 doi:10.20944/preprints202109.0339.v1

Article

Towards computationally guided design and engineering of a *Neisseria meningitidis* serogroup W capsule polymerase with altered substrate specificity

Subhadra Paudel [1], James Wachira [2,*] and Pumtiwitt C. McCarthy [3,*]

[1] Department of Computer Science; subhra.paudel@gmail.com
[2] Department of Biology; James.Wachira@morgan.edu
[3] Department of Chemistry; Pumtiwitt.McCarthy@morgan.edu
* Correspondence: James.Wachira@morgan.edu and Pumtiwitt.McCarthy@morgan.edu

Abstract: Heavy metal contamination of drinking water is a public health concern that requires the development of more efficient bioremediation techniques. Absorption technologies, including biosorption, provide opportunities for improvements to increase the diversity of metal ions removed and overall binding capacity. Microorganisms are a key component in wastewater treatment plants and they naturally bind metal ions through surface macromolecules but with limited capacity. The long-term goal of this work is to engineer capsule polymerases to synthesize molecules with novel functionalities. In previously published work, we showed that the *Neisseria meningitidis* serogroup W (NmW) galactose-sialic acid (Gal—NeuNAc) heteropolysaccharide binds lead effectively, thereby demonstrating the potential for using this capsular polysaccharide in environmental decontamination applications. In this study, computational analysis of the NmW capsule polymerase galactosyltransferase (GT) domain was used to gain insight into how the enzyme could be modified to enable the synthesis N-acetylgalactosamine-sialic acid (GalNAc—NeuNAc) heteropolysaccharide. Various computational approaches, including molecular modeling with I-TASSER and molecular dynamics simulations (MD) with NAMD, were utilized to identify key amino acid residues in the substrate binding pocket of the GT domain that may be key to conferring UDP-GalNAc specificity. Through these combined strategies and using BshA, a UDP-GlcNAc transferase, as a structural template, several NmW active site residues were identified as mutational targets to accommodate the proposed N-acetyl group in UDP-GalNAc. Thus, a rational approach for potentially conferring new properties to bacterial capsular polysaccharides is demonstrated.

Keywords: Polysaccharide biomaterials; capsule polymerases; galactosyltransferase; molecular dynamics simulations; bioremediation; protein engineering; *Neisseria meningitidis*

1. Introduction

One of the environmental implications of industrialization has been the contamination of water sources with heavy metals either directly (i.e., industrial wastes) or indirectly (i.e., through contaminated soils) [1]. Biosorption, which is the use of microorganisms and/or isolated microbial biomolecules as metal adsorption materials, could provide an environmentally-safer route for metal decontamination of water and several microorganisms and their polysaccharides have been investigated for the ability to bind metals [2,3]. Polysaccharides can be present as a component of the extracellular polymeric substance (EPS) associated with the surface of the bacteria [4]. In addition, some bacteria, particularly Gram-negative bacteria, have capsular polysaccharides that are firmly attached to the outer surface [4,5]. Bacterial capsules are relevant to the understanding of microbial physiology and genetics as well as bacterial host interactions [4,5]. Capsular polysaccharides exhibit great diversity between bacteria in terms of the constituent monosaccharides, the glycosidic linkages, and chemical modifications. Thus, they offer a rich source of

 doi:10.20944/preprints202109.0339.v1

diverse polysaccharides that could be further developed as biomaterials with varied applications [6]. This diversity underlies the pathogenicity of the Gram-negative bacteria *Neisseria meningitidis* [7] and our focus is on the capsular polysaccharide of serogroup W (NmW) [8,9].

In previous work, the lead-binding affinity of NmW capsular polysaccharides was demonstrated [10]. Solutions of lead were incubated in the absence of or presence of NmW polysaccharide (PS) and subsequently passed through a filtration device. The metal solution without NmW PS present was able to pass freely through the filtration device leading to equal concentrations of metal in filtrate and retentate. However, in the presence of PS, less metal was present in the filtrate indicating binding to the polysaccharide. From these promising results, we have embarked on exploring new ways to broaden the nucleotide-donor substrate specificity of the NmW capsule polymerase to create novel structures with enhanced metal-binding properties.

The NmW capsule polymerase, encoded by the SiaDw gene, is a 120 kDa glycosyltransferase comprised of an N-terminal GT domain, an intervening sequence, and a C-terminal sialyltransferase domain [11-15]. The two catalytic domains differ in the utilization of nucleotide-monosaccharide donors with GT domain using UDP-Gal and the sialyltransferase domain using CMP-NeuNAc. The N-terminal galactosyltransferase domain is member of the GT4 CAZy family, while the C-terminal sialyltransferase domain is within the newly classified GT97 CAZy family [13,16]. Mechanistically, GT4 enzymes transfer sugars with retention of stereochemistry at the anomeric carbon and GT97 family members transfer with inversion of stereochemistry at that position. For many of these enzymes, mechanisms are inferred from biochemical and mutational studies with further insight being provided by molecular dynamics simulations [17].

The structures of several members of the BshA family, also classified as GT4 family enzymes, are available. These include family members from pathogenic and non-pathogenic species such as *Bacillus subtilis*, *Bacillus anthracis*, *Staphylococcus epidermis*, and *Staphylococcus aureus* [18-20]. The crystal structures reveal the active site and interactions between the UDP and hexose moieties with specific amino acid residues. These studies revealed the basis for the proposed substrate assisted reaction mechanism and identified the putative residues involved in the catalytic mechanism [19]. Still, questions remain about the chemical events leading to catalysis.

The studies described here report on sequence analysis, homology modeling and molecular dynamics of the NmW GT. Our results have identified several "putative" ligand binding residues that confer substrate specificity. Simulations and molecular dockings could guide the engineering of glycosyltransferases for the synthesis of novel materials.

2. Materials and Methods

2.1. Sequence Alignments

The genomic sequence of *N. meningitidis* serogroup W capsule polymerase (NmW) (SiaDw) was obtained from NCBI [21] (accession no. ABW93688). The full-length sequence is 1037 aa residues and this work focuses on the residues 1-399, which contains the galactosyltransferase domain. The following sequences that encode for β3-GlcNAc-transferases were selected for alignments: LgtA (NCBI accession no. AAL12840) a β3-GlcNAc-transferase from *Neisseria meningitidis* [22-26] and BoGT56a (PDB code 4AYL), an alpha 3-GlcNAc-transferase from *Bacteroides ovatus* [27]. Alignments were performed using BLAST and EMBL-EBI Clustal Omega (ClustalO) [28,29]. All the BLAST parameters were set to default during the BLAST run. The ClustalO multiple sequence alignments were obtained under the default parameters.

2.2. Homology Modeling

The galactosyltransferase domain of the NmW capsule polymerase was modeled with I-TASSER [30]. The output from I-TASSER provided homology model prediction, ligand binding sites and active sites. For the subsequent MD simulations and docking

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 September 2021 doi:10.20944/preprints202109.0339.v1

studies, the best homology model was chosen based on quality metrics. The best template structure was the glycosyltransferase BshA from *Staphylococcus aureus* complexed with UDP (PDB code: 6d9t) [19]. Its RMSD value was 0.897, with the highest TM-score of 0.76. MolProbity assessment [31-33] was used to determine the model quality.

2.3. Molecular Dynamics Simulation

The systems for molecular dynamics simulations were prepared with CHARMM-GUI [34-37]. To simulate the apoenzyme state, ligand molecules were deleted from the template PDB structure (6d9t) using UCSF-Chimera [38] prior to uploading to the CHARMM-GUI webserver. The systems were prepared essentially using the default parameters. The protein was solvated with TIP3P water and neutralized with potassium chloride [39,40]. A water boundary of 10.0 Å and the KCl concentration 0.15M were used in the simulations.

The simulation steps (minimization, equilibration, and production runs) were performed sequentially with NAMD [41,42]. The all-atom molecular dynamics simulations were conducted using default NAMD parameters including Particle Mesh Ewald electrostatics calculations [43]. Equilibration was conducted over six steps as follows: 10ps of minimization followed by a 1.5ns equilibration steps that were conducted with Langevin dynamics parameters, the default CHARMM36 all-atom additive protein force field [44], and a constant temperature of 303.15K. Equilibration was conducted with restraints that were gradually reduced over time. Production simulation was conducted for 99ns in 1ns blocks each with a 2fs timestep. The simulations were conducted on XSEDE-PSC-Bridges supercomputer resources [45].

2.5. Data Analysis and Visualization

UCSF-Chimera [38], VMD (Visual Molecular Dynamics) [46], and Bio3D [47] were used for MD simulations trajectory analysis.

3. Results

3.1. Sequence Comparison of NmW capsule polymerase, LgtA, and 4AYL

The basis for the substrate specificities of the capsule polymerases from N. meningiditis serogroup Y and W towards UDP-Glc or UDP-Gal, respectively, has been mapped to a single amino acid residue in the conserved Ex_7E motif of GT4 family of glycosyltransferases [11]. Given our goal of engineering specificity for N-acetylgalactosamine, the similarity of NmW GT domain with other bacterial GalNAc transferases was investigated. Because of carbohydrate stereochemistry, glycosyltransferases can transfer sugars to an acceptor to create either α or β glycosyl bonds [48]. The NmW capsule polymerase synthesizes α-glycosidic bonds for both GT and ST activities. As such, we focused on two enzymes, LgtA (a β3-GlcNActransferase) and BoGT56a (an α3-GalNActransferase). We performed a sequence comparison of these three sequences using Clustal Omega. There was minimal sequence identity between the three sequences and only nine amino acid residues aligned identically (**Figure 1**). Previous research has shown that mutation of Pro310 to Gly in the NmW GT domain can switch specificity from UDP-Gal to UDP-Glu [11]. The alignment results show this Proline is aligned with a Glutamic acid (Glu) of LgtA and T - Threonine (Thr) of BoGT56a. Interestingly, all three sequences have an identical Tyrosine adjacent to this position (Tyr 311 in NmW GT NmW GT). In contrast, the BoGT56a enzyme has a highly conserved region Trp189 and Glu192 which is conserved in all glycosyltransferases of the GT6 family and is known to interact with acceptor substrate in the protein family members for which the three-dimensional structure is known [49]. The hydrophobic property of Trp 189 is conserved in the other two proteins as this position aligns with NmW Tyr 382 and LgtA Phe195. To better visualize areas of conserved sequence, among the NmW GT domain, LgtA and BoGT56a, the sequences were input into BLAST-Cobalt. This analysis revealed a conserved region among all three proteins that corresponds to NmW GT amino acids 235-285 (not shown).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 September 2021 doi:10.20944/preprints202109.0339.v1

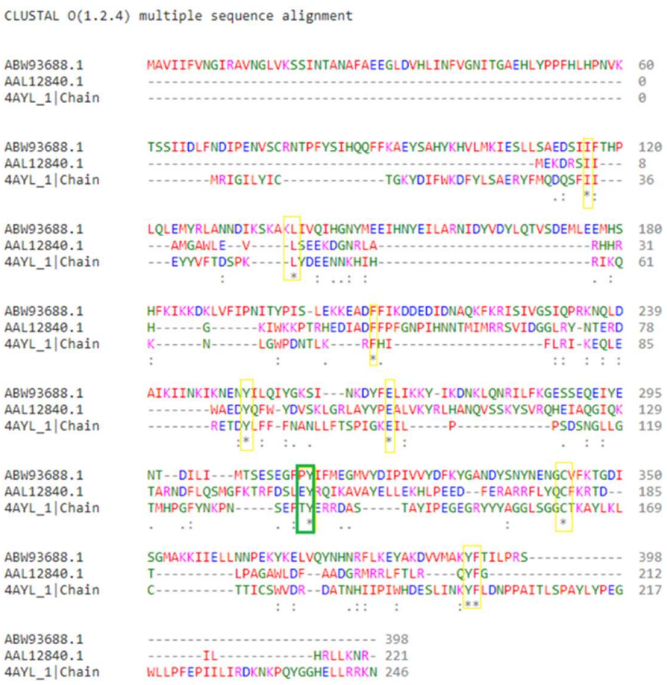

Figure 1. ClustalO multiple sequence alignment of NmW GT domain (ABW93688.1) aligned with LgtA (AAL12840.1) and BoGT56a (4AYL_1). Identical residues among all three are highlighted by a yellow box. The position of NmW Pro310 is highlighted in green and aligned with aligned residues in LgtA and BoGT56a. From Ref. [50].

3.2. Homology Modeling of NmW GT domain and sequence alignment with BshA

In the absence of a crystal structure of the NmW capsule polymerases, homology modeling of the NmW GT domain was performed using I-TASSER. Of the ten protein structures that were used as threading templates for the homology model, a UDP-GlcNActransferase from Staphlyoccocus aureus, BshA (PDB ID: 6d9t) was the top structure with the highest C-score of 0.31 (a measure of confidence in the model). This protein plays a functional role in bacillithiol biosynthesis. Both enzymes are members of the GT4 CAZy family (**Figure 2**).

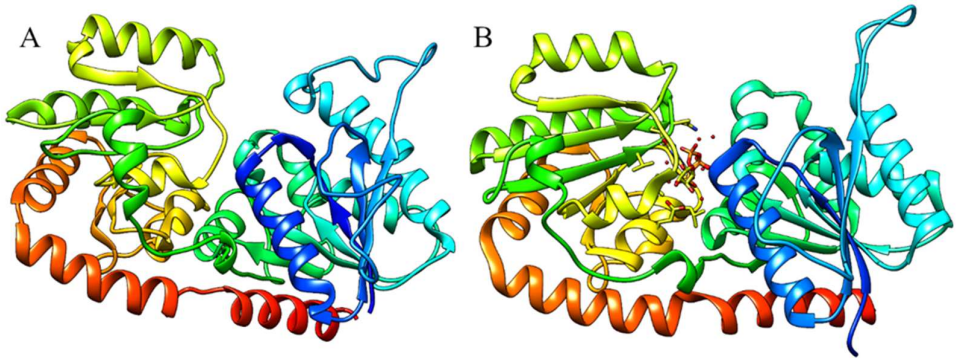

Figure 2. The predicted model of NmW-GT domain. A) I-TASSER generated model; B) BshA template (PDB: 6D9T). The models are colored from the N-terminus (blue) to the C-terminus (red). The overall topology is considered accurate based on the I-TASSER quality scores [30,51,52].

The different poses of UDP-hexoses are shown in Figure 3 based on crystal structures. The rotatable bonds notwithstanding, the differences in the hexose moieties of interest are located at 2 carbon atoms, C2, which carries an -OH or N-acetyl group and C4,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 September 2021 doi:10.20944/preprints202109.0339.v1

where the difference is in the orientation of the -OH group. Analysis of protein-ligand interactions was based on the crystal structures of BshA bound to UDP and with UDP and N-acetylglucosamine (GlcNAc) as reported in the literature [19]. NmW GT differs in substrate specificity in that it utilizes UDP-Gal and these studies seek to change the specificity of this domain to UDP-GalNAc (**Figure 3**). In further analysis, we highlight a series of functionally relevant residues that can inform our studies to alter specificity of the NmW GT domain.

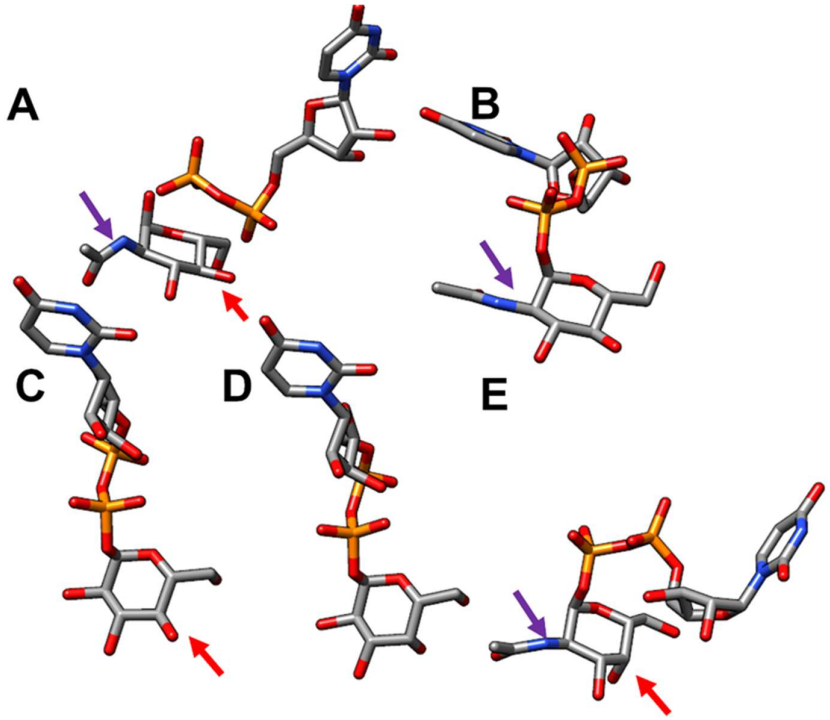

Figure 3. Selected molecular structures of UDP-sugars enzyme substrates. The purple arrows highlight the position of the N-acetyl moiety on C2 and the red arrows highlight the position of the -OH group on C4 whose orientation differs between glucose or galactose. A) UDP and N-acetylglucosamine from a co-crystal structure with *Staphylococcus aureus* BshA [19] B) UDP-N-acetylglucosamine from co-crystal structure with *E. coli* N-acetylglucosamine 1-phosphate uridyltransferase [53]; C) The structure of UDP-glucose structure from a complex with galactose-1-phosphate uridylyltransferase [54]; D) UDP-galactose structure from a complex with galactose-1-phosphate uridylyltransferase [53]; and E) the structure of UDP-N-Acetyl- Galactosamine from a complex with bovine Beta1,4- Galactosyltransferase [55]. The structures were obtained from the PDB and protein components deleted with UCSF-Chimera [38].

Based on the structure of BshA complexed with UDP and N-acetylglucosamine, a model of NmW-GT in complex with the natural ligand UDP-galactose and with UDP-N-acetylgalactosamine was generated within UCSF-Chimera. Through structural alignment in UCSF-Chimera and Matchmaker, N-acetylglucosamine was placed within the putative NmW-GT active site and the residues within 3 angstroms of the ligand identified as L16, H144, V170, R234, and the motif EGFPY 307-311. Among these nine residues, 4 are identical between the two proteins (H144, V170, E307, and F309) (**Figure 4**). It has been shown in other published studies that mutation of NmW GT E307A leads to the complete abolishment of galactosyltransferase activity in that domain [19]. In addition to the likely effects due the stereochemical differences of glucose and galactose moieties at carbon 4, the residues in proximity to the N-acetyl group (R234 and S306) are not conserved between the two proteins and are good targets for modification. The 2-D representations of putative ligand-protein interactions were generated with LigPlot+[56].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 September 2021 doi:10.20944/preprints202109.0339.v1

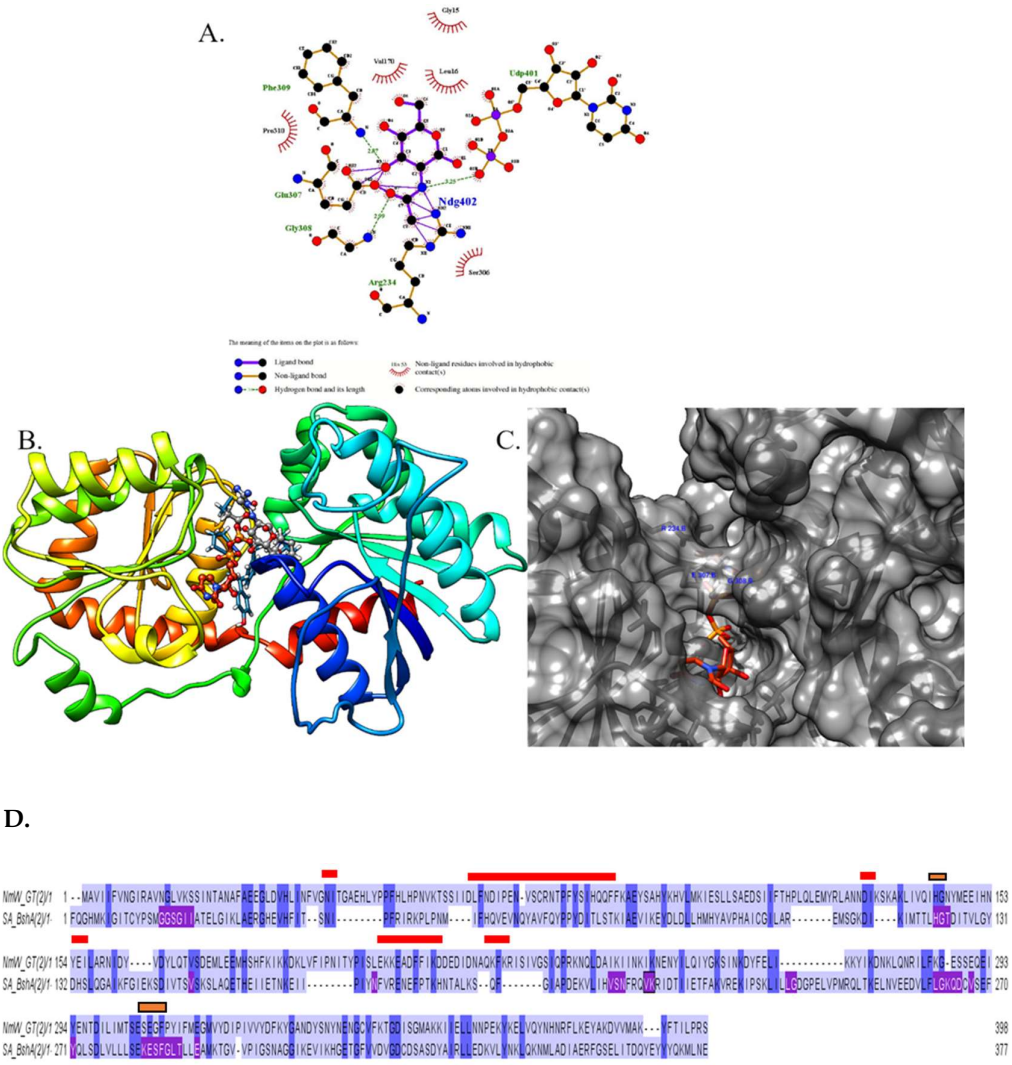

Figure 4. Model of the NmW GT active site residues. A) Proposed interactions of NmW GT with UDP-N-acetylglucosamine. 2-D interaction map was generated with LigPlus B) and C) Location of the active site identified through structural alignment. The NmW GT model and BshA crystal structure were aligned with Matchmaker in UCSF-Chimera, then BshA protein sequence was deleted. Residues potentially interacting with N-acetyl moiety and 4-OH groups are labeled in C. D) Pairwise sequence alignment of NmW-GT and BshA. Red indicates highly fluctuating regions in the molecular dynamics simulations, below, purple indicates the residues in BshA that interact with the ligand and gold represents residues that interact with the GlcNAc moiety in BshA and are conserved in NmW-GT.

The NmW-GT domain was aligned to selected BshA sequences to better map functional motifs as identified through X-ray crystallography of the *S. aureus* enzyme5. In panel C, residues highlighted in purple are determined to be within 5 angstroms of the ligands (BshA sequence) and they map to the same regions identified through experimental data by other authors. The region containing the ESFG motif (280-283; numbering is based on *S. aureus* BshA) that is identified as forming hydrogen bonds with GlcNAc (Ndg402) in *S. aureus* BshA 5 is highly conserved in BshA enzymes. Two residues that interact directly with GlcNAc in this motif (E and F) are also conserved in NmW-GT (see the gold paneling in **Figure 4D, Figure 5**). Of special note is that the G is replaced with P in NmW-GT Mutation of this residue to G in NmW-GT changes the enzyme's specificity for UDP-Gal in favor of UDP-Glu [11]. Other residues that form hydrogen bonds with GlcNAc in BshA are H118 and N171. H118 is conserved in NmW-GT but N171 is not (**Fig**

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 September 2021 doi:10.20944/preprints202109.0339.v1

4c). Whereas in the crystal structure this residue interacts with the C6 O and not the anomeric carbon (see the gold arrow in **Figure 5**), mutational data suggests a critical role in catalysis. Other functionally important residues are T120, which is replaced with N in NmW-GT, K209 and E288, which are conserved in NmW-GT, and E280, which is also conserved in NmW-GT (see the blue arrow in the left panel). E280 and E288 align with E307 and E314 of the NmW GT. These residues are part of a conserved EX_7E motif known to be critical for nucleotide donor sugar recognition in GT4 family members [57].

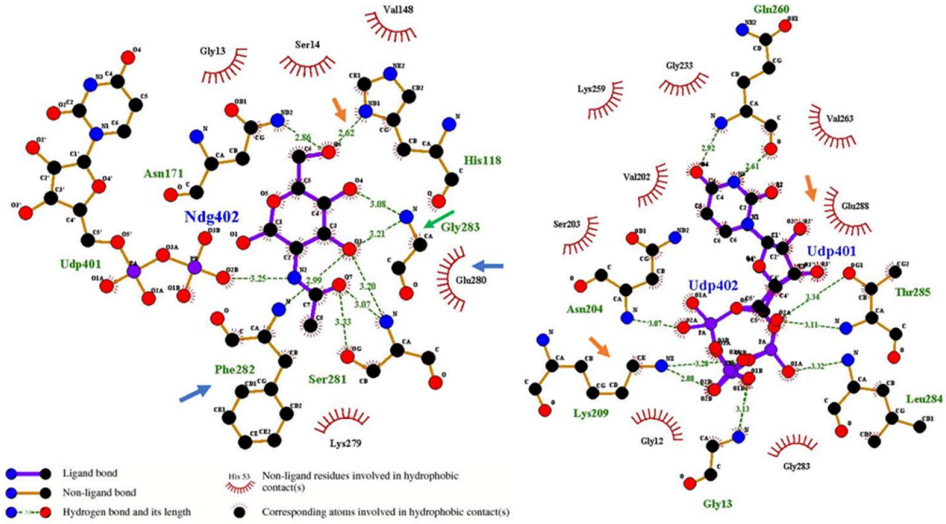

Figure 5. Putative ligand binding residues in NmW-GT based on the modeling template. Ligand-proximal residues in the template structures were identified with Ligplot+. Arrows point to the ligand-proximal residues that are conserved in BshA and NmW-GT.

A multiple sequence alignment was conducted with selected BshA orthologs and with NmW-GT (**Figure 6**) and data show the high conservation of these functionally important residues. Still, several substrate-proximal residues in BshA are not conserved in NmW-GT (see the gold bar in **Figure 6**). Structural alignment with Matchmaker [58] reveals the potential positioning of the substrate binding residues and a good agreement of the model with the template, with an RMSD of 0.540 angstroms between 265 pruned atom pairs or 3.586 across all 363 pairs (**Fig 8A**).

 doi:10.20944/preprints202109.0339.v1

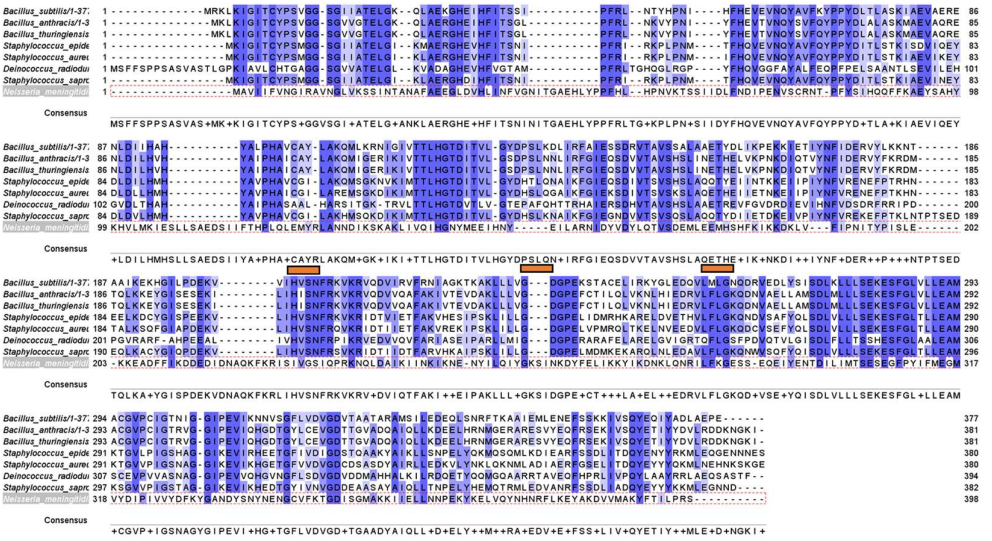

Figure 6. Comparison of NmW-GT with BshA sequences from different species. Blue highlights regions of conservation and the gold bars highlight regions that interact with the substrate in BshA enzymes and that are not conserved in NmW-GT.

3.3. Molecular Dynamics Simulations

The NmW-GT domain model used in these studies was generated with I-TASSER25 using the Staphylococcus aureus bacillithiol biosynthesis glycosyltransferase (BshA) as the template. The structure of *S. aureus* BshA was solved in the presence of UDP and in the presence of UDP and N-acetylglucosamine (GlcNAc) [19]. The structure revealed substrate binding residues and the involvement of the substrate in the catalytic mechanism. Molecular dynamics simulations were conducted to further investigate the conformational spaces that are accessible to the critical residues [19]. Among the highly flexible regjons, the Val-Ser-Asn (202-204) segment is near the substrate (**Figure 8A,** residues rendered in ball-and-stick format and highlighted in pink**).** K209, which is next to this region, is implicated in catalytic activity5 and, whereas the VSN sequence is not conserved in NmW-GT, K is conserved (**Figure 4D**). The flexible regions are highlighted in pink in BshA and in cyan in NmW-GT and partial overlaps are apparent (**Figure 8A**) and are visualized through principal component analysis for NmW-GT (**Figure 8B**). Regions contributing the most to the structural fluctuations are primarily in the terminal regions, as expected, the periphery of the structure, but also at the interface of the two domains, which is also proximal to the active site.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 September 2021 doi:10.20944/preprints202109.0339.v1

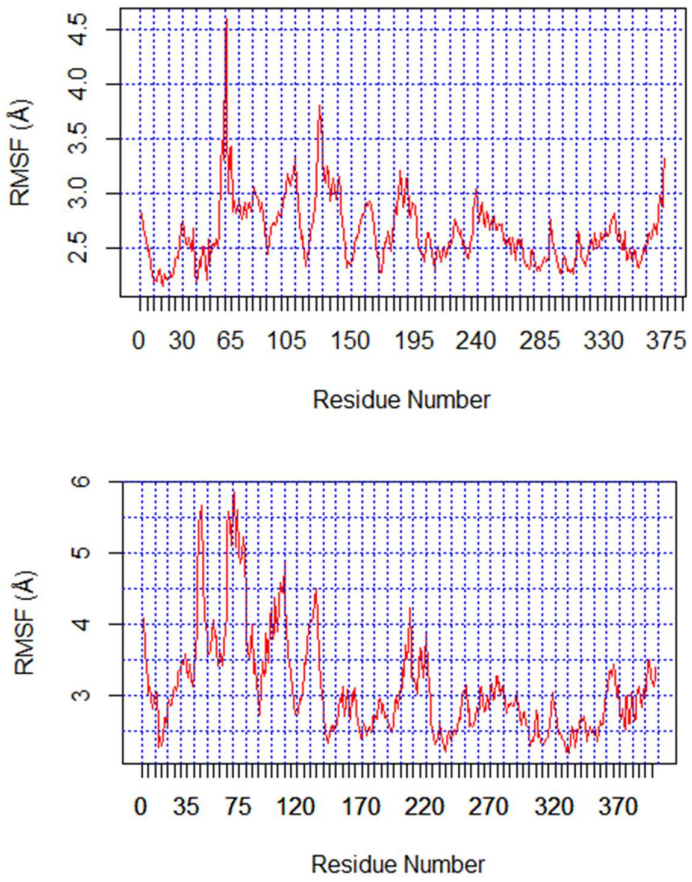

Figure 7. Residue-wise mobility during MD simulation. Top panel: *S. aureus* Bsha. Regions with the highest fluctuations are 55-65 (NQYAVFQYPPYD) and 125-135 (LQGAIKFGDHS). Bottom panel: NmW-GT domain. Highest fluctuations are observed in regions 40-50 (VGNITGAEHLY), 60-80 (KTSSIIDLFNDIPENVSCRNT), 100-110 (HVLMKIEESLLLS), 130-140(NNDIKSKAKLI), and 200-220(SLEKKEADFFIKDDEDIDNAQ). The vertical gridlines are spaced by 10 residues and tick marks by 5 residues. Panel A combined trajectories at the first 10ns, 35 to 45ns and the last 10ns of the total 99ns trajectory. Panel B combines the initial 10 ns and the last 10ns of the trajectory.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 September 2021 doi:10.20944/preprints202109.0339.v1

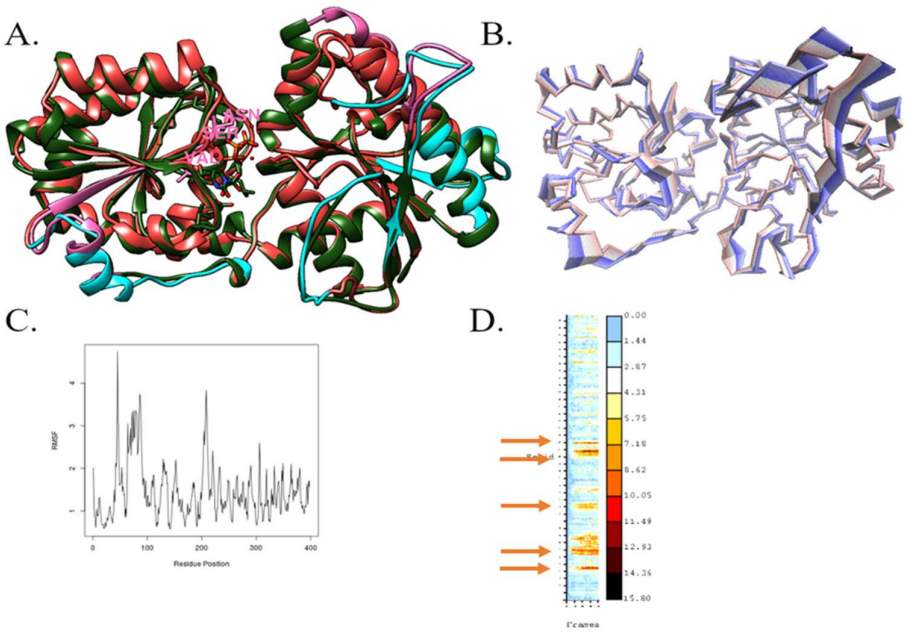

Figure 8. NmW-GT conformation changes. A) RMSF analysis was conducted on MD trajectories and the regions showing highest conformational changes were highlighted with UCSF-Chimera in pink (BshA) and cyan (NmW-GT). The template (PDB ID: 6d9t) and the NmW-GT model were aligned with Matchmaker in UCSF-Chimera. B) Segments with the highest flexibility in NmW-GT as revealed through principal component analysis (PCA). C) Plot of the contributions of each residue to fluctuations of the NmW-GT. D) Heatmap of the RMSD. The regions experiencing the highest mobility are as follows: Residues 42-48; 64-91; 128-136; 154-156; 201-214; and 219-223; gold arrows, also see the boxes in Fig. 9.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 September 2021 doi:10.20944/preprints202109.0339.v1

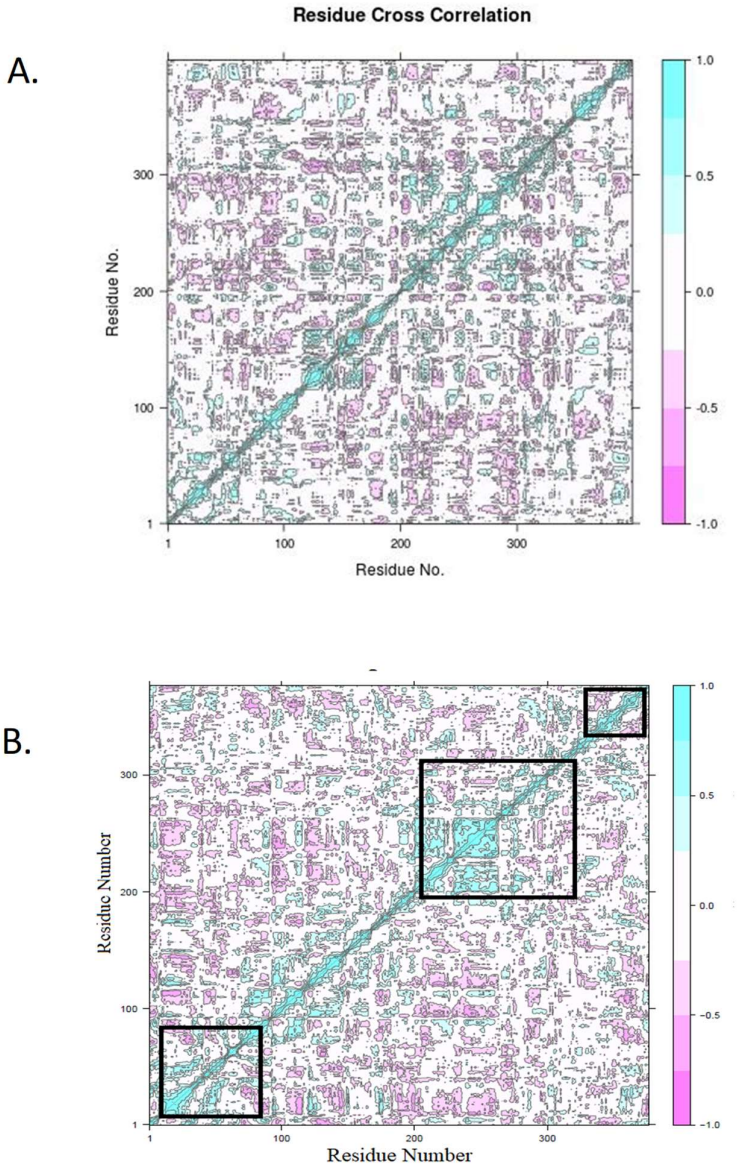

Figure 9. NmW-GT correlative motions. A) NmW-GT model; B) BshA structure. The DCCM matrix with constructed and plotted with Bio3D R library. Regions that move in a correlated manner are shown in green and anticorrelated ones in pink. From Ref. [50].

3.3.1. Cross-correlation of fluctuating residues

Dynamic cross-correlation matrix (DCCM) was constructed to determine the regions that move in concert during the trajectory (Figure 9). Incidentally, the regions with high degrees of fluctuation appear to be correlated in the NmW-GT model (boxes in Fig. 9). The DCCM statistics of the template structure (6d9t) (Figure 9) highlighted the following regions as positively correlated during MD simulation: (aa 21 – 90, starts with α-helix structure and end at a loop region), (aa 207 –328, begins and ends with α-helix structure with a mixture of structure in between), and (aa 356 – 376 falls under α-helix structure).

 doi:10.20944/preprints202109.0339.v1

4. Discussion

Heavy metals environmental pollution remains a challenging problem that could potentially be alleviated through biosorption. Microbial polysaccharides have been shown to effectively bind lead in water. To enhance chelation and hence the capacity for bioremediation, biosynthetic routes for capsular polysaccharides with desired properties are currently under development. The GT domain of the *Neisseria meningiditis* serogroup W capsule polymerase was modeled to inform experimental studies that seek to alter the substrate specificity of this domain. We identified putative active site residues through homology modeling and sequence alignments. Molecular dynamics simulations also revealed regions of fluctuation that surround the active site region and that might influence enzyme conformations and hence activity.

In the context of the full-length NmW capsule polymerase, others have shown that P310G mutation leads to a switch in nucleotide donor specificity from UDP-Gal to UDP-Glu [11]. This can be interpreted as an indication that, like BshA, the Gly then forms hydrogen bonding interactions with the C3 and C4 OH groups of glucose (Figure 5). Given that the main chain NH group that forms the hydrogen bond is lacking in proline, it is tempting to speculate that the proline and the subsequent tyrosine (Y311) enable interactions that are favorable to specificity for galactose. S281 forms H-binding interactions with the acetyl moiety of GlcNAc in BshA (Figure 5) and this residue is replaced with a G in NmW GT. Thus, this region merits further investigation. While our results await experimental validation, others have reported that the alteration of a single amino acid in β1,4-galactosyltransferase biases substrate utilization from UDP-Gal to UDP-GalNAc [55]. Further, conformational changes mediated by protein-protein interactions were important in determining substrate specificity [55].

Mutation of either H118, T120, K209, E280, or E288 led to the complete loss of BshA catalytic activity suggesting a role for these residues in the reaction mechanism [19]. This is consistent with other studies that reported that the highly conserved EX_7E motif of the GT4 family mediates catalysis. Still, while E280 is proximal to the reactive bond, E288 is positioned closer to the uracil group, which is consistent with the observation that the mutation of the initial E in the AceA α-mannosyltransferase from *Acetobacter xylinum*, also a GT4 enzyme, leads to a complete loss of enzymatic activity while mutants of the second E retain some residual activity [59]. This motif is conserved in NmW-GT but several ligand binding residues exhibit lower levels of conservation. K209 was identified of playing the catalytic role of interacting with a β-phosphate in the UDP-sugar moiety [19]. MD simulations determined that the main chain amide group of G13 and the side chain of K209 are proximal to β-phosphoryl group of the UDP moiety and hence could mediate the displacement of the leaving group [19]. In this study, structural alignments show the conservation of these residues in NmW-GT. There are three types of mechanisms proposed for retaining glycosyltransferases, SN_2, double displacement and substrate assisted SN_i. Recent crystallography and enzymatic studies have definitively ruled out two of the three. H118 was shown to play a critical role in stabilizing the oxocarbenium-like intermediate that develops by the SN_i-like mechanism. Although NmW GT domain has been characterized as a retaining glycosyltransferase, the mechanistic details of how this transfer takes place have not been determined due to the lack of a three-dimensional structure. Given that H118 and the aforementioned residues, are conserved in NmW GT except T120, the present work alludes to a similar catalytic mechanism for this domain.

Some BshA mutations have been reported to eliminate enzymatic activity despite not being proximal to the substrate reacting centers [19]. The results of MD simulations identified additional residues in the regions intervening the substrate binding residues that could contribute to conformational changes that mediate activity (**Figures 8 and 9**). In closing, these studies have utilized a combined strategy of sequence alignment, homology

modeling and MD simulations to provide new insight on potential active site residues of the NmW GT that can be modified to engineered new nucleotide-donor substrate specificity. Future work will target these residues for mutation and assess activity of the NmW capsule polymerase to transfer GalNAc to an acceptor.

Author Contributions: Conceptualization, P.C.M. and J.W.; methodology, S.P., J.W. and P.C.M.; software, X.X.; validation, S.P., J.W.; formal analysis, S.P., J.W. and P.C.M.; investigation, S.P., J.W.; resources, J.W. and P.C.M.; data curation, S.P., J.W. and P.C.M.; writing—original draft preparation, S.P., J.W. and P.C.M.; writing—review and editing, J.W. and P.C.M.; writing—final draft preparation, J.W. and P.C.M.; supervision, P.C.M. and J.W.; funding acquisition, P.C.M. and J.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Science Foundation's Division of Materials Science Biomaterials Program (DMR-BMAT) under grant number 2100978. The authors also acknowledge the support of the National Institute on Minority Health and Health Disparities through grant number 5U54MD013376 and National Institute of General Medical Sciences through grant number 5UL1GM118973.

Data Availability Statement: The data from this work is available upon reasonable request from the authors.

Acknowledgments: This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1548562.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Muthu, M.; Wu, H.F.; Gopal, J.; Sivanesan, I.; Chun, S. Exploiting microbial polysaccharides for biosorption of trace elements in aqueous environments-scope for expansion via nanomaterial intervention. *Polymers (Basel)* **2017**, *9*.
2. Heavy metals : Sources, toxicity and remediation techniques.
3. Mwandira, W.; Nakashima, K.; Kawasaki, S.; Arabelo, A.; Banda, K.; Nyambe, I.; Chirwa, M.; Ito, M.; Sato, T.; Igarashi, T., *et al*. Biosorption of pb (ii) and zn (ii) from aqueous solution by oceanobacillus profundus isolated from an abandoned mine. *Scientific Reports* **2020**, *10*, 21189.
4. Roberts, I.S. The biochemistry and genetics of capsular polysaccharide production in bacteria. *Annu Rev Microbiol* **1996**, *50*, 285-315.
5. Willis, L.M.; Whitfield, C. Structure, biosynthesis, and function of bacterial capsular polysaccharides synthesized by abc transporter-dependent pathways. *Carbohydrate research* **2013**, *378*, 35-44.
6. Schmid, J.; Sieber, V.; Rehm, B. Bacterial exopolysaccharides: Biosynthesis pathways and engineering strategies. *Frontiers in Microbiology* **2015**, *6*.
7. Tzeng, Y.L.; Stephens, D.S. Epidemiology and pathogenesis of neisseria meningitidis. *Microbes Infect* **2000**, *2*, 687-700.
8. Bhattacharjee, A.K.; Jennings, H.J.; Kenny, C.P.; Martin, A.; Smith, I.C. Structural determination of the polysaccharide antigens of neisseria meningitidis serogroups y, w-135, and bo1. *Can J Biochem* **1976**, *54*, 1-8.
9. Jennings, H.J.; Bhattacharjee, A.K.; Bundle, D.R.; Kenny, C.P.; Martin, A.; Smith, I.C. Strucutres of the capsular polysaccharides of neisseria meningitidis as determined by 13c-nuclear magnetic resonance spectroscopy. *J Infect Dis* **1977**, *136 Suppl*, S78-83.
10. Ghimire, S.; McCarthy, P.C. Capture of pb(2+) and cu(2+) metal cations by neisseria meningitidis-type capsular polysaccharides. *Biomolecules* **2018**, *8*, 23.
11. Claus, H.; Stummeyer, K.; Batzilla, J.; Mühlenhoff, M.; Vogel, U. Amino acid 310 determines the donor substrate specificity of serogroup w-135 and y capsule polymerases of neisseria meningitidis. *Mol Microbiol* **2009**, *71*, 960-971.

12. Romanow, A.; Haselhorst, T.; Stummeyer, K.; Claus, H.; Bethe, A.; Mühlenhoff, M.; Vogel, U.; von Itzstein, M.; Gerardy-Schahn, R. Biochemical and biophysical characterization of the sialyl-/hexosyltransferase synthesizing the meningococcal serogroup w135 heteropolysaccharide capsule. *The Journal of Biological Chemistry* **2013**, *288*, 11718-11730.
13. Romanow, A.; Keys, T.G.; Stummeyer, K.; Freiberger, F.; Henrissat, B.; Gerardy-Schahn, R. Dissection of hexosyl- and sialyltransferase domains in the bifunctional capsule polymerases from neisseria meningitidis w and y defines a new sialyltransferase family. *J Biol Chem* **2014**, *289*, 33945-33957.
14. Sharyan, A.; Gonzalez, C.; Ukaegbu, O.; Powell, K.; McCarthy, P.C. Determination of the binding affinities of neisseria meningitidis serogroup w capsule polymerase with two nucleotide sugar substrates. *BMC Res Notes* **2018**, *11*, 482.
15. Li, R.; Yu, H.; Muthana, S.M.; Freedberg, D.I.; Chen, X. Size-controlled chemoenzymatic synthesis of homogeneous oligosaccharides of neisseria meningitidis w capsular polysaccharide. *ACS Catalysis* **2020**, *10*, 2791-2798.
16. Cantarel, B.L.; Coutinho, P.M.; Rancurel, C.; Bernard, T.; Lombard, V.; Henrissat, B. The carbohydrate-active enzymes database (cazy): An expert resource for glycogenomics. *Nucleic Acids Res* **2009**, *37*, D233-238.
17. Ardèvol, A.; Rovira, C. Reaction mechanisms in carbohydrate-active enzymes: Glycoside hydrolases and glycosyltransferases. Insights from ab initio quantum mechanics/molecular mechanics dynamic simulations. *J Am Chem Soc* **2015**, *137*, 7528-7547.
18. Winchell, K.R.; Egeler, P.W.; VanDuinen, A.J.; Jackson, L.B.; Karpen, M.E.; Cook, P.D. A structural, functional, and computational analysis of bsha, the first enzyme in the bacillithiol biosynthesis pathway,. *Biochemistry* **2016**, *55*, 4654-4665.
19. Royer, C.J.; Cook, P.D. A structural and functional analysis of the glycosyltransferase bsha from staphylococcus aureus: Insights into the reaction mechanism and regulation of bacillithiol production. *Protein Science: A Publication of the Protein Society* **2019**, *28*, 1083-1094.
20. Parsonage, D.; Newton, G.L.; Holder, R.C.; Wallace, B.D.; Paige, C.; Hamilton, C.J.; Dos Santos, P.C.; Redinbo, M.R.; Reid, S.D.; Claiborne, A. Characterization of the n-acetyl-α-d-glucosaminyl l-malate synthase and deacetylase functions for bacillithiol biosynthesis in bacillus anthracis. *Biochemistry* **2010**, *49*, 8398-8414.
21. Database resources of the national center for biotechnology information. *Nucleic Acids Res* **2018**, *46*, D8-d13.
22. Blixt, O.; van Die, I.; Norberg, T.; van den Eijnden, D.H. High-level expression of the neisseria meningitidis lgta gene in escherichia coli and characterization of the encoded n-acetylglucosaminyltransferase as a useful catalyst in the synthesis of glcnac beta 1-->3gal and galnac beta 1-->3gal linkages. *Glycobiology* **1999**, *9*, 1061-1071.
23. Erwin, A.L.; Haynes, P.A.; Rice, P.A.; Gotschlich, E.C. Conservation of the lipooligosaccharide synthesis locus lgt among strains of neisseria gonorrhoeae: Requirement for lgte in synthesis of the 2c7 epitope and of the beta chain of strain 15253. *J Exp Med* **1996**, *184*, 1233-1241.
24. Gotschlich, E.C. Genetic locus for the biosynthesis of the variable portion of neisseria gonorrhoeae lipooligosaccharide. *J Exp Med* **1994**, *180*, 2181-2190.
25. Wakarchuk, W.; Martin, A.; Jennings, M.P.; Moxon, E.R.; Richards, J.C. Functional relationships of the genetic locus encoding the glycosyltransferase enzymes involved in expression of the lacto-n-neotetraose terminal lipopolysaccharide structure in neisseria meningitidis. *J Biol Chem* **1996**, *271*, 19166-19173.
26. Yang, Q.L.; Gotschlich, E.C. Variation of gonococcal lipooligosaccharide structure is due to alterations in poly-g tracts in lgt genes encoding glycosyl transferases. *J Exp Med* **1996**, *183*, 323-327.
27. Thiyagarajan, N.; Pham, T.T.K.; Stinson, B.; Sundriyal, A.; Tumbale, P.; Lizotte-Waniewski, M.; Brew, K.; Acharya, K.R. Structure of a metal-independent bacterial glycosyltransferase that catalyzes the synthesis of histo-blood group a antigen. *Scientific Reports* **2012**, *2*, 940.
28. Sievers, F.; Higgins, D.G. The clustal omega multiple alignment package. *Methods in Molecular Biology (Clifton, N.J.)* **2021**, *2231*, 3-16.

29. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T.J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J., *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using clustal omega. *Mol Syst Biol* **2011**, *7*, 539.
30. Zheng, W.; Zhang, C.; Bell, E.W.; Zhang, Y. I-tasser gateway: A protein structure and function prediction server powered by xsede. *Future generations computer systems : FGCS* **2019**, *99*, 73-85.
31. 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.; Richardson, D.C. Molprobity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **2010**, *66*, 12-21.
32. Davis, I.W.; Leaver-Fay, A.; Chen, V.B.; Block, J.N.; Kapral, G.J.; Wang, X.; Murray, L.W.; Arendall, W.B., 3rd; Snoeyink, J.; Richardson, J.S., *et al.* Molprobity: All-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* **2007**, *35*, W375-383.
33. Davis, I.W.; Murray, L.W.; Richardson, J.S.; Richardson, D.C. Molprobity: Structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res* **2004**, *32*, W615-619.
34. Brooks, B.R.; Brooks, C.L.; MacKerell, A.D.; Nilsson, L.; Petrella, R.J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S., *et al.* Charmm: The biomolecular simulation program. *Journal of computational chemistry* **2009**, *30*, 1545-1614.
35. Jo, S.; Cheng, X.; Lee, J.; Kim, S.; Park, S.-J.; Patel, D.S.; Beaven, A.H.; Lee, K.I.; Rui, H.; Roux, B., *et al.* Charmm-gui 10 years for biomolecular modeling and simulation. *Journal of computational chemistry* **2017**, *38*, 1114-1124.
36. Jo, S.; Kim, T.; Iyer, V.G.; Im, W. Charmm-gui: A web-based graphical user interface for charmm. *Journal of Computational Chemistry* **2008**, *29*, 1859-1865.
37. Lee, J.; Cheng, X.; Swails, J.M.; Yeom, M.S.; Eastman, P.K.; Lemkul, J.A.; Wei, S.; Buckner, J.; Jeong, J.C.; Qi, Y., *et al.* Charmm-gui input generator for namd, gromacs, amber, openmm, and charmm/openmm simulations using the charmm36 additive force field. *Journal of Chemical Theory and Computation* **2016**, *12*, 405-413.
38. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. Ucsf chimera--a visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **2004**, *25*, 1605-1612.
39. Beglov, D.; Roux, B. Finite representation of an infinite bulk system: Solvent boundary potential for computer simulations. *The Journal of Chemical Physics* **1994**, *100*, 9050-9063.
40. Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **1983**, *79*, 926-935.
41. Phillips, J.C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R.D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with namd. *Journal of Computational Chemistry* **2005**, *26*, 1781-1802.
42. Phillips, J.C.; Hardy, D.J.; Maia, J.D.; Stone, J.E.; Ribeiro, J.V.; Bernardi, R.C.; Buch, R.; Fiorin, G.; Hénin, J.; Jiang, W. Scalable molecular dynamics on cpu and gpu architectures with namd. *The Journal of Chemical Physics* **2020**, *153*, 044130.
43. Sagui, C.; Darden, T.A. Molecular dynamics simulations of biomolecules: Long-range electrostatic effects. *Annual Review of Biophysics and Biomolecular Structure* **1999**, *28*, 155-179.
44. Best, R.B.; Zhu, X.; Shim, J.; Lopes, P.E.M.; Mittal, J.; Feig, M.; Mackerell, A.D. Optimization of the additive charmm all-atom protein force field targeting improved sampling of the backbone φ, ψ and side-chain χ(1) and χ(2) dihedral angles. *Journal of Chemical Theory and Computation* **2012**, *8*, 3257-3273.
45. Towns, J.; Cockerill, T.; Dahan, M.; Foster, I.; Gaither, K.; Grimshaw, A.; Hazlewood, V.; Lathrop, S.; Lifka, D.; Peterson, G.D., *et al.* Xsede: Accelerating scientific discovery. *Computing in Science & Engineering* **2014**, *16*, 62-74.
46. Humphrey, W.; Dalke, A.; Schulten, K. Vmd: Visual molecular dynamics. *Journal of Molecular Graphics* **1996**, *14*, 33-38.
47. Grant, B.J.; Rodrigues, A.P.C.; ElSawy, K.M.; McCammon, J.A.; Caves, L.S.D. Bio3d: An r package for the comparative analysis of protein structures. *Bioinformatics* **2006**, *22*, 2695-2696.

48. Albesa-Jové, D.; Giganti, D.; Jackson, M.; Alzari, P.M.; Guerin, M.E. Structure-function relationships of membrane-associated gt-b glycosyltransferases. *Glycobiology* **2014**, *24*, 108-124.
49. Brockhausen, I. Crossroads between bacterial and mammalian glycosyltransferases. *Front Immunol* **2014**, *5*.
50. Paudel, S. Computationally- guided design and engineering of a neisseria meningitidis serogroup w capsule polymerase. M.S., Morgan State University, Ann Arbor, 2021.
51. Zhang, Y.; Skolnick, J. Scoring function for automated assessment of protein structure template quality. *Proteins* **2004**, *57*, 702-710.
52. Yang, J.; Yan, R.; Roy, A.; Xu, D.; Poisson, J.; Zhang, Y. The i-tasser suite: Protein structure and function prediction. *Nat Methods* **2015**, *12*, 7-8.
53. Brown, K.; Pompeo, F.; Dixon, S.; Mengin-Lecreulx, D.; Cambillau, C.; Bourne, Y. Crystal structure of the bifunctional n-acetylglucosamine 1-phosphate uridyltransferase from escherichia coli: A paradigm for the related pyrophosphorylase superfamily. *The EMBO journal* **1999**, *18*, 4096-4107.
54. Thoden, J.B.; Ruzicka, F.J.; Frey, P.A.; Rayment, I.; Holden, H.M. Structural analysis of the h166g site-directed mutant of galactose-1-phosphate uridylyltransferase complexed with either udp-glucose or udp-galactose: Detailed description of the nucleotide sugar binding site. *Biochemistry* **1997**, *36*, 1212-1222.
55. Ramakrishnan, B.; Qasba, P.K. Structure-based design of beta 1,4-galactosyltransferase i (beta 4gal-t1) with equally efficient n-acetylgalactosaminyltransferase activity: Point mutation broadens beta 4gal-t1 donor specificity. *The Journal of Biological Chemistry* **2002**, *277*, 20833-20839.
56. Laskowski, R.A.; Swindells, M.B. Ligplot+: Multiple ligand-protein interaction diagrams for drug discovery. *Journal of Chemical Information and Modeling* **2011**, *51*, 2778-2786.
57. Kapitonov, D.; Yu, R.K. Conserved domains of glycosyltransferases. *Glycobiology* **1999**, *9*, 961-978.
58. Meng, E.C.; Pettersen, E.F.; Couch, G.S.; Huang, C.C.; Ferrin, T.E. Tools for integrated sequence-structure analysis with ucsf chimera. *BMC Bioinformatics* **2006**, *7*, 339.
59. Abdian, P.L.; Lellouch, A.C.; Gautier, C.; Ielpi, L.; Geremia, R.A. Identification of essential amino acids in the bacterial alpha -mannosyltransferase acea. *J Biol Chem* **2000**, *275*, 40568-40575.