Organic & Biomolecular Chemistry

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: Y. Guo, L. Wang, J. Guo, G. Gu and Z. Guo, *Org. Biomol. Chem.*, 2017, DOI: 10.1039/C7OB00555E.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/obc

View Article Online
DOI: 10.1039/C7OB00555E

Biochemical Studies of Inositol *N*-Acetylglucosaminyltransferase Involved in Mycothiol Biosynthesis in *Corynebacterium diphtheria*

Yuchuan Guo,[a,b] Lizhen Wang,[a] Jiatong Guo,[b] Guofeng Gu,[a,*] and Zhongwu Guo[b,*]

[a] National Glycoengineering Research Center and Shandong Provincial Key Laboratory of Carbohydrate Chemistry and Glycobiology, Shandong University, 27 Shanda Nan Lu, Jinan 250100, China

[b] Department of Chemistry, University of Florida, 214 Leigh Hall, Gainesville, Florida 32611, United States

* Corresponding author E-mails: guofenggu@sdu.edu.cn; zguo@chem.ufl.edu

Abstract

Mycothiol (MSH) is the predominant low molecular weight thiol produced by actinomycetes, and it plays a pivotal role in the bacterial detoxication process. 1L-*myo*-Inositol-1-phosphate (1L-Ins-1-P) α-*N*-acetylglucosaminyltransferase (GlcNAc-T), known as MshA, is the only glycosyltransferase involved in MSH biosynthesis. In this work, the MshA from *Corynebacterium diphtheria*, named as CdMshA, was expressed, purified, and studied in detail. Its enzymatic activity to transfer GlcNAc to 1L-Ins-1-P was confirmed by the isolation and rigorous characterization of its reaction product 3-phospho-1-D-*myo*-inositol-2-acetamido-2-deoxy-α-D-glucopyranoside. CdMshA was shown to accept only UDP-GlcNAc and 1L-Ins-1-P as its substrates among various tested glycosyl donors, such as UDP-GlcNAc, UDP-Gal, UDP-Glc, UDP-GalNAc and UDP-GlcA, and glycosyl acceptors, such as *myo*-inositol, 1L-Ins-1-P and 1D-Ins-1-P. The results have demonstrated the strict substrate selectivity of CdMshA. Furthermore, its reaction kinetics with UDP-GlcNAc and 1L-Ins-1-P as substrates were characterized, while site-directed mutagenesis of CdMshA disclosed that its amino acid residues N28, K81 and R157 were essential for its enzymatic activity.

Key words: Mycothiol, Inositol *N*-acetylglucosaminyltransferase, *Corynebacterium diphtheria*, substrate selectivity, site-directed mutagenesis

View Article Online
DOI: 10.1039/C7OB00555E

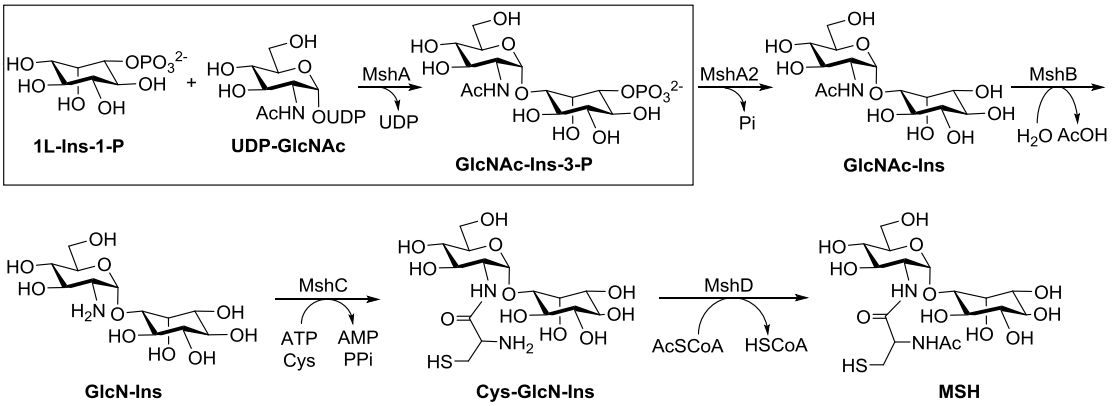
Introduction

Mycothiol (MSH), 1-D-*myo*-inositol 2-(*N*-acetyl-L-cysteinyl)amino-2-deoxy-α-D-glucopyranoside, is the principal low molecular thiol in actinobacteria and is the functional equivalent of glutathione found in Gram-negative bacteria and most eukaryotes.[1, 2] Similar to glutathione, mycothiol plays a pivotal role in protecting bacteria against oxidative stresses and in the detoxication of xenobiotics.[3] In mycobacteria where MSH has been extensively investigated, MSH-deficient mutants exhibited increased sensitivities to oxidative stresses, alkylating agents, and a broad range of antibiotics,[4, 5] suggesting the significance of MSH in mycobacterial survival and pathogenicity.[6] Consequently, targeted disruption of MSH biosynthesis in *Mycobacterium tuberculosis* Erdman produced no viable bacterial clones, indicating that MSH is essential for *M. tuberculosis* Erdman.[6] In addition, it has been demonstrated that MSH was important for *Corynebacterium glutamicum* resistance to alkylating agents, glyphosate, ethanol, antibiotics, heavy metals, and aromatic compounds[7] and that increasing the intracellular MSH content in *C. glutamicum* could improve its robustness to various stresses.[8] Therefore, the MSH biosynthetic pathway represents a promising target for the development of new antibacterial therapeutics or novel agents that can improve the efficacies of current antibiotics.[9, 10]

As shown in Scheme 1, MSH biosynthesis consists of five enzymatic steps[3, 11] mediated respectively by an inositol *N*-acetylglucosaminyltransferase (GlcNAc-T) (MshA),[12-14] a phosphatase (MshA2), a GlcNAc deacetylase (MshB),[15-20] an ATP-dependent ligase (MshC),[21-24] and an acetyltransferase (MshD).[25, 26] Except for MshA2, the other four enzymes from different bacterial species have been identified. As outlined, the first dedicated step for MSH biosynthesis is MshA-mediated transfer of a GlcNAc unit from α-GlcNAc uridine diphosphate (UDP-GlcNAc) to 1L-*myo*-insitol-1-phosphate (1L-Ins-1-P), which affords 3-phospho-1-D-*myo*-inositol-2-acetamido-2-deoxy-α-D-glucopyranoside (GlcNAc-Ins-3-P). An MshA from *C. glutamicum* (CgMshA) was studied by X-ray crystallography which helped the understanding of this enzyme.[12] However, in general, information about MshA, and other enzymes involved in MSH biosynthesis as well, is limited and more detailed studies on these enzymes are necessary to gain insights into their functions and application.

View Article Online
DOI: 10.1039/C7OB00555E

Scheme 1. The MSH biosynthetic pathway



We described herewith the expression, purification, and characterization of a novel MshA from *C. diphtheria*, named as CdMshA. The biological properties of this enzyme, including its reaction kinetics and substrate specificity, were studied in detail employing commercial sugar donors and chemically synthesized inositol and sugar derivatives as glycosyl acceptors. Moreover, site-directed mutagenesis of CdMshA was performed to investigate its structure-enzymatic activity relationship.

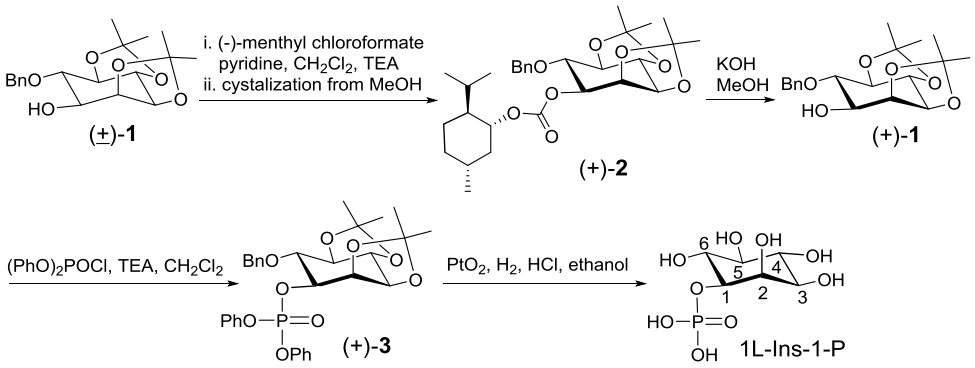
Results and Discussion

Expression and purification of CdMshA. The MshA-encoding gene *mshA* was derived from *C. diphtheria*, synthesized by Sangon biotech, and ligated into the *Nde*I/*Xho*I sites of the pET-22b expression vector. Thereafter, the recombinant plasmid carrying *mshA* was transferred into *E. coli* BL21 (DE3), and the transformant was cultured in LB medium containing ampicillin at 37 ^{o}C overnight and then transferred into a fresh medium in a 1:100 (v/v) ratio. Enzyme expression was induced upon the addition of IPTG after the cell culture reached an optical density (OD) value of 0.60~0.80 at 600 nm. Following continuous cultivation at 16 ^{o}C for 12 h, cells were harvested and disrupted by ultrasonication. The cell lysate was centrifuged at 4 ^{o}C and 12000 rpm for 30 min to obtain the crude enzyme in the supernatant, which was further purified to homogeneity by Nickel-chelation chromatography. The purified enzyme showed a single band corresponding to ~50 kDa in SDS-PAGE (Figure S1), which was in accordance with the theoretical mass prediction (47,416 Da) of CdMshA.

View Article Online
DOI: 10.1039/C7OB00555E

Enzymatic activity, substrate specificity, and reaction kinetics of CdMshA. The natural substrate of MshA, 1L-Ins-1-P, was synthesized by a reported procedure (Scheme 2)[27, 28] and used to assess the enzymatic activity of purified CdMshA to catalyze the transfer of GlcNAc to inositol. First, *myo*-inositol was converted into racemic (±)-**1**,[28-30] which was then resolved in two steps, including acylation using (-)-menthyl chloroformate and separation of the resultant diastereomers through recrystallization from methanol to obtain isomer (+)-**2**, followed by its saponification to produce enantiomerically pure (+)-**1**. Thereafter, (+)-**1** was phosphorylated smoothly to give (+)-**3** that was finally deprotected to provide 1L-Ins-1-P. The analytical data of all synthetic intermediates and the final product agreed well with their structures and the reported data.[28-31]

Scheme 2. Synthesis of 1L-Ins-1-P



Assays of the enzymatic activity of CdMshA were performed in Tris-HCl buffer (50 mM, pH 8.0) containing 1L-Ins-1-P (0.5 mM) and UDP-GlcNAc (1 mM) in a 1:2 ratio, and the concentration of CdMshA was 60 μg/mL. The reaction mixture was kept at 37 ^{o}C, and its aliquots were taken at 1, 2 and 4 h, respectively, and quenched with preheated (60 ^{o}C) acetonitrile and incubation at 60 ^{o}C for 10 min. After the reaction mixtures were cooled with ice and centrifuged (13,000 × g, 10 min) to remove the protein, they were subjected to reversed-phase HPLC analysis. As the reaction product, GlcNAc-Ins-3-P, did not have UV absorptions, we analyzed the reaction mixtures by observing the consumption of UDP-GlcNAc and the generation of UDP. As depicted in Figure 1, after 1, 2 and 4 h of incubation, 15%, 32% and 48% of UDP-GlcNAc was transformed into UDP in the presence of CdMshA, while no conversion or decomposition of UDP-GlcNAc was observed under the same

View Article Online
DOI: 10.1039/C7OB00555E

conditions without CdMshA. These data suggested that the conversion rates of 1L-Ins-1-P under the specified conditions were 30%, 64%, and 96%, respectively. Finally, the product was isolated from the reaction mixture by ion-exchange HPLC to verify the formation and yields of GlcNAc-Ins-3-P, and its structure was characterized with MS and NMR spectroscopy.[12] For example, the coupling constant of the anomeric proton (J_{H1-2} = 3.6 Hz) of its GlcNAc confirmed the α-configuration. These results proved unambiguously the enzymatic activity of CdMshA to catalyze GlcNAc transfer to 1L-Ins-1-P. Eventually, CdMshA was used to realize milligram-scale synthesis of GlcNAc-Ins-3-P, which offered a 95% isolated yield.

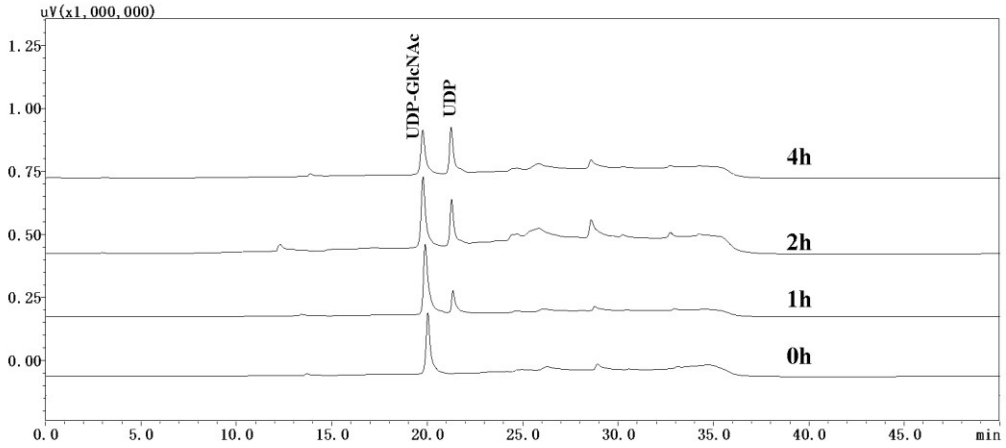


Figure 1. HPLC results of CdMshA-catalyzed reaction between UDP-GlcNAc and 1L-Ins-1-P for 1, 2, and 4 h, respectively. The reaction was carried out in Tris-HCl buffer (50 mM, pH 8.0) containing UDP-GlcNAc (1 mM) and 1L-Ins-1-P (0.5 mM), and the concentration of CdMshA was 60 μg/mL. HPLC conditions: C18 column (4.6 x 150 mm), 1 mL/min flow rate, gradient eluents using buffer A (2 mM tetrabutylammonium hydroxide in water at pH 6.5) and buffer B (10 mM tetrabutylammonium hydroxide and 77 mM KH_2PO_4 in 50% aqueous methanol at pH 5.5) and UV detection at 260 nm.

After the enzymatic activity of CdMshA was verified, we continued to investigate its substrate specificity. In this case, all reactions were carried out and analyzed under the same conditions described above. First, we used the proved glycosyl acceptor 1L-Ins-1-P to probe various glycosyl donors, including UDP-Gal, UDP-Glc, UDP-GalNAc and UDP-GlcA, as enzymatic substrates and found that they were not accepted by CdMshA. Then, using UDP-GlcNAc as the glycosyl donor, we examined the reactivity of *myo*-inositol, 1D-Ins-1-P derived from (±)-**1** by the same method outlined in Scheme 2 with (+)-menthyl carbonate as the resolution reagent, and a series of sugar

View Article Online
DOI: 10.1039/C7OB00555E

phosphates including glucose-6-phosphate, galactose-6-phosphate, mannose-6-phosphate, glucose-1-phosphate in the presence of CdMshA. We found that these compounds were not accepted by CdMshA as glycosyl acceptors either. The results demonstrated the rigid substrate selectivity of CdMshA, which was in agreement with previous reports.[12, 13]

The biochemical properties of CdMshA were further characterized through studying its reaction kinetics and the influence of various conditions, such as temperature, pH and metal cations, on its enzymatic activity. First, we examined the reaction of UDP-GlcNAc and 1L-Ins-1-P using CdMshA in Tris-HCl buffer (50 mM, pH 8.0) at various temperatures. As depicted in Figures 2A, CdMshA showed the maximal activity at 40 ^{o}C, but it could tolerate a wide range of temperature, as about 90% of the catalytic activity was retained from 20 to 50 ^{o}C. Next, we studied the influence of pH on this enzymatic reaction. As shown in Figures 2B, CdMshA exhibited good activity (>80%) at the pH ranging from pH 7 to 10, and the maximal activity was observed at pH 8.0.

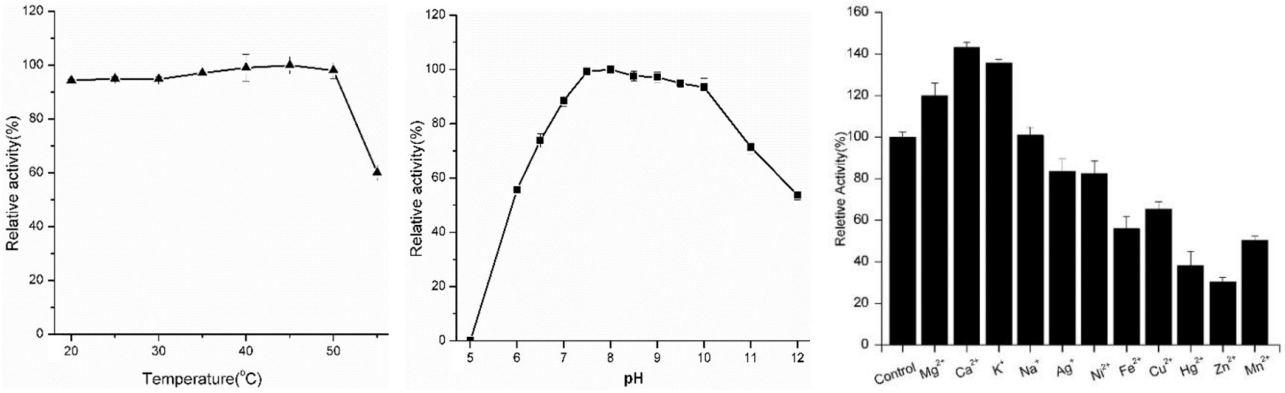


Figure 2. The influences of temperature (A), pH (B), and metal cations (C) on the activity of CdMshA

It was also discovered that divalent metal ions were not strictly necessary for the catalytic activity of CdMshA (Figure 2C), as the enzyme exhibited good activity in the absence of a divalent metal cation. In this respect, CdMshA was similar to other reported GT-4 family glycosyltransferases, such as L-malate α-*N*-acetylglucosaminyltransferase BshA involved in bacillithiol biosynthesis.[32] However, metal cations did exhibit some influences on the enzymatic activity of CdMshA. For example, Mg^{2+}, Ca^{2+}, and K^{+} could slightly enhance the activity of CdMshA, whereas Ag^{+}, Ni^{2+}, Fe^{2+}, Cu^{2+}, Hg^{2+}, Zn^{2+}, and Mn^{2+} demonstrated minor to significant inhibitions on CdMshA. These

View Article Online
DOI: 10.1039/C7OB00555E

findings were consistent with a previous report[32] of slight activity enhancement of BshA by Mg^{2+} and an observation made with the crystal structure of *C. glutamicum* MshA, namely that MshA contains Mg^{2+} but the metal cation is not located at the active site.[12]

After the optimal reaction conditions for CdMshA (Tris-HCl buffer, pH 8.0, 40 ^{o}C) were established, they were employed to examine the reaction kinetics of this enzyme and determine the K_m and V_{max} values for both substrates. Initial reaction velocities for 1L-Ins-1-P and UDP-GlcNAc were obtained with varied 1L-Ins-1-P concentrations/saturated UDP-GlcNAc and varied UDP-GlcNAc concentrations/saturated 1L-Ins-1-P, respectively. The data were utilized to calculate the Michaelis constants (K_m) for 1L-Ins-1-P and UDP-GlcNAc, which were 0.485 ± 0.049 mM and 0.185 ± 0.047 mM (Figure S2), respectively. The V_{max} values for 1L-Ins-1-P and UDP-GlcNAc were 0.437 ± 0.022 and 0.328 ± 0.028 nmol Min^{-1}, and the K_{cat} values were 6.907 ± 0.347 Min^{-1} and 5.184 ± 0.442 Min^{-1}, respectively. The K_m value of CdMshA for UDP-GlcNAc was comparable to those of the crude MshAs derived from *M. smegmatis* (0.17 ± 0.02 mM) and *C. glutamicum* (0.208 ± 0.017 mM).[12, 13] However, the K_m value of CdMshA for 1L-Ins-1-P was about two times higher than those of the crude MshAs from *M. smegmatis* (0.15 ± 0.01 mM) and *C. glutamicum* (0.236 ± 0.013 mM). These results indicated that CdMshA had a slightly lower affinity towards 1L-Ins-1-P than other similar enzymes.

Mutagenic studies of CdMshA. To identify the key amino acids responsible for the enzymatic activity of CdMshA, we probed site-directed mutagenesis of this enzyme. The target amino acids were selected according to a reported crystal structure of CgMshA,[12] whose amino acid sequence was *ca.* 72% identical to that of CdMshA. It was reported that MshA follows an ordered sequential kinetic mechanism with UDP-GlcNAc binding first. After binding with UDP-GlcNAc, MshA has a large conformational change to bring UDP-GlcNAc and 1L-Ins-1-P into close proximity.[12] The crystal structure of CgMshA also revealed that at its active site (Figure 3A), the side chains of K78, R154, T134, and Y110 could form hydrogen bonds with the phosphate group of 1L-Ins-1-P and that the M24, N25 and D20 residues had polar interaction with the 3-OH, 4-OH, and 5-OH groups of 1L-Ins-1-P, respectively. The 2-OH group in 1L-Ins-1-P pointed into a small pocket generated

View Article Online
DOI: 10.1039/C7OB00555E

by the side chains of M24, Y110, T134, and UDP-GlcNAc.[12] Our analysis of MshAs from various species further suggested that amino acids D20, M24, N25, K78, Y110, T134, and R154 were highly conserved (Figure S3). Based on the crystal structure of MshA and the above analyses, we constructed a homology model for the active site of CdMshA (Figure 3B). This model indicated that its D23, M27, N28, K81, Y113, T137, and R157 residues were correspondent to the conserved amino acids of MshAs identified above. Therefore, in this study, we mutated these amino acids to alanine and examined the activities of the resultant mutant enzymes.

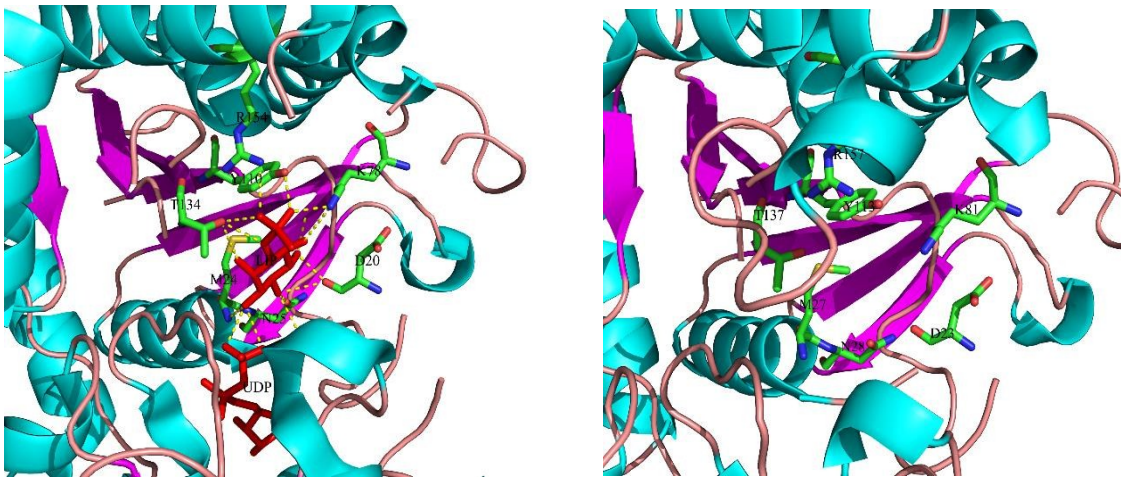


Figure 3. Crystal structure of the active center of CgMshA (A),[12] and the homology model of CdMshA (B). Homology modeling of CdMshA was carried out using the Phyre2 program, and the images were created by the software Pymol-1.4.1.

Site-directed mutation of the CdMshA gene was achieved by standard protocols,[33] and the mutant enzymes were expressed and purified as described above. Assays of the enzymatic activities of the CdMshA mutants were carried out under the above-mentioned optimal reaction conditions, and the results were depicted in Figure 4. Compared to wild type CdMshA, the D23A, M27A, and Y113A mutants exhibited a mild reduction in catalytic activity, while the T137A mutant retained only *ca*. 30% of activity. In contrast, N28A, K81A, and R157A mutations abolished the enzymatic activity almost completely (<5% of activity retained). Clearly, N28, K81, and R157, as well as T137, played a critical role in the enzymatic activity of CdMshA. According to the models depicted in Figure 3, K81/R157 and N28 should form hydrogen bonds with the phosphate and the 4-OH groups of

View Article Online
DOI: 10.1039/C7OB00555E

1L-Ins-1-P, respectively, which would help keep 1L-Ins-1-P in proper orientation at the active site of CdMshA and hold its 3-OH group close to the pyrophosphate group of UDP-GlcNAc. This should be critical for the enzymatic reaction to occur. T137 could also form a hydrogen bond with the phosphate group of 1L-Ins-1-P and, in the meantime, work together with M27 and Y113 to create a small pocket to accommodate the axial 2-OH group of 1L-Ins-1-P and limit its freedom to facilitate the reaction. Comparably, the roles of D23 and M27 were relatively insignificant, as the activity of CdMshA was only mildly affected after mutation at these sites. The requirement of multiple hydrogen bonds formed between CdMshA and the phosphate and hydroxyl groups of 1L-Ins-1-P for the enzyme to function properly and the tight space at the active center of CdMshA, as proposed in the model, may also explain the rigid substrate selectivity of CdMshA.

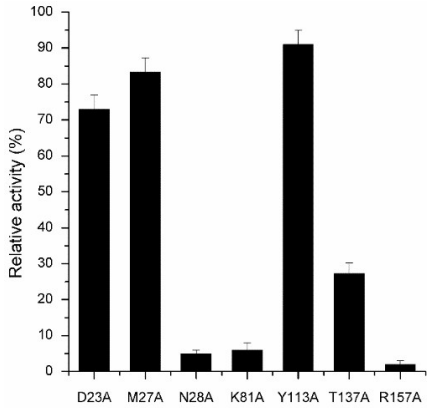


Figure 4. Relative enzymatic activities of various CdMshA mutants

Finally, employing UDP-GlcNAc as the glycosyl donor, we assessed the activities of these mutant enzymes, especially D23A, M27A, Y113A and T137A, to incorporate unnatural glycosyl acceptors *e.g.*, 1D-Ins-1-P, *myo*-inositol, glucose-6-phosphate, galactose-6-phosphate, mannose-6-phosphate and glucose-1-phosphate. HPLC analysis of the reaction mixtures revealed no UDP production, indicating that the mutant enzymes did not use these compounds as substrates. Thus, mutagenesis at the identified sites could not relax the substrate specificity of CdMshA.

Conclusion

In this research, we have accomplished the expression, purification, and characterization of a new MshA from *C. diphtheria* and named it as CdMshA. We have demonstrated that CdMshA could

View Article Online
DOI: 10.1039/C7OB00555E

effectively catalyze the initial step of mycothiol biosynthesis, that is, transferring GlcNAc from UDP-GlcNAc to 1L-Ins-1-P to generate GlcNAc-Ins-3-P, which was isolated and characterized. CdMshA exhibited relatively rigid substrate selectivity, as it did not accept a number of common glycosyl donors and acceptors including 1D-Ins-1-P, a diastereomer of 1L-Ins-1-P. Our studies further revealed that CdMshA tolerated a relatively wide range of pH (7-10) and temperature (20 to 50 ^{o}C). Similar to MshAs from other bacteria, the enzymatic activity of CdMshA did not need divalent metal cations but certain cations like Mg^{2+}, Ca^{2+} and K^{+} could slightly enhance its activity. However, kinetics analysis indicated that compared to other MshAs, CdMshA exhibited a slightly lower affinity towards 1L-Ins-1-P. Furthermore, mutagenic studies on CdMshA combined with molecular modeling of its active center revealed that N28, K81, T137 and R157 in its peptide sequence are the key amino acids for the enzymatic activity. This study has provided some detailed insights into CdMshA, which should be very helpful for the design and development of inhibitors against the enzyme and other potential applications.

Materials and Method

Materials and general methods. Restriction endonucleases were purchased from New England Biolabs. Resins and columns used for Nickel chelation chromatography were obtained from GE Healthcare. PCRs were performed with Q5® High-Fidelity DNA Polymerase that was obtained from NEB. The gel extraction, PCR purification, and plasmid extraction kits were purchased from Sangon Biotech. Other reagents and chemicals were purchased from commercial sources and used without further purification or treatment if not specified otherwise.

Bacterial strains and growth condition. Bacterial strains and plasmids used in this study are listed in Table S1. *E. coli* was cultured in shaking LB broth (180 rpm), and if needed, ampicillin (100 μg/mL) was used as the antibiotic.

DNA isolation and manipulation. General DNA manipulations, transformations, and agarose gel electrophoresis were carried out by standard protocols. Restricted enzyme digestion, ligation, and plasmid purification were performed in accordance with the manufacturers' instructions. DNA

View Article Online
DOI: 10.1039/C7OB00555E

sequencing, gene synthesis, and primer synthesis were carried out by Sangon Biotech, and primers used in this study are listed in Table S1.

Expression and purification of recombinant proteins. After recombinant pET-22b-cdmsha genes were transformed into the *E. coli* BL21 (DE3) host, the recombinant bacterial strains were cultured in LB broth at 37 ^{o}C to an OD_{600} of 0.6~0.8. The culture was cooled to 16 ^{o}C, and then isopropyl β-D-1-thiogalactopyranoside (IPTG, 0.4 mM) was added to induce protein expression, while the cells were allowed to grow for additional 12 h. The cells were harvested through centrifugation, and then re-suspended in the binding buffer (50 mM Tris-HCl, 500 mM NaCl, pH 8.0, four times the original volume), followed by disruption with sonication and centrifugation at 12,000 rpm and 4 ^{o}C for 30 min to remove cell debris. Protein purification was achieved via affinity chromatography using a Nickel chelation column. Accordingly, after cell supernatants obtained above were loaded onto His trap columns, the columns were washed first with Tris-HCl buffer and then with 100 mM imidazole in buffer (50 mM Tris-HCl containing 500 mM NaCl) to elute target proteins carrying a His-tag. Protein samples were analyzed with 10% SDS-PAGE and visualized by Coomassie brilliant blue staining. Pure proteins were dialyzed three times against Tris-HCl buffer (50 mM, pH 8.0) at 4 ^{o}C and finally concentrated by ultrafiltration (Millipore). Protein concentration was determined using the Bradford protein assay kit following the manufacturer's instructions with bovine serum albumin (BSA) as standard. The proteins were stored at -80 ^{o}C.

Evaluation of the enzymatic activity of CdMshA. To the solution of UDP-GlcNAc (1 mM) and 1L-Ins-1-P (0.5 mM) in Tris-HCl buffer (50 mM, pH 8.0, 300 μL) was added CdMshA (60 μg/mL). After the mixture was incubated at 40 ^{o}C, 50 μL aliquots were taken from the mixtures at 1, 2, and 4 h and quenched by adding the same volume of heated acetonitrile (60 ^{o}C) and then incubation at 60 ^{o}C for 10 min. After the resultant samples were cooled with ice and centrifuged (13,000 × g，10 min) to remove insoluble proteins, they were subjected to HPLC analysis. HPLC conditions were: C18 column (4.6 x 150 mm); 10 μL sample for each injection; UV monitor at 260 nm;[13] 1 mL/min flow rate; gradient eluent [0 to 1 min: 100% buffer A; 1 to 11 min: 75% buffer A and 25% buffer B; 11 to 31 min: 100% buffer B; thereafter: 100 % buffer A. Buffer A: tetrabutylammonium hydroxide

View Article Online
DOI: 10.1039/C7OB00555E

(Bu_4NOH) aqueous solution (2 mM, pH 6.5); buffer B: Bu_4NOH solution (10 mM, pH 5.5) in 50% aqueous methanol containing 77 mM of KH_2PO_4]. Under these conditions, the retention times for UDP-GlcNAc and UDP were 20.2 and 21.5 min, respectively.

Enzymatic synthesis of GlcNAc-Ins-3-P.[12] To a solution of UDP-GlcNAc (10 mM) and 1L-Ins-1-P (10 mM) in Tris-HCl buffer (50 mM, pH 8.0, 0.5 mL) was added CdMshA (100 μg). The reaction progress was monitored by HPLC according to the above-described protocol. When the reaction was complete (in *ca*. 4 h), the reaction product, GlcNAc-Ins-3-P, was isolated by HPLC using a 1 mL-Mono Q ion-exchange column following a protocol similar to that previously reported.[12] Thus, the reaction mixture was injected into the Mono Q column and fractionated using a programmed gradient eluent: 0-5 min with 100% buffer A, 20 min with 65% buffer A and 35% buffer B, 25-30 min with 100% buffer B, 33 min with 100% buffer A, where buffer A is 20 mM of aq. ammonium bicarbonate solution and buffer B is 600 mM of aq. ammonium bicarbonate solution. The flow rate was 1 mL/min, monitored at 260 nm. The resultant fractions (1 mL each) were then treated with alkaline phosphatase and assayed for inorganic phosphate using malachite green phosphate assay. Fractions that were found to contain phosphate but did not correlate to the peak at 260 nm were pooled and lyophilized to afford GlcNAc-Ins-3-P (2.2 mg, 95%) as a white solid. ^{1}H NMR (600 MHz, D_2O, with DHO at δ 4.78 ppm as internal reference): δ 5.12 (d, 1 H, J = 3.6 Hz, H-1^{GlcNAc}), 4.38 (br s, 1 H, H-2^{Ins}), 3.96-3.91 (m, 2 H, H-2^{GlcNAc}, H-3^{Ins}), 3.91-3.85 (m, 2 H, H-5,6a^{GlcNAc}), 3.82 (t, 1 H, *J* = 9.6 Hz, H-3^{GlcNAc}), 3.81-3.74 (m, 3 H, H-6b^{GlcNAc}, H-5,6^{Ins}), 3.61 (br d, 1 H, *J* = 10.2 Hz, H-1^{Ins}), 3.50 (t, 1 H, *J* = 9.6 Hz, H-4^{GlcNAc}), 3.35 (t, 1 H, *J* = 9.6 Hz, H-4^{Ins}); ^{13}C NMR (150 MHz, D_2O): δ 174.26, 99.46, 79.20, 74.66, 74.01, 72.38, 71.78, 71.66, 71.03, 70.80, 69.83, 62.37, 53.78, 21.85; ^{31}P NMR (242 MHz, D_2O): δ 2.50. HR ESI-MS *m*/*z*: calcd for $C_{14}H_{25}NO_{14}P^{-1}$ $(M - H)^{-1}$, 462.1018; found, 462.0998.

Evaluation of the substrate selectivity of CdMshA. These assays were performed in triplicate under the same conditions used for the enzymatic activity assays of CdMshA. With 1L-Ins-1-P as the glycosyl receptor, the glycosyl donors evaluated included UDP-GlcNAc, UDP-Gal, UDP-Glc, UDP-GalNAc, and UDP-GlcA. With UDP-GlcNAc as the glycosyl donor, the glycosyl receptors

View Article Online
DOI: 10.1039/C7OB00555E

tested included 1L-Ins-1-P, 1D-Ins-1-P, *myo*-inositol, glucose-6-phosphate, galactose-6-phosphate, mannose-6-phosphate, and glucose-1-phosphate.

Evaluation of the influences of pH, temperature and metal cation on the activity of CdMshA. To evaluate the influence of pH on the activity of CdMshA, the enzymatic reaction was examined in a buffer containing 6.01 g/L of citric acid, 3.89 g/L of KH_2PO_4, 1.77 g/L of boric acid, and 5.27 of g/L barbitone with the pH values adjusted in a range of 4 to 12 using NaOH. The concentrations of UDP-GlcNAc, 1L-Ins-1-P, and CdMshA utilized were the same as that in above experiments. The reaction was kept at 37 ^{o}C for 30 min before it was quenched and subjected to HPLC analysis by the protocols described above. To study the influence of temperature on the activity of CdMshA, the reaction conditions described for the activity assays were used, but the reaction was carried out at varied temperatures including 20, 25, 30, 35, 40, 45, 50 and 55 ^{o}C. After the reaction mixture was kept at certain temperature for 30 min, it was quenched and analyzed by HPLC as described above. Similar experimental setup and conditions were also employed to evaluate the influence of mental ions on the activity of CdMshA, and cations tested included Na^{+}, K^{+}, Ag^{+}, Ni^{2+}, Fe^{2+}, Cu^{2+}, Hg^{2+}, Zn^{2+}, Mn^{2+}, Mg^{2+}, and Ca^{2+} and their concentration was set at 1 mM. The reaction was kept at 37 ^{o}C for 30 min before it was quenched and analyzed.

Analysis of the kinetics of CdMshA-catalyzed reactions. These reactions were carried out under the optimized conditions described for the activity assays (Tris-HCl buffer, pH 8.0, 40 ^{o}C, and 60 μg/mL of CdMshA) employing varied concentrations of UDP-GlcNAc or 1L-Ins-1-P. In one set of experiments, the concentration of 1L-Ins-1-P was 1 mM, while the concentration of UDP-GlcNAc varied from 0.05 to 1 mM. In the other set of experiments, the concentration of UDP-GlcNAc was kept at 1 mM, and the concentration of 1L-Ins-1-P varied from 0.05 to 1 mM. The reactions were quenched at 30 min and analyzed by HPLC as described. The data obtained were used to calculate initial velocities and determine the Michaelis constants (K_m) and maximal velocities (V_{max}) with GrapPad Prism. All assays were performed in triplicate.

Mutagenesis of MshA and evaluation of the enzymatic activity of the mutants. Site-directed mutagenesis of MshA with its D23, M27, N28, K81, Y113, T137 and R157 amino acids substituted

View Article Online
DOI: 10.1039/C7OB00555E

with an alanine residue, respectively, was performed with an Easy Mutagenesis kit according to standard protocols.[33] Briefly, forward and reverse primers were designed, synthesized, and utilized to replace the designated residue (Table S1) with the pET-22b-mshA vector as the template. PCR reactions were performed to obtain the amplified fragments, which were then treated with *Dpn* I enzyme to remove the template and transformed into *E. coli* DH5α. The mutant plasmids were extracted from *E. coli* and sequenced to confirm the mutations in *mshA* gene. The mutant proteins were expressed and purified (SDS-PAGE results in Figure S4) by the above-mentioned protocols. The activities of these mutant enzymes were assessed under the same conditions described for the wild type enzyme. The reactions were kept at 40 ^{o}C for 30 min before they were quenched and analyzed by HPLC. Similarly, the selectivity of these mutant enzymes for glycosyl acceptors were also assessed using UDP-GlcNAc as glycosyl donor. The glycosyl acceptors examined included 1D-Ins-1-P, *myo*-inositol, glucose-6-phosphate, galactose-6-phosphate, mannose-6-phosphate, and glucose-1-phosphate as glycosyl acceptor. All these assays were performed in triplicate.

Synthesis of (+)-2. To a solution of (±)-**1** (200 mg, 0.857 mmol) in the mixture of DCM (2 mL) and pyridine (2 mL) were added triethylamine (TEA, 0.5 ml) and (-)-methyl chloroformate (0.5 ml). The solution was stirred at rt overnight. After TLC showed the completion of reaction, the solvent was removed, and the residue was purified by column chromatography (petroleum ether and ethyl acetate 15:1) to afford the product as colorless syrup. This product was then dissolved in MeOH and allowed to stand at 2 ^{o}C overnight. The resultant crystal was filtered and washed with cold methanol to give (+)-**2** (80 mg, 21%). ^{1}H NMR (600 MHz, $CDCl_3$): δ 7.36-7.25 (m, 5 H, *Ph*), 5.14 (dd, 1 H, J = 4.2, 2.4 Hz, H-2), 4.73 (s, 2 H, CH_2Ph), 4.53 (dt, 1 H, J = 10.8, 4.2 Hz, H-1′), 4.49 (dd, 1 H, J = 7.2, 4.2 Hz, H-1), 4.37 (t, 1 H, J = 7.2 Hz, H-6), 4.02 (dd, 1H, J = 10.8, 7.2 Hz, H-5), 3.87 (dd, 1H, J = 7.8, 2.4 Hz, H-3), 3.60 (dd, 1H, J = 10.8, 7.8 Hz, H-4), 2.09-2.04 (m, 1 H), 1.98-1.89 (m, 1 H), 1.70-1.63 (m, 2H), 1.48, 1.44, 1.43, 1.33 (4 s, 4 × 3 H, 2 × C(CH_3)$_2$), 1.43-1.37 (m, 1H), 1.09-0.97 (m, 2 H), 0.90 (d, 3 H, J = 6.6 Hz, -CH_3), 0.88 [d, 3H, J = 7.2 Hz, CH(CH_3)$_2$), 0.88-0.82 (m, 1H), 0.77 [d, 3H, J = 7.2 Hz, CH(CH_3)$_2$]. The data agreed well with those reported in the literature.[30]

View Article Online
DOI: 10.1039/C7OB00555E

Synthesis of (+)-1. A mixture of (+)-**2** (80 mg, 150 μmol) and KOH in MeOH (3 mL, 25% w/v) was refluxed until TLC indicated the completion of reaction. The solvent was removed in vacuum, and the residue was suspended in saturated aq. ammonium chloride solution and extracted with CH_2Cl_2. The organic phases were pooled, dried (Na_2SO_4), filtered, and concentrated. The residue was purified by column chromatography to produce (+)-**1** as colorless syrup (36 mg, 68%). [α] = +19.5° (c = 0.5, $CHCl_3$); ^{1}H NMR (600 MHz, $CDCl_3$): δ 7.39-7.26 (m, 5H, *Ph*), 4.78 (d, 1 H, *J* = 10.8 Hz, C*H*$_2$Ph), 4.65 (d, 1 H, *J* = 10.8 Hz, C*H*$_2$Ph), 4.45 (dd, 1 H, *J* = 7.2, 3.6 Hz, H-1), 4.35 (t, 1 H, *J* = 7.2 Hz, H-6), 4.18 (dd, 1 H, *J* = 10.2, 7.2 Hz, H-5), 4.05-4.01 (m, 1H, H-2), 3.92 (dd, 1 H, *J* = 7.8, 2.4 Hz, H-3), 3.56 (dd, 1 H, *J* = 10.2, 8.4 Hz, H-4), 2.58 (d, 1 H, *J* = 1.2 Hz, -O*H*), 1.53, 1.44, 1.37 [3 s, 4 × 3 H, 2 × C(C*H*$_3$)$_2$]. Its ^{1}H NMR spectrum was identical to that of (±)-**1** and agreed well with that of the reported.[28-30]

Synthesis of (+)-3. To a solution of (+)-**1** (8 mg, 23 μmol) dissolved in the mixture of pyridine (0.5 mL) and DCM (0.5 mL) was added TEA (0.1 ml) followed by diphenyl chlorophosphate (18 mg, 69 μmol). The solution was stirred at rt overnight. After TLC showed the completion of reaction, the solvents were evaporated in vacuum, and the residue was purified by column chromatography (petroleum ether and ethyl acetate 3:1) to give (+)-**3** as colorless syrup (10 mg, 81%). ^{1}H NMR (600 MHz, $CDCl_3$): δ 7.35-7.13 (m, 15 H, Ph), 4.90 (td, 1 H, *J* = 9.6, 3.0 Hz, H-1), 4.69 (d, 1 H, *J* = 12.0 Hz, C*H*$_2$Ph), 4.65 (d, 1 H, *J* = 12.0 Hz, C*H*$_2$Ph), 4.52-4.47 (m, 1 H, H-2), 4.36 (t, 1 H, *J* = 7.2 Hz, H-6), 4.00 (dd, *J* = 10.8, 7.8 Hz, 1 H, H-5), 3.95 (dd, *J* = 7.8, 3.0 Hz, 1 H, H-3), 3.54 (dd, *J* = 10.8, 8.4 Hz, 1 H, H-4), 1.49, 1.43, 1.38, 1.32 [4 s, 4 × 3 H, 2 × C(C*H*$_3$)$_2$]. The data agreed well with those reported in the literature.[28-30]

Synthesis of 1L-Ins-1-P. To a solution of (+)-**3** (3 mg, 5 μmol) in EtOH (1 mL) were added PtO_2 (1 mg) and 3 drops of concentrated HCl. The mixture was vigorously stirred under H_2 atmosphere at rt for 2 d. The mixture was filtered, and the filtrate was evaporated in vacuum to give 1L-Ins-1-P as a white solid (1 mg, 90%), which did not need additional purification. ^{1}H NMR (600 MHz, D_2O): δ 4.20 (t, 1 H, *J* = 3.0 Hz, H-2), 3.80 (dt, 1 H, *J* = 9.6, 3.0 Hz, H-1), 3.70 (t, 1 H, *J* = 9.6 Hz, H-6), 3.59 (t, 1 H, *J* = 9.6 Hz, H-5), 3.51 (dd, 1 H, *J* = 9.6, 3.0 Hz, H-3), 3.28 (t, 1 H, *J* =

View Article Online
DOI: 10.1039/C7OB00555E

9.6 Hz, H-4). The data agreed well with those reported in the literature.[31]

Synthesis of 1D-Ins-1-P. This synthesis followed exactly the same procedure described for the synthesis of 1L-Ins-1-P, except that (+)-methyl chloroformate, instead of (-)-methyl chloroformate, was used as the auxiliary agent for the racemic resolution. Similar reaction yields were obtained, and the spectroscopic and other analytical data of the synthetic intermediates and final product agreed well with those reported in the literature. [31]

Acknowledgements

This research was supported by the National Basic Research (973) Program (2012CB822102) and the National Natural Science Foundation (21472114 and 21672129) of China. We thank Mr. Zhicheng Gu for some NMR measurements.

References

1. V. K. Jothivasana and C. J. Hamilton, *Nat. Prod. Rep.*, 2008, **25**, 1091-1117.
2. R. C. Fahey, *Biochim. Biophys. Acta*, 2013, **1830**, 3182-3198.
3. G. L. Newton, N. Buchmeier and R. C. Fahey, *Microbiol. Mol. Biol. Rev.*, 2008, **72**, 471–494.
4. M. Rawat, G. L. Newton, M. Ko, G. J. Martinez, R. C. Fahey and Y. Av-Gay, *Antimicrob. Agents Chemother.*, 2002, **46**, 3348-3355.
5. K. S. Ung and Y. Av-Gay, *FEBS Lett.*, 2006, **580**, 2712-2716.
6. D. Sareen, G. L. Newton, R. C. Fahey and N. A. Buchmeier, *J. Bacteriol.*, 2003, **185**, 6736-6740.
7. Y. B. Liu, M. X. Long, Y. J. Yin, M. R. Si, L. Zhang, Z. Q. Lu, Y. Wang and X. H. Shen, *Arch. Microbiol.*, 2013, **195**, 419-429.
8. Y. B. Liu, C. Chen, M. T. Chaudhry, M. R. Si, L. Zhang, Y. Wang and X. H. Shen, *Biotechnol. Lett.*, 2014, **36**, 1453-1459.
9. G. L. Newton and R. C. Fahey, *Arch. Microbiol.*, 2002, **178**, 388–394.
10. M. Hernick, *Expert. Rev. Antiinfect. Ther.*, 2013, **11**, 49-67.
11. F. Fan, M. W. Vetting, P. A. Frantom and J. S. Blanchard, *Cur. Opin. Chem. Biol.*, 2009, **13**,

View Article Online
DOI: 10.1039/C7OB00555E

451-459.

12. M. W. Vetting, P. A. Frantom and J. S. Blanchard, *J. Biol. Chem.*, 2008, **283**, 15834-15844.
13. G. L. Newton, P. Ta, K. P. Bzymek and R. C. Fahey, *J. Biol. Chem.*, 2006, **281**, 33910-33920.
14. G. L. Newton, T. Koledin, B. Gorovitz, M. Rawat, R. C. Fahey and Y. Av-Gay, *J. Bacteriol.*, 2003, **185**, 3476-3479.
15. S. G. Broadley, J. C. Gumbart, B. W. Weber, M. J. Marakalala, D. J. Steenkamp and B. T. Sewell, *Acta Crystallogr. D Biol. Crystallogr.*, 2012, **68**, 1450-1459.
16. G. L. Newton, M. Ko, P. Ta, Y. Av-Gay and R. C. Fahey, *Protein Expr. Purif.*, 2006, **47**, 542-550.
17. A. A. McCarthy, N. A. Peterson, R. Knijff and E. N. Baker, *J. Mol. Biol.*, 2004, **335**, 1131-1141.
18. A. A. McCarthy, R. Knijff, N. A. Peterson and E. N. Baker, *Acta Crystallogr. D Biol. Crystallogr.*, 2003, **59**, 2316-2318.
19. J. T. Maynes, C. Garen, M. M. Cherney, G. Newton, D. Arad, Y. Av-Gay, R. C. Fahey and M. N. James, *J. Biol. Chem.*, 2003, **278**, 47166-47170.
20. G. L. Newton, Y. Av-Gay and R. C. Fahey, *J. Bacteriol.*, 2000, **182**, 6958-6963.
21. D. Sareen, M. Steffek, G. L. Newton and R. C. Fahey, *Biochemistry*, 2002, **41**, 6885-6890.
22. F. Fan, A. Luxenburger, G. F. Painter and J. S. Blanchard, *Biochemistry*, 2007, **46**, 11421-11429.
23. L. W. Tremblay, F. Fan, M. W. Vetting and J. S. Blanchard, *Biochemistry*, 2008, **47**, 13326-13335.
24. F. Fan and J. S. Blanchard, *Biochemistry*, 2009, **48**, 7150-7159.
25. T. Koledin, G. L. Newton and R. C. Fahey, *Arch. Microbiol.*, 2002, **178**, 331-337.
26. M. W. Vetting, S. L. Roderick, M. Yu and J. S. Blanchard, *Protein Sci.*, 2003, **12**, 1954-1959.
27. J. P. Vacca, S. J. deSolms, J. R. Huff, D. C. Billington, R. Baker, J. J. Kulagowski and I. M. Mawer, *Tetrahedron*, 1989, **45**, 5679-5702.
28. L. W. Leung and R. Bittman, *Carbohydr. Res.*, 1997, **305**, 171-179.

View Article Online
DOI: 10.1039/C7OB00555E

29. L. W. Leung, C. Vilcheze and R. Bittman, *Tetrahedron Lett.*, 1998, **39**, 2921-2924.

30. A. Aguilo, M. Martinlomas and S. Penades, *Tetrahedron Lett.*, 1992, **33**, 401-404.

31. K. M. Pietrusiewicz, G. M. Salamonczyk, K. S. Bruzik and W. Wieczorek, *Tetrahedron*, 1992, **48**, 5523-5542.

32. K. R. Winchell, P. W. Egeler, A. J. VanDuinen, L. B. Jackson, M. E. Karpen and P. D. Cook, *Biochemistry*, 2016, **55**, 4654-4665.

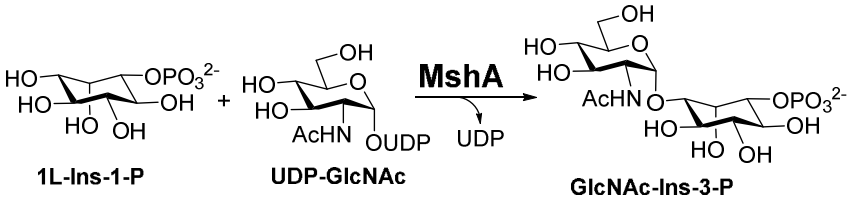
33. S. N. Ho, H. D. Hunt, R. M. Horton, J. K. Pullen and L. R. Pease, *Gene*, 1989, **77**, 51-59.

View Article Online
DOI: 10.1039/C7OB00555E

Biochemical Studies of Inositol *N*-Acetylglucosaminyltransferase Involved in Mycothiol Biosynthesis in *Corynebacterium diphtheria*

Yuchuan Guo,[a,b] Lizhen Wang,[a] Jiatong Guo,[b] Guofeng Gu,[a,*] and Zhongwu Guo[b,*]

TOC graphic:



First-time expression, isolation, biochemical characterization, and mutagenesis studies of a MshA from *Corynebacterium diphtheria* involved in its mycothiol biosynthesis