.* 1993; 231:1049–1067. [PubMed: 8515464]
- (30). DeLano, WL. The PyMOL Molecular Graphics System. DeLano Scientific; San Carlos, CA, USA: 2002.

Figure 1.

Structure of GlmA determined at pH 5.0. Shown in (a) is the observed electron density for acetyl CoA binding in the N-terminal domain. The electron density corresponding to the CoA fragment in the C-terminal domain is displayed in (b). Both maps, contoured at $\sim 3\sigma$, were calculated with coefficients of the form $(F_0 - F_c)$, where F_0 was the native structure factor amplitude and F_c was the calculated structure factor amplitude. Importantly, the maps were calculated before the ligands had been included in the X-ray coordinate file, and thus there is no model bias. A ribbon representation of the GlmA dimer is depicted in (c) with subunits 1 and 2 colored in cyan and purple, respectively. The ligands are drawn in space filling representations. All figures were created with PyMOL.³⁰

Figure 2. Superposition of the N- and C-terminal domains of GlmA. The N- and C-terminal domains are highlighted in purple and gray, respectively.

Figure 3.

Structure of GlmA determined at pH 8.0. Electron densities observed for the cofactors in the N- and C-terminal domains are shown in (a) and (b), respectively. The omit maps were calculated as described in the legend to Figure 1. The adenosine 3-phosphate groups of the cofactors in the N-terminal domains of GlmA solved at pH 5.0 versus pH 8.0 are rotated by nearly 180° as can be seen in (c). These different orientations are most likely due to crystal packing effects.

Figure 4.

Structure of the GlmA/CoA/GlcN abortive complex. An omit map showing the electron density for the CoA fragment and the GlcN ligand is displayed in (a). A close-up view of the active site in the C-terminal domain is provided in (b). The CoA fragment, the sugar molecule, and the HEPPS buffer are depicted in purple bonds. Possible hydrogen bonding interactions are indicated by the dashed lines. Water molecules are represented by the red spheres.

Figure 5. Observed electron density in the C-terminal domain for the GlmA/CoA/GlcNAc abortive complex. An omit map showing the electron density for the CoA fragment and the GlcNAc ligand is shown.

Figure 6.

Model of the Michaelis complex. In the Y297F mutant protein structure, electron density was observed for the acetyl group of the cofactor as can be seen in (a). By superimposing the coordinates for the GlmA/CoA/GlcN complex and the Y297F mutant protein a model of the Michaelis complex was built as depicted in (b). The C-2 amino group is ideally positioned to attack the re face of the acetyl group as indicated by the dashed line. Likewise, the side chain of Tyr 297 is located near the sulfur atom of acetyl CoA.

Figure 7. Structure of the GlmA/CoA/GalN complex. The observed electron density in the C-terminal domain for the cofactor and the GalN moiety is presented in (a). A superposition of the GlmA/CoA/GlcN and GlmA/CoA/GalN complexes is displayed in (b). The GalN and GlcN groups are drawn in purple and gray, respectively. The GalN ligand, due to its axial configuration about C-4, is shifted out of the GlmA active site.

Figure 8.

Structure of the GlmA/CoA/(GlcN)₂ complex. Electron density observed in the C-terminal domain for the ligands is presented in (a). A close-up view of the region surrounding the (GlcN)₂ ligand is depicted in (b) with potential hydrogen bonds indicated by the dashed lines. Superimposed are the binding locations for the GlcN and HEPPS molecules observed in the GlmA/CoA/GlcN model. A surface representation of the GlmA dimer with the bound disaccharide is displayed in (c). There is a wide opening extending from the active site to the solvent that allows the enzyme to function on larger substrates including (GlcN)₃.

Figure 9. Comparison of GlmA and MshD. A superposition of the ribbon drawings for GlmA (purple) and MshD (white) is shown in (a). The ligands are drawn as sticks. A close-up view of the substrate binding regions for these two proteins is given in (b). The (GlcN)₂ substrate is highlighted in purple whereas the desacetylmethylol ligand (abbreviated DAM) is colored in green.

Scheme 1.

Scheme 2.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

X-ray Data Collection Statistics

	Apo pH 5.0	Apo pH 8.0	GlcN pH 8.0	GlcNAc pH 8.0	D287N pH 5.0	Y297F pH 5.0	GalN pH 8.0	(GlcN) ₂ pH 7.0
resolution limits (Å)	50-2.0 (2.1-2.0) ^b	50-1.9 (2.0-1.9) ^b	50-1.9 (2.0-1.9) ^b	50-1.5 (1.55-1.5) ^b	50-1.9 (2.0-1.9) ^b	50-1.8 (1.9-1.8) ^b	50-1.9 (2.0-1.9) ^b	50-1.8 (1.9-1.8) ^b
number of independent reflections	46396 (5725)	25764 (3583)	25315 (3253)	53438 (5337)	54576 (7506)	66070 (9489)	25686 (3480)	63868 (8901)
completeness (%)	95.4 (87.2)	99.0 (97.1)	95.1 (86.4)	99.3 (99.8)	98.5 (95.6)	98.1 (94.6)	97.9 (93.6)	96.6 (90.5)
redundancy	3.5 (1.8)	4.7 (2.8)	2.3 (1.4)	4.8 (3.8)	4.6 (2.5)	6.6 (3.0)	4.4 (2.5)	3.4 (2.0)
avg I/avg σ(I)	10.0 (1.9)	10.9 (2.4)	9.5 (2.6)	48.4 (6.3)	7.3 (2.1)	11.3 (2.2)	11.6 (2.4)	10.3 (2.4)
R _{sym} (%) ^a	7.0 (30.9)	7.5 (32.5)	5.3 (20.1)	6.6 (37.7)	8.0 (38.9)	8.0 (37.9)	7.2 (30.2)	7.0 (28.2)

$$^a R_{\text{sym}} = \frac{\sum I}{\sum I} \times 100$$

^b Statistics for the highest resolution bin.

Table 2

Refinement Statistics

	Apo pH 5.0	Apo pH 8.0	GlcN pH 8.0	GlcNAc pH 8.0	D287N pH 5.0	Y297F pH 5.0	GalN pH 8.0	(GlcN) ₂ pH 7.0
resolution limits (Å)	50-2.0	50-1.9	50-1.9	50-1.50	50-1.9	50-1.80	50-1.90	50-1.80
<i>R</i> -factor ^a (overall)%/no. reflections	18.7/46385	17.4/25762	16.3/25315	17.6/53401	17.3/54576	20.7/66061	17.0/25685	15.9/63853
<i>R</i> -factor (working)%/no. reflections	18.4/44132	17.1/24463	16.0/24057	17.4/50753	17.0/51862	20.5/62715	16.7/24410	15.7/60566
<i>R</i> -factor (free)%/no. reflections	22.8/2253	23.1/1299	21.0/1258	20.6/2648	22.4/2714	24.9/3346	21.5/1275	19.8/3287
number of protein atoms	5207	2638	2635	2642	5242	5205	2636	5258
number of heteroatoms	404	261	314	370	797	584	292	756
average B values								
protein atoms (Å ²)	33.1	24.7	19.9	20.3	19.3	26.0	22.4	18.7
ligand (Å ²)	33.5	21.3	16.9	19.2	20.9	26.7	21.7	19.7
solvent (Å ²)	41.0	30.4	25.6	30.7	29.0	30.6	26.3	26.1
weighted RMS deviations from ideality								
bond lengths (Å)	0.014	0.011	0.013	0.013	0.009	0.009	0.012	0.013
bond angles (deg)	1.75	2.00	2.00	2.00	1.85	1.80	1.85	1.80
planar groups (Å)	0.007	0.008	0.009	0.010	0.007	0.007	0.008	0.008
Ramachandran regions (%)^b								
most favored	89.9	90.3	91.0	90.6	89.2	89.4	89.9	89.1
additionally allowed	9.5	8.7	8.7	8.3	10.2	10.0	9.4	10.3
generously allowed	0.2	0.7	0.3	0.7	0.2	0.2	0.7	0.2
disallowed	0.4	0.3	0	0.4	0.4	0.4	0	0.4

^a *R*-factor = $(|F_O - F_C| / |F_O|) \times 100$ where F_O is the observed structure-factor amplitude and F_C is the calculated structure-factor amplitude.

^b Distribution of Ramachandran angles according to PROCHECK.²⁹

Table 3
Steady-State Kinetic Parameters for the GlmA Reaction^a

k_{cat}	$2.5 \pm 0.1 \text{ s}^{-1}$
$k_{\text{cat}}/K_{\text{AcCoA}}$	$2.1 (\pm 0.2) \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$
$k_{\text{cat}}/K_{\text{GlcN}}$	$1.2 (\pm 0.1) \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$
K_{AcCoA}	$12 \pm 1 \text{ }\mu\text{M}$
K_{GlcN}	$210 \pm 20 \text{ }\mu\text{M}$
$K_{\text{i,AcCoA}}$	$3 \pm 1 \text{ }\mu\text{M}$

^aThe values of the parameters and errors shown were determined by fitting the data to eq 1. The term $K_{\text{i,AcCoA}}$ corresponds to the K_{iA} term in eq 1 and describes the dissociation constant of AcCoA from GlmA.

Table 4

Inhibition Patterns Obtained with GlcNAc

variable substrate	fixed substrate, mM	pattern
GlcN	AcCoA, 0.015	NC
GlcN	AcCoA, 0.75	NC
AcCoA	GlcN, 0.15	UC
AcCoA	GlcN, 7.5	<i>a</i>

^aNo inhibition observed.

Table 5

Kinetic Parameters of GlmA Wild-Type and Mutant Proteins

protein	K_{GlcN} (mM)	k_{cat} (s^{-1})	$k_{\text{cat}}/K_{\text{GlcN}}$ ($\text{M}^{-1} \text{s}^{-1}$)
wild-type	0.13 ± 0.02	2.8 ± 0.1	$2.1 (\pm 0.3) \times 10^4$
D287N	11.8 ± 0.5	1.05 ± 0.03	89 ± 5
Y297F	7.1 ± 0.5	1.30 ± 0.06	180 ± 20