# **Chapter 1: Introduction**

#### **1.1** The importance of NTD

3,3'-Neotrehalosadiamine (NTD) (Figure 1.1) has been reported to inhibit the growth of *Klebsiella pneumoniae* and *Staphylococcus aureus*.<sup>1-3</sup> NTD's activity against *S. aureus* is noteworthy, as Methicillin-resistant *S. aureus* (MRSA), which is resistant to the  $\beta$ -lactam family of antibiotics, is a well known hospital-acquired infection and is the cause of significant morbidity and mortality.<sup>4</sup> Standard treatment for MRSA is the administration of vancomycin, a glycopeptide that inhibits cell wall synthesis. This antibiotic was sufficient for nearly 40 years until the isolation of vancomycin-intermediate *S. aureus* (VISA) from a clinical setting in 1996.<sup>5</sup> In 2002, fully vancomycin-resistant *S. aureus* (VRSA) was isolated from a clinical setting.<sup>6</sup> This, combined with the fact that *S. aureus* strains have shown varying levels of resistance to several other classes of antibiotics, makes it clear that investigation of any potential antibiotic is a pressing concern.<sup>4</sup>



Figure 1.1: Structure of NTD.

## **1.2** The nature NTD

Structurally, NTD is a 1,1'-linked aminodisaccharide composed of two kanosamine residues linked in an  $\alpha$ , $\beta$ -fashion. This peculiar linkage classifies NTD as a neotrehalose (Figure 1.2), a rare class of disaccharides. In fact, NTD is the only known naturally occurring example of a neotrehalose composed of two aminosugars.<sup>7</sup> It belongs to a group of similar aminosugars that possess antibiotic properties along with trehalosamine, 3-trehalosamine, 4-trehalosamine, and mannosyl glucosaminide. NTD's kanosamine subunits are also a distinguishing feature in the important antibiotic, kanamycin (Figure 1.3). Naturally produced by strains of *Bacillus* 

*pumilis* and *B. circulans*,<sup>1</sup> the production of NTD has been demonstrated in a rifampicin resistant mutant of *B. subtilis*.<sup>3</sup>







Figure 1.3: Structure of kanamycin A. The kanosamine unit is highlighted in blue.

### 1.2.1 Identification of the Ntd operon

*B. subtilis* does not normally produce NTD in any detectable level. However, Ochi *et al.* discovered that a S487L mutation in the  $\beta$ -subunit of RNA polymerase from a rifampicin resistant strain resulted in the up-regulation of NTD production. Transposon mutagenesis

revealed that disruption of the *yhjLKM* (renamed to *ntdABC*) polycistronic operon resulted in the abolishment of NTD production. Sequencing revealed the upstream presence of a LacI family transcriptional regulator, yhjM (renamed to *ntdR*), disruption of which also abolished the production of NTD. It was also found that the presence of NTD caused NtdR, the NTD promoter, to bind upstream of *NtdA*. This evidence, combined with the fact that addition of NTD to a culture of wild type *B. subtilis* actually induces NTD production, indicates NTD acts as its own inducer.<sup>3</sup>

## **1.2.2** The role of NTD in the cell

In order to understand this phenomenon, Ochi et al. analyzed the transcripts of ntdABC, revealing that glcP, a gene encoding a proton-dependent glucose/mannose symport permease and possessing no transcriptional initiator of its own, is cotranscribed with *ntdABC* in approximately 18% of transcripts (Figure 1.4). It was also found that inactivation of glcP up-regulated NTD production, while overexpression of *glcP* down-regulated NTD production. Transposon mutagenesis had already revealed that disruption of *ccpA*, a gene which encodes a transcriptional regulator of carbon catabolite expression associated with the glucose-specific phosphoenolpyruvate-dependent phosphotransferase system (glucose-PTS), impaired NTD production. Mutants lacking glucose-PTS also demonstrated a lowered production of NTD, but in both cases, addition of exogenous NTD restored endogenous NTD production. This indicated that glucose-PTS was not merely providing the substrates for NTD biosynthesis, but was actually affecting its regulation, most likely through some downstream process. Ochi makes the conclusion that NTD operates as a signaling molecule for quorum sensing (the only known example of an aminosugar operating in this role) with GlcP-mediated glucose uptake under its control.<sup>8</sup>

Ochi goes to some lengths to contradict himself in a later publication. Glucose uptake assays performed by his group could not detect any change in glucose uptake rates when either *glcP* was deleted or exogenous NTD was added.<sup>9</sup> According to Paulsen *et al.*, deletion of *glcP* should have resulted in a 30% decrease in glucose uptake,<sup>10</sup> with the corollary conclusion being that the addition of NTD, which up-regulates *glcP* transcription, should increase glucose uptake. Although Paulsen makes a very convincing argument, Ochi disagrees, concluding that the

function of GlcP is not to provide an alternative route of glucose uptake in order to supplement metabolism, but to provide a regulatory element for NTD biosynthesis<sup>9</sup> by sensing free glucose in the cell.<sup>8</sup>



**Figure 1.4:** The *ntdABC* operon and associated elements. In the absence of NTD, NtdR cannot bind upstream of the *ntdABC* operon transcription initiation sequence, inhibiting the transcription of the operon. In the presence of NTD, NtdR binds, promoting transcription of the *ntdABC* operon, along with *glcP* as a readthrough product.

Real-time PCR experiments using glcP deletion mutants showed that NTD biosynthesis was responsible for the down-regulation of the genes alsD (acetolactate synthase), alsS (acetolactate decarboxylase), and yufR (Na<sup>+</sup>/malate symporter) as well as the up-regulation of *licBCAH* operon (components of the lichenan-PTS associated lichenan uptake and metabolism), the *yyaH-maa* operon (an operon with an unidentified function), and *yyaJ* (a probable transporter). Any regulation of these genes by GlcP itself or free glucose was ruled out as *ntdABC* deletion mutants did not show any change in the amount of these genes' transcripts in either the presence or absence of GlcP expression. Finally, perhaps most interesting is that although NTD is directly responsible for its own regulation, the addition of NTD to a culture of *B. subtilis* during the exponential phase only had a significant effect on the expression of the *yyaH-maa* operon. Ochi explains this only by saying that "activation" of the NTD biosynthetic

operon is responsible for the regulation of all of the above genes,<sup>9</sup> but it is likely that some intermediate in NTD biosynthesis is acting as a regulatory element.

# 1.3 Kanosamine biosynthesis

As stated above, NTD is composed of two kanosamine units. However, the origin of these units is unknown. Kanosamine has been known as a component of kanamycin for over fifty years,<sup>11</sup> but was not isolated as its free form until ten years after its initial discovery.<sup>12</sup> It was the first deoxyamino sugar to be isolated and has been isolated from several other *Bacillus* species since,<sup>13-15</sup> although it is unclear whether it is produced in its free form or whether it is the result of the hydrolysis of NTD due to the acidic conditions used in kanosamine's isolation.

Kanosamine is probably best known as the carbon and nitrogen source for 3-amino-5hydroxybenzoic acid (AHBA), a precursor of the rifamycin series of antibiotics (Figure 1.5). This pathway has been highly studied in *Amycolatopsis mediterranei*.<sup>16</sup> Scheme 1.1 displays the method by which *A. mediterranei* converts UDP-glucose to kanosamine and, eventually, kanosamine 6-phosphate. NtdC has been putatively assigned as an oxidoreductase, NtdA as an aminotransferase, and NtdB as a member of the haloacid dehalogenase superfamily,<sup>3</sup> of which many members are phosphatases. A prudent assumption would be to assume that NtdC, NtdA, and NtdB catalyze the same reactions on the same substrates as RifL, RifK, and RifM. However, protein sequence alignments indicate that these enzymes only share 19%, 17%, and 14% sequence identities, respectively. This indicates that the likelihood of them utilizing the same substrates is very low, and is not surprising, as no members of the *Bacillus* genus are known to produce rifamycins. Hence, the NTD operon may code for a novel biosynthesis of kanosamine as well as NTD.



Rifamycin B

**Figure 1.5: Structures of AHBA and rifamycin B.** The rifamycin subunit derived from AHBA is highlighted in blue.



Scheme 1.1: Biosynthetic route to kanosamine and kanosamine 6-phosphate in A. mediterranei.<sup>16</sup>

## 1.4 Proposed biosynthetic pathway for NTD

Members of the genus *Escherichia* are not known to produce either neotrehaloses or kanosamine, but *Ochi et al.* have shown that *Escherichia coli* BL21(DE3) cells that have been transformed with the pUC18-*ntdABC* construct, a construct that contains the *ntdABC* 

biosynthetic operon cloned between the *Eco*RI and *Bam*HI sites of the pUC-18 vector, are capable of producing NTD. Barring some unreported ability for *E. coli* to produce neotrehaloses, this exogenous expression of NTD in *E. coli* suggests that NtdA, NtdB, and NtdC are the only proteins required for the synthesis of NTD from some metabolite common to both *E. coli* and *B. subtilis*.<sup>3</sup> Given that a BLAST<sup>17</sup> search agrees with Ochi *et al.*'s original assignment that the gene products encoded by *ntdA*, *ntdB*, and *ntdC* are members of the pyridoxal phosphate (PLP) dependent aminotransferase family, the haloacid dehalogenase (HAD) hydrolase superfamily, and the nicotinamide adenine dinucleotide (NADH) dependent dehydrogenase family, respectively, it is possible to make an educated guess as to the origin of this metabolite (Scheme 1.2).



Scheme 1.2: Presumed biosynthetic route for NTD.

The simplest route imaginable is to assume that glucose is the starting material. In this case, NtdC catalyzes a selective oxidation, and subsequently, NtdA catalyzes a reductive amination of carbon three to make kanosamine. NtdB is a member of the HAD superfamily and could theoretically work in reverse to condense two kanosamine residues into the final product, NTD. However, biological glycosylations usually require a sugar nucleotide diphosphate, indicating that glucose may be too simple a precursor. A more thorough analysis of what is known and what can be deduced about the individual enzymes is presented below.

### 1.4.1 NtdA

In the original paper describing its cloning, Ochi *et al.* assigned NtdA to the family of PLP-dependent aminotransferases,<sup>3</sup> the general mechanism of which can be found in most introductory biochemistry texts. Briefly, a conserved active site lysine residue covalently binds PLP as an internal aldimine via the formation of a Schiff base. The amino donor then displaces the lysine to form an external aldimine. The lysine residue now acts as a base, catalyzing tautomerization of the external aldimine to an external ketamine. This process is possible due to the electron withdrawing properties of the protonated pyridoxal ring which increase the acidity of the hydrogens attached to the C- $\alpha$ . Nucleophilic attack by water at C- $\alpha$  of the amino donor disrupts the Schiff base of the external ketamine, producing pyradoxamine phosphate (PMP) and a ketone. From here, the process acts in reverse, with attack by PMP on the C- $\alpha$  of the amino acceptor generating a second external aldimine, lysine catalyzing a tautomerization to an external ketamine, and finally, lysine displacing the amino acceptor to release the newly formed amine and reform the internal aldimine. The process follows a two stage, ping-pong mechanism where the first substrate must leave before the second can bind.<sup>18</sup>

A protein BLAST search<sup>17</sup> revealed several proteins of known structure and function that share a sequence identity of 30% or greater with NtdA. The two highest identities were achieved with DesV (PDB ID: 2OGA) and QdtB (PDB ID: 3FRK), both of which are members of the aspartate aminotransferase family. Both these enzymes catalyze the transfer of an amino group from glutamate to C-3 of (d)TDP-3-oxo-sugar pyranoses, placing them into subfamily VI<sub>β</sub> of the aspartate aminotransferases, as opposed to subfamily VI<sub>α</sub> which accept NDP-4-oxo-sugar pyranoses and subfamily VI<sub>γ</sub> which accept *scyllo*-inosose. It is predicted that NtdA also acts on C-3 of a 3-oxo-sugar pyranose. A sequence alignment between NtdA, DesV, and QdtB is displayed in Figure 1.6. The crystal structure of NtdA as the internal aldimine has also been solved and agrees well with the structures of the other members of the VI<sub>β</sub> subfamily (Figure 1.7).<sup>19, 20</sup>



Figure 1.6: Sequence alignment of NtdA with other PLP-dependent aminotransferases from subfamily  $VI_{\beta}$ . Sequence alignment generated with *ClustalW*<sup>21</sup> using the Blosum scoring matrix:<sup>22</sup> Opening gap penalty = 10; Extending gap penalty = 0.05; End gap penalty = 10; Separation gap penalty = 0.05. Image generated using *ESPrit* V2.2.<sup>23</sup> The abbreviations used, references to the published sequences, and databank accession numbers are as follows: NtdA, the gene product of *ntdA* from *B. subtilis*;<sup>3</sup> DesV, glutamate/TDP-4,6-dideoxy-3-oxo- $\alpha$ -D-glucose aminotransferase of *Streptomyces venezuelae* (PDB ID: 2OGA);<sup>24</sup> QdtB, glutamate/TDP-6-deoxy-3-oxo- $\alpha$ -D-glucose aminotransferase of *Thermoanaerobacterium thermosaccharolyticum* (PDB ID: 3FRK).<sup>25</sup> The N-terminal hexa-histidine tag of DesV and the C-terminal hexa-histidine tag of QdtB have been removed in order to improve the alignment.

Kim *et al.* identified four motifs common to the sugar aminotransferases. Motif I has a consensus sequence of (G,D)DEV(I/V) which is completely conserved among the sequences in Figure 1.6. Kim *et al.* report that this motif is unique to sugar aminotransferases, and hence, is most likely to be involved in nucleotide binding.<sup>26</sup> They make reference to a paper on the

nucleotide-binding motif of the chromatin binding protein, Cdc6p, as support for their claim, but the only similarity between the two motifs seems to be the aspartate-glutamate pair found in the Walker B motif of Cdc6p (VLDEMD). In the Walker B motif, this pair of residues is responsible for coordinating a magnesium ion which in turn coordinates the phosphate chain of ATP.<sup>27</sup> However, no sugar aminotransferase structures possess magnesium within the active site, and the crystal structure of QdtB complexed with the external aldimine of its sugar substrate, PLP:TDP-3-aminoquinovose aldimine,<sup>25</sup> shows that motif I is situated on a 7-stranded mixed  $\beta$ sheet located within the interior of the protein with the nucleotide tail of the substrate pointed in the opposite direction (Figure 1.8). Although this motif may be unique to sugar aminotransferases, it does not appear to function directly in nucleotide binding.

Motif II, VHXXG, is within the active site, but makes no obvious contacts with the external aldimine sugar in the QdtB structure. Motif III, (I/L/V)X(I/L/V)(I/L/V)EDXA(Q/H/E), contains a conserved aspartate residue, D222 in NtdA and D157 in QdtB, which is critical for efficient catalysis. In the sugar aminotransferase structures, this residue is positioned in such a way that the  $\gamma$ -carboxyl group cradles the nitrogen atom of the pyridoxal ring, ensuring that it remains in its protonated form. Motif IV,  $S(F/L)(F/Y/H)XXKX_{5-6}(E/D)GG$ , contains the catalytic lysine (K247 in NtdA and K186 in QdtB). The active site and its important residues are displayed in Figure 1.9.



**Figure 1.7: Crystal structure of NtdA as the internal aldimine.**<sup>20</sup> V196, H197, G200, the internal aldimine between PLP and D222, and K247 are represented as their stick figures. Oxygen and nitrogen atoms are coloured bright red and dark blue, respectively, while phosphorous atoms are coloured orange. Image generated using PyMOL V0.99rc6.<sup>28</sup>



**Figure 1.8: Crystal structure of QdtB complexed with PLP:TDP-3-aminoquinovose aldimine (PDB ID: 3FRK).**<sup>25</sup> PLP:TDP-3-aminoquinovose aldimine and the residues of motif I are displayed as their stick structures. Oxygen and nitrogen atoms are coloured bright red and dark blue, respectively, while phosphorous atoms are coloured orange. Carbon atoms of the external aldimine are coloured green. Image generated using *PyMOL* V0.99rc6.<sup>28</sup>



**Figure 1.9: Structure of the active site of QtdB complexed with PLP:TDP-3-aminoquinovose aldimine (PDB ID: 3FRK).**<sup>25</sup> PLP:TDP-3-aminoquinovose aldimine, V131, H132, G135, D157, and K186 are displayed as their stick structures. Oxygen and nitrogen atoms are coloured bright red and dark blue, respectively, while phosphorous atoms are coloured orange. Carbon atoms of the external aldimine are coloured green. Image generated using *PyMOL* V0.99rc6.<sup>28</sup>

# 1.4.2 NtdB

The primary structure of NtdB aligns with members of the haloacid dehalogenase (HAD) superfamily. This ubiquitous group of proteins is named in honor of the first members to be identified. However, most of the members of this family are actually phosphotransferases, with the bulk of these being ATPases and phosphatases.<sup>29</sup> All members of the HAD superfamily possess an  $\alpha/\beta$ -core hydrolase fold similar to the Rossmann fold, where a  $\beta$ -sheet is sandwiched between two layers of  $\alpha$ -helices.<sup>30</sup> This core presents four catalytic loops which are responsible for recognizing the substrates, catalyzing the reaction, and in the case of phosphotransferases, binding magnesium.<sup>29</sup> The crystal structure of NtdB containing Mg<sup>2+</sup> has been reported (PDB ID: 3GYG) (Figure 1.10),<sup>31</sup> suggesting that it is a phosphotransferase. The large,  $\beta$ -sandwich cap domain (yellow) inserted between catalytic loops two (green) and three (magenta) indicates

that NtdB belongs to subfamily II of the HAD superfamily along with trehalose 6-phosphate phosphatase (T6PP) and sucrose 6-phosphate phosphatase (S6PP).



**Figure 1.10: Crystal structure of NtdB from** *B. subtilis* (**PDB ID: 3GYG**).<sup>31</sup> The core and cap domains are coloured blue and yellow, respectively. Loops one through four are highlighted in cyan, green, magenta, and red, respectively. The lone magnesium atom is represented by a grey sphere. D25, D27, T29, T65, K209, D232, and D236 are represented as their stick figures. Oxygen and nitrogen atoms are coloured bright red and dark blue, respectively. Image generated using *PyMOL* V0.99rc6.<sup>28</sup>

Figure 1.11 displays the sequence alignment of these three enzymes. Although the sequence identities are poor at only 10% between NtdB and T6PP and 19% between NtdB and S6PP, motifs I, II, and II are well conserved. Motif I, DXDX(T/V), is situated on loop one (cyan), and is highly conserved among HAD phosphatase and phosphomutase family members. The first aspartate residue corresponds to D25 in NtdB and is instrumental in the formation of the phosphoaspartate intermediate common to HAD phosphatases (Scheme 1.3), as well as coordinating Mg<sup>2+</sup>. The second aspartate, D27, is a distinguishing feature of HAD phosphatases and phosphomutases and is believed to function as an acid or base catalyst.<sup>32</sup> The threonine residue corresponds to T29 and possibly forms a stabilizing hydrogen bond with the phosphate moiety, as well as forms an electrostatic interaction between its backbone carbonyl and Mg<sup>2+</sup>. Motif II, (S/T)GX, is located on loop two (green). The threonine in motif I, by donating a

hydrogen bond to the phosphate moiety. Motif III,  $K(X)_{16-30}(G/S)(D/S)XXX(D/N)$ , spans loops three (magenta) and four (red). The lysine residue, K209 in NtdB, is believed to form a salt bridge with the phosphate moiety. This lysine also forms a hydrogen bond with the nucleophilic aspartate in motif I (D25), stabilizing its negative charge. Finally, the two aspartate residues, D232 and D236 in NtdB, are responsible for coordinating Mg<sup>2+</sup>.<sup>33, 34</sup> For comparison, RifM only conserves the first and last residues of motif I, does not conserve motif II, and does not conserve the lysine in motif III. Figure 1.12 displays the active site of NtdB, along with the residues predicted to be important for its function.



Figure 1.11: Sequence alignment of NtdB with some HAD subfamily II members, as well as RifM. Sequence alignment generated with  $ClustalW^{21}$  using the Blosum scoring matrix:<sup>22</sup> Opening gap penalty = 10; Extending gap penalty = 0.05; End gap penalty = 10; Separation gap penalty = 0.05. Image generated using *ESPrit* V2.2.<sup>23</sup> The abbreviations used, references to the published sequences, and databank accession numbers are as follows: NtdB, the gene product of *ntdB* from *B. subtilis*;<sup>3</sup> T6PP, trehalose 6-phosphate phosphatase of *Thermoplasma acidophilum* TA1209 (PDB ID: 1U02);<sup>35</sup> S6PP, sucrose 6-phosphate phosphatase of *Synechocystis* sp PCC 6803 (PDB ID: 1S2O);<sup>30</sup> RifM, RifM of *Amycolatopsis mediterranei* (NCBI Reference Sequence: NC\_014318.1).<sup>36</sup>



Scheme 1.3: General mechanism of HAD phosphorylases.



Figure 1.12: Structure of the active site of NtdB from *B. subtilis* (PDB ID: 3GYG).<sup>31</sup> The core and cap domains are coloured blue and yellow, respectively. Loops one through four are highlighted in cyan, green, magenta, and red, respectively. The lone magnesium atom is represented by a grey sphere. D25, D27, T29, T65, K209, D232, and D236 are represented as their stick figures. Oxygen and nitrogen atoms are coloured bright red and dark blue, respectively. Image generated using *PyMOL* V0.99rc6.<sup>28</sup>

#### 1.4.3 NtdC

A protein BLAST search<sup>17</sup> identified NtdC as a member of the GFOR/IDH/MocA-C superfamily of oxidoreductases, a group of enzymes that use NAD(P)<sup>+</sup>/NAD(P)H to oxidize/reduce various sugar and inositol substrates. This class of enzymes displays two domains, a nicotinamide dinucleotide-binding domain and a catalytic, sugar-binding domain (Figure 14). A sequence alignment between NtdC and its three closest relatives with known structure, inositol dehydrogenase from *B. subtilis* (IDH) (22% identity),<sup>37, 38</sup> 1,5-anhydro-D-fructose reductase *from Sinorhizobium morelense* (AFR) (PDB ID: 2GLX) (19% identity),<sup>39</sup> and glucose-fructose oxidoreductase from *Zymomonas mobilis* (GFOR) (PDB ID: 1RYD) (17% identity),<sup>40</sup> was performed using ClustalW<sup>21</sup> and the Blosum scoring matrix<sup>22</sup> (Figure 1.14). The sequence of RifL from *Amycolatopsis mediterranei* was also included for comparison, <sup>36</sup> although no structure is available.



**Figure 1.13:** Crystal structure of AFR with NADP<sup>+</sup> bound.<sup>39</sup> NADP<sup>+</sup> is represented as its stick figure. The N-terminal, nicotinamide dinucleotide-binding domain and the C-terminal, catalytic sugar-binding domain are coloured yellow and orange, respectively. Oxygen and nitrogen atoms of the cofactor are coloured bright red and dark blue, respectively, while phosphorous and carbon atoms are coloured orange and green, respectively. Image generated using *PyMOL* V0.99rc6.<sup>28</sup>

The bulk of the conserved residues in these proteins fall within the nicotinamide dinucleotide-binding domain, which is composed of the first 120 or so residues. Although the

number of conserved residues in this domain is limited, the structure of this domain, named the Rossmann fold after the researcher who first recognized its importance in nicotinamide dinucleotide-binding proteins, is well conserved among most NAD(P)<sup>+</sup>/NAD(P)H-dependent oxidoreductases. Rossmann identified a  $\beta\alpha\beta\alpha\beta$ -topology consisting of a three-stranded, parallel  $\beta$ -sheet layered atop the two  $\alpha$ -helices (Cyan in Figure 1.15).<sup>41</sup> Subsequent investigations have revealed that the  $\beta$ -sheet of the Rossmann fold must be extended by at least one additional  $\beta$ -strand in order to properly bind its cofactor. In lactate dehydrogenase (LDH) type oxidoreductases, this fourth  $\beta$ -strand is contributed by a second Rossmann fold (Magenta in Figure 1.15) related to the first by roughly a C2 axis between the folds' interface.

As can be seen in Figure 1.15, the cofactor is held in an extended conformation, cradled within the valley created by the turns connecting the individual strands of the  $\beta$ -sheet to the  $\alpha$ -helices. Each Rossmann fold contains a single conserved motif positioned along these loops. The first conserved motif is part of the so-called fingerprint region and has a general consensus sequence of GXGXX(G/A) (Figure 1.16). The first glycine, G8 in NtdC and G7 in AFR, is strictly conserved and is believed to be necessary in order to create the sharp turn of the first loop. The second glycine, G10 in NtdC, is less well conserved, as is evident from its substitution to a serine at position nine of AFR, but is thought to allow a close contact of the main chain with the phosphate, enhancing the interaction between the negatively charged 2'-phosphate moiety of the NADP<sup>+</sup>/NADPH and the positive dipole of the N-terminus of the first  $\alpha$ -helix.<sup>42</sup> The third glycine appears to be most strongly conserved in NAD<sup>+</sup>/NADH-binding proteins, while an alanine, A13 in NtdC and A12 in AFR, is more common in NADP<sup>+</sup>/NADPH binding proteins.<sup>43</sup> The side chain of alanine is believed to slightly distort the first  $\alpha$ -helix, widening the binding pocket for the accommodation of the 2'-phosphate group in NADP<sup>+</sup>/NADPH.<sup>42</sup> Therefore, the alanine at this position in NtdC suggests that NtdC might be NADP<sup>+</sup>/NADPH-dependent.

																										1						11	0						2	Q.							- 7	3 0	5	
NtdC GFOR AFR IDB RifL	883 						Å.	22		01	· · · ·				21	101	·		· · · ·	11	R	· · ·	- Ch	E	R . NN	28. 1999	NFN1V FCFFF	I Y N I A	Charles and and		terister and	Sel & Lan	COMPANY OF	1211112	AP AP GP	ANTINE	HISHI	I I V NIL	-CO-14 (CO)	LIGGHL	SPATE	HORXE	KCALH	101 PA	AND AN	HALLA MA HALLAN	I I V I V	1010101	UI LI VI AI	日日の日日日
NtdC GFOR AFR IDH Rifl	LI O TAVA	O E E E A	4 NAASA	O.F.K.Y.O.O	SIA12		KATER	E SAL	in a change			· R · · ·		an a paracter a	0.2221		N N S N A		N M N N N	LILLA	E LACAT	o'Annna	on A, A, CA AC	. XIII	No to to to	GAAAL	11 Vi Vi		2110112	911-91	0 0 0 0 0 0 0	In strain 24		NARDI	T IN CALL IN	AATVA	111122	P A A A A A	100000	GKRKA	KHAGAG	I NNN N	10 10 10 10 10 10	VVVV	9 10 10 10	O- N N N N N			122223	1111111
NtdC GFOR AFR IDH RifL	10 LEAD VAD AED SAT	O'ACACH	SPERCO	THE R. P. LEWIS CO., LANSING MICH.	KI VI	22222	AAAEE	1 EKRIS	IN LA LA VIO	9120121	7RKVRL	A1111	SNI SI VQ	C C N V S	13 F) 31 A)	O YRHMR	RGLRV	Y1 R3 R3	90400	Y PA BA	23 23 23 23	1 IRRVL	a LAAGR	O'EANIE		TLAAP		NEL NG	EN CHE	1 LORVI	40 310 10 10 10	IN RE .		I X	K 1 3 3	HTAA	1 F B D R D R D R D R D R D R D R D R D R D	O-EGEIX	NURRE	SV ACR	A1 MC AF	R	RNRNV	1 100 000	L FALVA	0-1-0-000	100 01 01		0111 ···	NR R DR
NtdC GFOR AFR IDH Rifl	I I SELS NYX E SELS NYX E NYX	0.10010	00000	21 22 22 22 22			ATTACA A	1 8 3 1 3 1 1 1 1 3	o lava al	THURST STATE	100 ED 100 ED 100	201-21-	VW IR IR IR	and the second second		0 EGNNG	01400A		IIAYE	THE HT	2977 277 27	2 VAA01	O RYIVG	O ATSIL		N	73	(V					RA	· · · · A	RP		NDAND	2 8 8 8 8 9 8 9 8 9 8 9 8 9 8 9 8 9 8 9	0. HUGKG	C III N N A	R GR	VVE PA	D III D III III			N N N N N N N	P AI VI		YOUVE	30RREL
NtdC GFOR AFR IDH Rifl	LIFETR	a 12 0 0 0 0	Q FILS	VSALV	5 X0 H X		E SIDEA	12125		0.01110.0	211220	ETAGA		G .GIV	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	SIVUEI	0-HOHHH		O Main o	HAGDS	K 9 1 9 1 9 1 9 1 9 1 9 1 9 1 9 1 9 1 9	E LILLE	2 FXGXA	6 X DR101	POPULA	1111111111	00701	I Y H I G	II OO OT	V 121	27 2K	0.13		C1 1+1+104 64		VHLRA	I VIAN ROSE	2 COATS	<0.0 MI0 0.0	V XGI	PASAR	00000	ad \$4.175 Fr	110.01		9 PAPEL	9 9 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	EQUAP	N 10 10 10 10	Y GAL
NtdC GFOR AFR IDH RifL	WAI QLI ALA EIQ	O SHADE	O'LL LL LL	0000000	EAA80	0.03 01 11 11	N D		11 A	0.8.	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	O N H N N	0000000	ヤヤキヤ日	HR SSR	N D O A O A	0-010-1114	FIGAL	C.L. Lol. C.A. L. Lol	40000		NO DIO	3 QVLAV	BERAVS	ALL DIT	MO-LOW	11 A A A A A A A A A A A A A A A A A A	EYVV	E H H H H H H H H	HAAAA	SANAES	0.51112.0	000000	0.01.02.02.03	NYA ERG	VTIVA	CINE CIN	S VGGKA	0.2511111	V. KO	ROPE	- P - P -	- 10 -	2 .01						

Figure 1.14: Sequence alignment of NtdC with members of the GFOR/IDH/MocA-C superfamily. Sequence alignment generated with  $ClustalW^{21}$  using the Blosum scoring matrix:<sup>22</sup> Opening gap penalty = 10; Extending gap penalty = 0.05; End gap penalty = 10; Separation gap penalty = 0.05. Image generated using *ESPrit* V2.2.<sup>23</sup> The abbreviations used, references to the published sequences, and databank accession numbers are as follows: NtdC, the gene product of *ntdC* from B. subtilis;<sup>3</sup> GFOR, glucose-fructose oxidoreductase of *Zymomonas mobilis* (PDB ID: 1RYD);<sup>4040</sup> AFR, 1,5-anhydrofructose reductase of *Sinorhizobium morelense* (PDB ID: 2GLX);<sup>39</sup> IDH, *myo*-inositol dehydrogenase of *B. subtilis* (GeneBank ID: AAA22543.1);<sup>44</sup> RifL, RifL of *Amycolatopsis mediterranei* (GeneBank ID: ADJ42448.1).<sup>36</sup>



Figure 1.15: Structure of the nucleotide-binding domain of AFR complexed with NADP<sup>+</sup>.<sup>39</sup> The nicotinamide nucleotide-binding domain is composed of two Rossmann fold,  $\beta\alpha\beta\alpha\beta$  topologies linked by an  $\alpha$ -helix (green). The N-terminal and C-terminal Rossmann folds are coloured cyan and magenta, respectively. NADP<sup>+</sup> is represented as its stick figure. Oxygen and nitrogen atoms of the cofactor are coloured bright red and dark blue, respectively, while phosphorous and carbon atoms are coloured orange and green, respectively. Image generated using *PyMOL* V0.99rc6.<sup>28</sup>



Figure 1.16: Structure of the fingerprint region of AFR complexed with NADP<sup>+</sup>.<sup>39</sup> The N-terminal and C-terminal Rossmann folds are coloured cyan and magenta, respectively, and are linked by a single  $\alpha$ -helix in green. NADP<sup>+</sup>, G7, S9, A12, and S33 are represented as their stick figures. Oxygen and nitrogen atoms are coloured bright red and dark blue, respectively, while phosphorous and carbon atoms of the cofactor are coloured orange and green, respectively. Image generated using *PyMOL* V0.99rc6.<sup>28</sup>

However, NAD<sup>+</sup>/NADH-binding proteins usually possess a negatively charged residue within the second turn that hydrogen bonds with the 2'-hydroxyl group of NAD<sup>+</sup>/NADH, (D33 in NtdC) while NADP<sup>+</sup>/NADPH-binding proteins usually possess a small uncharged residue at a corresponding position (S33 in AFR). The substitution is necessary in order to avoid an unfavourable electrostatic interaction between the negative side chain of the residue at this position and the negatively charged 2'-phosphate moiety of NADP<sup>+</sup>/NADPH. On its own, the fact that NtdC possesses an aspartate at this position suggests that NtdC should be NAD<sup>+</sup>/NADH-dependent.<sup>45</sup> It has been shown that an A13G mutation in AFR afforded dual-cofactor specificity to the enzyme,<sup>39</sup> while a S116D mutation accomplished the same in GFOR.<sup>45</sup> This seems to suggest that NtdC may possess a natural, dual-cofactor specificity.

The second motif, AGKHVXCEKP, is located on the second loop of the second Rossmann fold structure and appears to be unique to oxidoreductases with sugars or inositols as substrates.<sup>45</sup> The CEKP section of this motif is especially important, as it makes several contacts with the nicotinamide moiety of the cofactor, and may interact with the sugar substrate and catalytic residues (Figure 1.17). The cysteine residue, C89 in NtdC and C92 in AFR, is positioned such that its side chain may form a hydrogen bond with the ring nitrogens of a histidine that is conserved in the GFOR/IDH/MocA-C superfamily of oxidoreductases (H73 in NtdC and H76 in AFR). This histidine forms a hydrogen bond with the 3'-hydroxyl group of the nicotinamide ribose moiety and is supported by possible hydrogen bonding interactions with the main chain nitrogen of the conserved glutamate and the main chain oxygens of the conserved glutamate, lysine, and proline. The side chain carboxylate of the conserved glutamate, E90 in NtdC and E93 in AFR, hydrogen bonds with the nitrogen of the amide moiety on the nicotinamide ring, as well as possibly forms a C-H····O hydrogen bond with NC2 of the nicotinamide ring. The lysine residue, K91 in NtdC and K94 in AFR, may form a cation- $\pi$ interaction with the nicotinamide ring. Mutagenesis studies in AFR have shown that a K94G mutation does not lower that enzyme's affinity for NADPH,<sup>39</sup> supporting the observation made by Kingston et al. that this lysine might only be positioned to interact in this manner when the nucleotide is in its reduced form.<sup>46</sup> It does, however, increase the Michaelis constant for 1,5anhydrofructose and drastically lower the enzymes catalytic efficiency, indicating that it might interact with the sugar substrate and/or the catalytic residues. Figure 1.17 shows how this lysine in AFR is positioned close enough to hydrogen bond with the first aspartate in the catalytic motif, as well as an acetate molecule that may occupy the sugar binding site. The role of the conserved proline, P92 in NtdC and P95 in AFR, is less clear, but the fact that the peptide bond between the lysine and proline adopts a *cis*-conformation suggests that it might be necessary to force the second loop into the correct orientation for cofactor binding.<sup>39</sup>



**Figure 1.17: Structure of the CEKP and catalytic motifs of AFR.**<sup>39</sup> The N-terminal and C-terminal Rossmann folds are coloured cyan and magenta, respectively, and are linked by a single  $\alpha$ -helix in green. NADP<sup>+</sup>, H76, C92, E93, K94, P95, D176, H180, and D183 are represented as their stick figures. Oxygen, nitrogen, and sulfur atoms are coloured bright red, dark blue, and yellow, respectively, while phosphorous and carbon atoms of the cofactor are coloured orange and green, respectively. Image generated using *PyMOL* V0.99rc6.<sup>28</sup>

The catalytic motif, DXXXHXXD, is conserved in NtdC, AFR, and IDH, but not GFOR; GFOR is proposed to possess a type of catalytic dyad between K129 (corresponding to K91 of the CEKP motif in NtdC) and T217 (corresponding to H180 in NtdC) where the lysine hydrogen bonds to the tyrosine, lowering tyrosine's pK<sub>a</sub> so that it may better act as a general acid/base for catalysis.<sup>46</sup> AFR and IDH are proposed to possess another kind of catalytic dyad where the first aspartate residue, D176 in NtdC and AFR and D172 in IDH, forms a hydrogen bond with the conserved histidine residue, H180 in NtdC and AFR and H176 in IDH, depressing histidine's pK<sub>a</sub> so that it may better act as a general acid/base at physiological pH (Scheme 1.4). In AFR, mutagenesis of D176 to alanine increased the Michaelis constant of 1,5-anhydrofructose by 6-fold and decreased the turnover number by nearly 200-fold, which, oddly, is a greater decrease than a H180A mutation of the primary catalytic residue which had no effect on the binding of 1,5-anhydrofructose and only a 60-fold decrease in the turnover number. Unfortunately, the

effect of this mutation on the Michaelis constant of NADPH was not studied.<sup>39</sup> The corresponding mutations in IDH had slightly different results. A D172N mutation increased the  $K_m$  of NAD<sup>+</sup> 14-fold, the  $K_m$  of *myo*-inositol by 15-fold, and decreased the turnover number by 31-fold. A H176A mutation increased the  $K_m$  of NAD<sup>+</sup> four-fold, the  $K_m$  of *myo*-inositol by 26-fold, and decreased the turnover number by 250-fold, so that overall, the mutation of the primary catalytic residue was more detrimental to the enzyme's activity. The role of the second aspartate residue, D183 in NtdC and AFR and D179 in IDH, is less defined, but it is probably involved in substrate binding, as a D179N mutation in IDH resulted in a five-fold increase in the  $K_m$  of NAD<sup>+</sup> and a six-fold decrease in the  $K_m$  of *myo*-inositol, while increasing the turnover number by only 26%.<sup>47</sup>



Scheme 1.4: Proposed mechanism of NAD(P)<sup>+</sup>/NAD(P)H-dependent oxidoreductases.

#### **1.5 Glycosylation reactions**

Glycosylation reactions are vital to the cell, as they are necessary for the formation of highly branched sugars for energy storage, complex structural sugars found in the cell wall of some bacteria, and for glycosylated proteins necessary for proper location of the protein and for signaling interactions. The formation of a glycosidic bond requires energy. As a result, most glycosylation reactions require an activated sugar in the form of a sugar nucleotide diphosphate. Sugar 1-phosphates are converted to NDP-sugars by displacement of the  $\beta$ - and  $\gamma$ -phosphate groups to form inorganic pyrophosphate (PP<sub>i</sub>). The rapid and highly energetically favourable hydrolysis of PP<sub>i</sub> by inorganic pyrophosphatase drives the equilibrium forward. The NDP group of the newly formed NDP-sugar is a good leaving group for nucleophilic displacement, promoting the glycosylation reaction, as well as provides a large surface area for protein-substrate interactions. Glycosylation reactions can also be carried out though the use of glycosyltransferases, enzymes which derive the energy necessary to form a glycosidic bond by hydrolyzing another glycosidic bond, such as is the case with glycosyl-(4-6)-transferase which transfers six- or seven-residue 1,4-linked glucose oligosaccharides from one branch of glycogen to form a new 1,6-linked branch of the glycogen chain.<sup>48</sup>

Glycosylation reactions are a staple of carbohydrate chemistry. Fischer reported one of the first reliable glycosylation reactions using glucose dissolved in methanolic hydrogen chloride to produce methyl  $\alpha$ -D-glucopyranoside. The reaction, however, was cumbersome, as it required the heating of the reaction mixture in a sealed tube.<sup>49</sup> Fischer's contemporaries, Koenigs and his student Knorr, reported the synthesis of  $\beta$ -glucose and mannose glycosides using the 1-bromoderivatives of peracetylated sugars as glycosyl donors and silver salts as catalysts.<sup>50</sup> Despite being published at the turn of the last century, this venerable method is still in use today,<sup>51</sup> as the reaction usually suffers from low yields due to the hydrolysis of the anomeric halogen, self-coupling of the glycosyl donor, and formation of the orthoester between glycosyl donors that possess a participating group at C-2.<sup>52</sup> Many attempts have been made to improve this process, most notably the substitution of mercuric salts<sup>53</sup> or mercuric cyanide instead of the expensive and light-sensitive silver compounds used by Koenigs and Knorr.

Several other glycosylating agents have been developed using the same  $S_N1$  mechanism employed by the Koenigs-Knorr method (Figure 1.18). These include orthoesters,<sup>54</sup> thioglycosides,<sup>55</sup> thioimidates,<sup>56</sup> trichloroacetimidates<sup>57</sup> (an evolution of older trichloroacetate<sup>58</sup> and imidate chemistry<sup>59</sup>), phosphates,<sup>60</sup> 4-*n*-pentenyl glycosides,<sup>61</sup> and 1-hydroxy glycosides,<sup>62</sup> as well as others (For reviews see Toshima *et al.*<sup>63</sup> and Demchenko).<sup>64</sup>



Glycosyl halides



Orthoesters







Thioglycosides

Thioimidates





Trichloroacetimidates

4-n-Pentenyl glycosides

1-Hydroxy glycosides

Figure 1.18: Structures of several glycosylating agents.

## **1.5.1** The challenge of stereochemistry with two anomeric centers

NTD's most intriguing feature is its  $\alpha$ , $\beta$ -linkage. Naturally occurring neotrehalose can be found in honey,<sup>65</sup> but in low enough quantities to make separation and purification from this natural source impractical. This is disheartening, as the reliable, stereospecific formation of 1,1'-glycosidic bonds is very difficult. It is so troublesome that the production of trehalose analogues, including 3,3'-trehalosadiamine,<sup>66</sup> is accomplished through the careful protection and deprotection of trehalose, which is easily obtained from a variety of sources.<sup>67-69</sup>

Lemieux *et al.* reported the first synthesis of a naturally occurring 1,1'-linked sugar with the synthesis of sucrose from Brigl's anhydride (tri-*O*-acetyl-1,2-anhydro- $\alpha$ -D-glucopyranose) and 1,3,4,6-tetra-*O*-acetyl-D-fructofuranose.<sup>70</sup> This synthesis was particularly challenging, as not only did the linking of the anomeric centers of two disparate sugars result in the possibility of four stereoisomers, of which only one was desired, but the glycosyl acceptor was also an unreactive tertiary hydroxyl group. Theoretically, NTD should be less challenging to synthesize as it is composed of two identical kanosamine units. Hence, the  $\alpha$ , $\beta$ - and  $\beta$ , $\alpha$ -isomers are equivocal, and only three diastereomers of the trehalose, neotrehalose, and isotrehalose forms are possible. Additionally, the glycosyl donor is merely a secondary alcohol and should be more reactive. Further evidence for this is given by the fact that trehaloses were often produced as self-condensation products in earlier attempts at the synthesis of sucrose.<sup>71</sup> However, this is not to say that the synthesis of neotrehalose is trivial.

Disregarding the yield of the glycosylation reaction, there is still the challenge of synthesizing and isolating the desired stereoisomer. Controlling the stereochemistry of at least one of the anomeric centers should be possible. It is widely reported that pyranoses possessing an acyl group at C-2 predominantly form the 1,2-*trans*-product. This is due to the intramolecular, acyloxonium ion that is formed by these "participating" groups, restricting nucleophilic attack from one side of the ring (Scheme 1.5).<sup>72</sup>



Scheme 1.5: Intramolecular, acyloxonium ion participation and its effect on the anomeric configuration of a glycosylation product.

Forming a 1,2-*cis*-product is more challenging. Using an ether type protecting group at C-2, a "non-participating" group, allows attack from both sides of the ring. Scheme 1.6 displays this process for a halageno-glucose molecule. Under Lewis acid catalyzed conditions, route D and E are the major competing processes. Although the  $\alpha$ -product formed by route E is usually the thermodynamically favoured product due to the anomeric effect, route D is usually favoured, as equatorial attack is usually less sterically hindered, and glycosylation reactions are usually not reversible.

Increased  $\alpha$ -selectivity has been achieved under neutral conditions by introducing additional halide ions through the use of tetraalkylammonium bromide<sup>73</sup> or iodide.<sup>74</sup> In this case, a rapid equilibrium is formed between the  $\alpha$ - and  $\beta$ -isomers of the glycosyl halide. This would have little significance if the reaction proceeded through an S<sub>N</sub>1 mechanism, as the reactive species would still be the central, unshielded, oxocarbenium ion, but the increase in observed  $\alpha$ selectivity for these reactions indicates that this is not the case. Later papers attribute this phenomenon to an S<sub>N</sub>2 style reaction between the  $\beta$ -isomer of the glycosyl halide and the incoming nucleophile in the absence of a Lewis acid promoter, but Lemieux *et al.* originally attributed this phenomenon to the S<sub>N</sub>i style reaction depicted in Scheme 1.6. The  $\alpha$ -isomer is highly stabilized due to the anomeric effect, resulting in a slow step B, while the  $\beta$ -isomer, although present in small quantities, experiences no such stabilization and reacts much more quickly through H.<sup>73</sup>



**Scheme 1.6: Mechanism of Koenigs-Knorr glycosylations in the presence and absence of added halide ions.** The competing routes D and E are in bold.

Unfortunately, controlling the stereochemistry of the glycosyl acceptor is more challenging. Sugars possessing a free hydroxyl group at C-2 can easily anomerize under glycosylating conditions in the presence of even a trace amount of water, making the use of difficult to obtain anomerically pure sugars superfluous. This actually appears to be beneficial in the case of 2,3,4,6-tetra-*O*-acetyl-D-glucopyranoside, as the  $\beta$ -anomer seems to react much more readily than the  $\alpha$ -anomer, primarily yielding the  $\beta$ -configuration for the resulting glycoside.<sup>75</sup>

Selectivity for the  $\alpha$ -anomer appears to be obtainable by using 2,3,4,6-tetra-*O*-benzyl-D-glucopyranoside as the glycosyl acceptor, as trehalose and neotrehalose, but not isotrehalose were apparent after the coupling of this compound with 2,3,4,6-tetra-O-benzyl-1-O-(trichloroacetamidyl)-D-glucopyranose.<sup>76</sup>

# **1.6** List of Objectives

The long term goal of this project is the complete elucidation of the process by which NTD is synthesized in *B. subtilis*, including the nature of the substrates, the reactions catalyzed by each of the three enzymes, and the mechanisms by which these enzymes facilitate these reactions and recognize their substrates. This thesis focuses on what is predicted to be the first enzyme in the pathway, NtdC. The objectives of this thesis were:

- The individual subcloning of the three NTD biosynthetic genes
- The expression and purification of NtdC
- The identification of the substrates of and the reaction catalyzed by NtdC
- The derivation of kinetic constants for the reaction catalyzed by NtdC under the chosen reaction conditions
- The determination of the kinetic mechanism displayed by NtdC under the chosen reaction conditions
- The investigation of the importance of NtdC's predicted catalytic residues, D176 and H180
- The synthesis of kanosamine for the eventual study of NtdA and NtdB
- The development of a methodology for synthesizing NTD by synthesizing neotrehalose
- The synthesis of  $3-[^{2}H]$ -D-glucose for investigation of a kinetic isotope effect
- The synthesis of 3-deoxy-3-fluoro-D-glucose as an inhibitor

# **Chapter 2: Results and Discussion**

# 2.1 Subcloning of *ntdA*, *ntdB*, and *ntdC*

The NTD operon was received as the pUC18-*ntdABC* construct, in which the DNA encoding the three synthetic genes had been cloned between the *Eco*RI and *Bam*HI sites of the pUC-18 vector.<sup>3</sup> In order to study the individual genes, it was necessary to subclone each of the *ntd* genes separately.

The vector, pET-28b, was selected as the host vector, as it possesses a number of features that make it desirable for protein expression, most notably the ability to control gene expression using isopropyl- $\beta$ -D-1-thiogalactopyranoside (IPTG), thanks to the presence of the T7*lac* promoter/operator,<sup>77</sup> and the ability to introduce an N-terminal hexa-histidine tag to the target protein, providing a simple means of purification via nickle-chelation.<sup>78</sup> The multiple cloning site for pET-28b is displayed in figure Figure 2.1. The genes were subcloned between the *NdeI* and *Bam*HI sites of the pET-28b vector, as the *NdeI* site contains a start codon in its sequence, minimizing the need for frame shifting and optimizing the length of the linker between the hexa-histidine tag and the recombinant protein, and because a *Bam*HI site had already been introduced at the terminus of *ntdC* during the original cloning of the *ntdABC* operon.<sup>3</sup>

T7 Pro	moter		lac O	perator		<u>Xbal</u>							
AATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCCTCTAGAAATAA													
	rbs	_	Ncol	_	His-Tag								
TTTTGTTTAACT	'TTAAGAAGGA	AGATAT	TACCATG	GGCAGCA	GCCATCATC.	ATCATCATCAC							
	Thrombir	n	Ndel	Nhel	T7-Tag								
AGCAGCGGCCTG	GTGCCGCGC	GGCAGC	CATATG	GCTAGCA	IGACTGGTG	GACAGCAAGGT							
BamHI Eco	RI Sacl	Sall	HindIII	Eagl Notl	Xhol	His-Tag							
CGGGATCCGAAT	TCGAGCTCC	GTCGAC	CAAGCTT	GCGGCCG	CACTCGAGC.	ACCACCACCAC							
CACCAC <mark>TGA</mark> GAT	CCGGCTGCT	AACAAA	AGCCCGA	AAGGAAG	CTGAGTTGG	CTGCTGCCACC							

## **T7 Terminator**

 ${\tt GCTGAGCAATAACTAGCATAACCCCTTGGGGGCCTCTAAACGGGTCTTGAGGGGGTTTTTTG}$ 

Figure 2.1: The multiple cloning site of pET-28b. The start and stop codons are highlighted in green and red, respectively.

Ligation-based cloning was chosen over PCR-based cloning as our lab had more experience with the technique. However, the gene, *ntdA*, possessed an internal *NdeI* site which would have prevented the cloning of this gene. Fortunately, a silent T1152C mutation (numbered according to the *ntdA* gene) could be made that would eliminate the *NdeI* site without altering the primary structure of the gene product. Primers *ntdA*-T1152C-Forward and *ntdA*-T1152C-Reverse were designed and used for the PCR based mutagenesis of the target nucleotide. Figure 2.2 displays *NdeI* restriction maps for pUC18-*ntdABC* both before and after the mutation. Wild type pUC18-*ntdABC* should be cut twice, leaving two fragments, one of 2309 bp and one of 3592 bp, while mutant pUC18-*ntdABC*-T1152C plasmids should be cut once, leaving a 5901 bp fragment. Removal of the offending *NdeI* site was verified by performing *NdeI* digests on both pUC18-*ntdABC* and pUC18-*ntdABC*-T1152C and visualizing their products on a 1% agarose gel stained with ethydium bromide (Figure 2.3).



Figure 2.2: Ndel restriction map of pUC18-ntdABC and pUC18-ntdABC-T1152C. Figure generated using NEBcutter V2.<sup>79</sup>



**Figure 2.3: Ethydium bromide stained 1% agarose gel of an** *Nde***I digest of the T1152C mutant of pUC18***-ntdABC***.** Lane 1: NEB 1 kbp DNA ladder. Lane 2: Wild-type pUC18*-ntdABC*. Lane 3: pUC18*-ntdABC*-T1152C.

PCR-based gene amplification was used to simultaneously amplify the desired genes while introducing *Nde*I and *Bam*HI restriction sites at their 5'- and 3'-termini, respectively. Due to difficulties with yield, PCR products were not visualized, and hence not purified, using gel electrophoresis. Instead, the parental DNA was digested using *Dpn*I and the PCR products purified using a QIAquick PCR Purification kit. This kit removed reaction buffers and reagents, as well as any DNA fragments less than 50 bp, leaving only the desired PCR products in the eluant. Gene fragments were produced by digesting the PCR products with *Nde*I and *Bam*HI and were then ligated into doubly digested pET-28b using T4 DNA ligase.

The BglII restriction maps for empty pET-28b and pET28b-*ntdC* are depicted in Figure 2.4. Empty pET-28b possesses only one BglII restriction site, resulting in a single 5368 bp fragment, while pET28b-*ntdC* contains two BglII sites, resulting in two fragments, one of 1142 bp and one of 5244 bp. A picture of an agarose gel verifying a successful insertion of the *ntdC* gene is shown in Figure 2.5.

The *Eco*RV restriction maps for empty pET-28b and pET28b-*ntdB* are depicted in Figure 2.6. Empty pET-28b possesses only one Eco*RV* restriction site, resulting in a single 5368 bp fragment as above, while pET28b-*ntdB* contains two *Eco*RV sites, resulting in two fragments, one of 2073 bp and one of 4121 bp. A picture of an agarose gel verifying a successful insertion of the *ntdB* gene is shown in Figure 2.7.

The *Xmn*I restriction maps for empty pET-28b and pET28b-*ntdA* are depicted in Figure 2.8. Empty pET-28b possesses two *Xmn*I restriction sites, resulting in two fragments, one of 2033 bp and one of 3335 bp, while pET28b-*ntdA* contains three *Xmn*I sites, resulting in three fragments, one of 1480 bp, one of 2033 bp, and one of 3146 bp. A picture of an agarose gel verifying a successful insertion of the *ntdA* gene is shown in Figure 2.9.



Figure 2.4: BglII restriction map of pET-28b and pET28b-ntdC. Figure generated using NEBcutter V2.79



**Figure 2.5: Ethydium bromide stained 1% agarose gel of a** *Bg/***II digest of pET28b-***ntdC*. Lane 1: Invitrogen 1 kb Plus DNA Ladder. Lane 2: Empty pET-28b. Lane 3: pET28b-*ntdC*.



Figure 2.6: EcoRV restriction map of pET-28b and pET28b-ntdB. Figure generated using NEBcutter V2.79



Figure 2.7: Ethydium bromide stained 1% agarose gel of an *Eco*RV digest of pET28b-*ntdB*. Lane 1: NEB 1 kbp DNA ladder. Lane 2: pET28b-*ntdB*.



Figure 2.8: XmnI restriction map of pET-28b and pET28b-ntdA. Figure generated using NEBcutter V2.79



**Figure 2.9: Ethydium bromide stained 1% agarose gel of an** *Xmn***I digest of pET28b***-ntdA***.** Lane 1: NEB 1 kbp DNA ladder. Lane 2: Empty. Lane 3: pET28b*-ntdA*.
DNA sequencing confirmed the successful subcloning of all three genes. However, the sequences of *ntdB* and *ntdC* both differed from those reported in the *B. subtilis* genome by a single nucleotide.<sup>80</sup> In the case of *ntdB*, the deletion of a thymidine nucleotide at position 845 (numbered according to the *ntdB* gene) would result in a truncated NtdB protein of 282 instead of 286 residues. For *ntdC*, the mutation of a guanidine nucleotide to an adenine nucleotide at position 680 would result in a conservative R227K mutation in the NtdC protein. In order to determine whether these discrepancies were the result of subcloning, sections of the pUC18-*ntdABC* plasmid containing these points were sequenced. The results showed that both of these mutations were present in the plasmid received from Japan. The assumption was made that these two mutations were either artefacts of the original *B. subtilis* genome sequencing project or mutations produced during the creation of the pUC18-*ntdABC* plasmid. As it had been shown that the proteins encoded by the pUC18-*ntdABC* plasmid were still capable of synthesizing NTD in *E. coli*,<sup>3</sup> it was not necessary to mutate the subcloned genes at this time – in either case, a set of functioning genes had been individually subcloned successfully.

The *ntdC* gene was also subcloned into pET-17b. This was done so that the gene product could be expressed in its native form in case the His-tagged fusion protein was not active or proved difficult to crystallize. The vector, pET-17b, encodes a T7-tag. However, the start codon for this tag is part of an *NdeI* site, meaning any gene cloned into the *NdeI* site will be expressed without a tag. As *ntdC* had already been cloned into the *NdeI* and *Bam*HI sites of pET-28b, and pET-17b possesses a *Bam*HI site downstream of its *NdeI* site, subcloning into pET-17b was simply a matter of performing an *NdeI/Bam*HI digest on pET28b-*ntdC* and ligating the gene fragment into a likewise digested pET-17b. Figure 2.10 displays *Bgl*II restriction maps for pET-17b and pET17b-*ntdC*. Empty pET-17b should be cut once, leaving a 3306 bp fragment, while pET17b-*ntdC* plasmids should be cut twice, leaving two fragments, one of 1075 bp and one of 3225 bp. Figure 2.11 shows a picture of an agarose gel which verifies the successful insertion of the *ntdC* gene.



Figure 2.10: Bg/II restriction map of pET-17b and pET17b-ntdC. Figure generated using NEBcutter V2.79



**Figure 2.11: Ethydium bromide stained 1% agarose gel of a** *Bgl***II digest of pET17b**-*ntdC*. Lane 1: NEB 1 kbp DNA ladder. Lane 2: Empty pET-17b. Lane 3: pET17b-*ntdC*.

## 2.2 Expression and purification of NtdC

In their demonstration of the exogenous production of NTD in *E. coli* cells, Ochi *et al.* showed that NtdA, NtdB, and NtdC were expressible in BL21 (DE3) cells and that these genes were not toxic to the cell.<sup>3</sup> As a preliminary experiment to determine whether the His-tagged versions of these proteins would behave similarly, BL21 (DE3) cell cultures which had been transformed with plasmids containing one of the *ntd* genes were grown in kanamycin supplemented Luria broth and induced with IPTG. After three hours, the gene products were identifiable in the supernatants of the crude cell lysates (Figure 2.12). Bands were identifiable at 50 kDa (NtdA), 33 kDa (NtdB), and 39 kDa (NtdC), showing expression was good for all three proteins. Attempts to express untagged NtdC in BL21 (DE3) were unsuccessful. However, changing cell lines to BL21 Gold® improved expression dramatically. Figure 2.13 shows an acrylamide gel demonstrating the excellent expression of untagged NtdC in BL21 Gold®.



**Figure 2.12:** Coomassie blue stained 4.5% acrylamide gel of the supernatants of crude cell lysates of BL21 (DE3) cells transformed with various plasmids. Lane 1: Invitrogen Benchmark ladder. Lane 2: pET28b-*ntdA*. Lane 3: pET28b-*ntdB*. Lane 4: pET28b-*ntdC*. Lane 5: Induced pET-28b. Lane 6: Uninduced pET-28b.



**Figure 2.13:** Coomassie blue stained 4.5% acrylamide gel of the supernatants of crude cell lysates of BL21 Gold® cells transformed with pET17b-*ntdC* at various times after induction. Lane 1: SigmaMarker Low Range molecular marker. Lane 2: 0 hr. Lane 3: 3 hr. Lane 4: Over night. Lane 5: Uninduced overnight. Lane 6: Induced pET-17b overnight.

As the assumed principle enzyme in the NTD pathway, NtdC was chosen as the first protein for purification and kinetic analysis. His-tagged NtdC was expressed in a manner similar to IDH and purified by nickel affinity chromatography using HiTrap Chelating HP columns. When purified according to the manufacturer's instructions, NtdC started to precipitate before if could be dialyzed into its storage buffer. Supplementing all purification buffers with 20% glycerol stabilized the protein long enough for it to be dialyzed without precipitating. Figure 2.14 shows an acrylamide gel of column eluants containing purified NtdC. On average, 1.6 mg of purified NtdC was obtained from a 50 mL culture.



**Figure 2.14: Coomassie blue stained 4.5% acrylamide gel of purified NtdC fractions.** Lane 1: NEB protein marker. Lane 2: Empty. Lanes 3-11: Eluant fractions 1-9. Lane 12: Pure NtdC.

## 2.3 The assay of NtdC

The biosynthetic route to NTD proposed in Scheme 1.2 indicates that glucose should be the simplest sugar substrate oxidized by NtdC. This was found to be the case, as a positive increase in absorbance could be measured when glucose was reacted with NAD<sup>+</sup> in the presence of NtdC under conditions identical to those previously used to assay IDH.<sup>81</sup> A pH-rate profile (Figure 2.15) revealed that the pH optimum for the oxidation of glucose was pH 9.0, with greater activity in tris(hydroxymethyl)aminomethane titrated with hydrochloric acid (Tris·HCl) than in 2-amino-2-methyl-1-propanol titrated with HCl (AMP·HCl). Assays at physiological pH were not investigated, as pH 7 is well below the pH optimum for the enzyme. As well, high pH optimums for the oxidation reactions of GFOR/IDH/MocA-C family members are common, such as is the case with IDH possessing a pH optimum of 9 for the oxidation of *myo*-inositol.<sup>81</sup> Neither the addition of 1 mM MgSO<sub>4</sub> nor the addition of 1 mM ethylenediaminetetraacetic acid (EDTA) had any discernible effect on the reaction rate. Addition of 10 mM 2-mercaptoethanol did not increase the rate of the reaction.



Figure 2.15: NtdC activity using glucose and NAD<sup>+</sup> at varying pH. Conditions: 20 mM glucose, 1 mM NAD<sup>+</sup>, 25 °C, (♦) 100 mM Tris·HCl, (■) 100 mM AMP·HCl.

### 2.3.1 NtdC has no hydrolytic activity

NtdC possesses the characteristic active site DXXXHXXD motif common to sugar dehydrogenases, the GXGXXA and KHVXCEKPXA motifs common to the Rossmann-fold, and only a low identity with the family 4 glycosidases (<15%) (Figure 2.16), a family of NAD<sup>+</sup>- dependent hydrolases.<sup>82-83</sup> However, to rule out any possible glycosidase activity, such as might be the case if NtdC catalyzed the conversion of UDP-glucose to 3-oxo-glucose, NtdC was tested with the two activated glycosides, *p*-nitrophenyl- $\beta$ -D-galactopyranoside and *p*-nitrophenyl- $\alpha$ -D-galactopyranoside, under reducing conditions and in the presence of Mg<sup>2+</sup> and NAD<sup>+</sup>, an experiment similar to those previously reported.<sup>82, 83</sup> Neither at pH 9.0, the pH optimum for glucose oxidation, nor at pH 7.0, the pH optimum for family 4 glycosidases, was any hydrolysis observed. Therefore, it was assumed that NtdC possessed no hydrolytic activity.



Figure 2.16: Sequence alignment of NtdC with members of glycosidase family 4. Sequence alignment generated with  $ClustalW^{21}$  using the Blosum<sup>22</sup> scoring matrix:<sup>22</sup> Opening gap penalty = 10; Extending gap penalty = 0.05; End gap penalty = 10; Separation gap penalty = 0.05. Image generated using ESPrit v2.2.<sup>23</sup> The abbreviations used, references to the published sequences, and databank accession numbers are as follows: NtdC, the gene product of *ntdC* from *B. subtilis*;<sup>3</sup> GlvA, maltose 6-phosphate hydrolase of *B. subtilis* (NCBI Reference Sequence: NP\_388699.1);<sup>80</sup> BgIT, cellobiose 6'-phosphate hydrolase of *Thermotoga maritima* (SwissProt identifier: Q9X108);<sup>82</sup> MalH, maltose 6-phospate hydrolase of *Fusobacterium mortiferum* (GeneBank ID: AAB63015.1);<sup>84</sup>; AgaL, alpha-galactosidase of *E. coli* (SwissProt identifier: P06720).<sup>85</sup>

#### 2.3.2 NtdC is an exclusively NAD<sup>+</sup>/NADH-dependent oxidoreductase

Experiments performed by Hongyan Zheng using glucose as the sugar substrate determined that NtdC is exclusively NAD<sup>+</sup>-dependent. As discussed in the introduction, the fingerprint region of NtdC's Rossmann fold possesses characteristics associated with both NAD<sup>+</sup>/NADH- and NADP<sup>+</sup>/NADPH-dependent oxidoreductases. However, as mentioned in the results, NtdC is strictly NAD<sup>+</sup>/NADH-dependent. As stated previously, an A13G mutation in AFR afforded dual-cofactor specificity to the enzyme,<sup>39</sup> while a S116D mutation accomplished the same in GFOR.<sup>45</sup> What was not mentioned was that these two mutations produced different results in the two enzymes.

In the case of AFR, in addition to conferring dual substrate specificity to AFR, the A13G mutation also lowered the Michaelis constant for NADPH five-fold while doubling the turnover number. In contrast to the results obtained for the S116D mutation of GFOR, the S33D mutation of AFR completely abolished activity. The A13G/S33D double mutant, like the A13G mutant was active with both NADH and NADPH, but the Michaelis constant and turnover number for NADPH decreased by 50-fold and 64-fold, respectively, while the Michaelis constant and turnover number for NADH remained basically unchanged.<sup>39</sup> Therefore, in AFR, A13 appears to function as an obstacle to NADH binding, even at the cost of some catalytic efficiency when using NADPH, while the S33D functions to destabilize existing interactions with NADPH.

In GFOR, the A95G mutation had no discernable effect on substrate specificity or the rate of catalysis, in contrast to the S116D mutation, which conferred dual substrate specificity. The S116D mutation caused a roughly three-fold decrease in the turnover number for NADP<sup>+</sup> while introducing glucose dehydrogenase activity to the enzyme. This new activity was made possible due to the weakening of the binding of NADP<sup>+</sup>/NADPH to the active site, as wild-type GFOR binds its cofactor so tightly that it cannot dissociate under reaction conditions. As a consequence, cofactor had to be added to the reaction mixture in order to facilitate the reaction. As with the A13G/S33D double mutant of AFR, no synergistic effect was observed for GFOR's A95G/S116D mutation. For this protein the A13G mutation does not appear to force a large enough change in secondary structure to affect substrate binding, while S116D appears to not only destabilize NADP<sup>+</sup>/NADPH binding, but also to stabilize NAD<sup>+</sup>/NADH binding.<sup>45</sup>

It is possible to explain NtdC's single cofactor specificity by assuming its A13 behaves more like GFOR's A95 than AFR's A13. If this is the case, an A13G mutation would be expected to have little effect on cofactor binding or catalysis. A D33S mutation should be far more exciting, as it may confer dual cofactor specificity or even reverse cofactor specificity, since the resulting consensus sequence would fully mimic that of an NADP<sup>+</sup>/NAHPH-dependent oxidoreductase.

# 2.3.3 NtdC recognizes monosaccharides through interactions at C-1, C-2, and C-3

Different sugar substrates were tested in order to elucidate the binding requirements for the sugar substrate. Figure 2.17 displays those sugars that showed activity when reacted with NAD<sup>+</sup> in the presence of NtdC at pH 9 in Tris·HCl. Of the sugars in Figure 16, glucose 6-phosphate (G6P), glucose, and *myo*-inositol exhibited the most activity. At 100 mM concentration and in the presence of 5 mM NAD<sup>+</sup>, G6P proved to be the most active with an initial rate of 15.3  $\mu$ mol·min<sup>-1</sup>·mg<sup>-1</sup>, followed by glucose at 7.7  $\mu$ mol·min<sup>-1</sup>·mg<sup>-1</sup> and *myo*-inositol at 4.9  $\mu$ mol·min<sup>-1</sup>·mg<sup>-1</sup>.



Figure 2.17: Sugars oxidized by NtdC.

Figure 2.18 displays those sugars that did not show activity under the above conditions. A special note must be made in the case of glucose 1-phosphate, as a 50 mM assay at pH 9.0 in 100 mM Tris·HCl buffer and in the presence of 1 mM NAD<sup>+</sup> showed roughly 14% the activity of a 50 mM G6P solution under the same conditions. It seems likely that this activity is actually just the result of a small amount of contaminating G6P in the glucose 1-phosphate stock, as a one micromolar solution of G6P, representing 0.2% contamination, would exhibit the same amount of activity. Both methyl- $\alpha$ -D-glucose and 2-deoxy-2-fluoro-*scyllo*-inositol were tested for inhibitory activity with respect to glucose, with methyl- $\alpha$ -D-glucose inhibiting the reaction, but with no effect from 2-deoxy-2-fluoro-*scyllo*-inositol. Glucose 1-phosphate did not inhibit the oxidation of G6P when an equimolar amount of both compounds was reacted with NAD<sup>+</sup> in the presence of NtdC.



Figure 2.18: Sugars not oxidized by NtdC.

A less extensive study was carried out on the reverse reaction catalyzed by NtdC. Figure 2.19 displays the structure of the three sugars tested for activity, 3-oxo-glucose, *scyllo*-inosose, and gluconolactone. As NtdC and IDH exhibited similar pH optima for the forward reaction, it seemed prudent to assume that the reverse reactions would also exhibit similar pH optima. As such, assays for the reverse reaction were performed at pH 7.0 in Tris·HCl buffer according to

Ramaley *et al.*<sup>81</sup> A decrease in the concentration of NADH could be observed when reacted with either 3-oxo-glucose or *scyllo*-inosose, but no evidence of a redox reaction could be observed when gluconolactone was used.



Substrates

Not a Substrate

Figure 2.19: Sugars tested for redox activity in the reverse reaction catalyzed by NtdC.

Inositols make good glucose analogues for studying the structural requirements of the sugar substrate. The validity of the assumption that inositol C-2 corresponds to glucose C-1, and hence all of the assignments made herein are accurate, is supported by the fact that myo-inositol, with an axial hydroxyl group at inositol C-2, is a substrate, whereas scyllo-inositol, which possesses no axial hydroxyl groups, is not a substrate. This indicates that an axial hydroxyl group is critical for sugar binding, with the inference that the position of the axial group is used to properly index the sugar ring within the active site. As glucose possesses no axial hydroxyl groups except when in the  $\alpha$ -conformation, inositol C-2 must correspond with glucose C-1. 2-Deoxy-2-fluoro-myo-inositol was used as an analogue for myo-inositol as the axial fluorine is capable of accepting a hydrogen bond, but is incapable of donating one. The fact that NtdC showed activity in the presence of 2-deoxy-2-fluoro-myo-inositol implies that a hydrogen bond is accepted by the axial hydroxyl group of *myo*-inositol and  $\alpha$ -D-glucose, and this hydrogen bond is sufficient for the proper binding and orientation of the sugar substrate within the active site. The fact that 2-deoxy-2-fluoro-scyllo-inositol was neither a substrate nor an inhibitor further supports this argument, as these results show that without the axial group, 2-deoxy-2-fluoro-scylloinositol could not bind to the active site of the enzyme. This cannot be the consequence of a steric clash between the axial fluorine group and a nearby residue, as 2-deoxy-2,2-difluoroinositol, which possesses both an axial and equatorial fluorine at inositol C-2, was a substrate. A surprising result was that although *scyllo*-inositol was not a substrate for the oxidation reaction, *scyllo*-inosose was a substrate for the reduction reaction. Therefore, an equatorial hydrogen bond acceptor is not necessary for the proper positioning of the sugar substrate for the reverse reaction. Gluconolactone was not a substrate, possibly due to a steric interaction between the hydroxyl group of C-6 with the protein. In this case, the sugar molecule would need to be rotated  $120^{\circ}$  in the active site so that C-5 occupied the position normally occupied by C-1.

The structural and sequence similarities of NtdA with members of subfamily  $VI_{\beta}$  of the PLP-dependent aminotransferases strongly suggests that NtdA acts on a (d)TDP-sugar. This would suggest that NtdC should also act on a (d)TDP-sugar. NtdC's requirement for an  $\alpha$ -configured sugar agrees with this theory. Unfortunately, neither TDP-glucose nor dTDP-glucose were available for testing. However, after examining NtdC's ability to accept additional functionalities at glucose C-1, it seems unlikely that (d)TDP-glucose would be a substrate. Methyl- $\alpha$ -D-glucose was shown to bind to NtdC, as it inhibited the oxidation of glucose, but it did not appear to be a substrate itself, indicating that it did not bind productively. The oxygen of the axial *O*-methyl group must still have been able to form the crucial hydrogen bond necessary for substrate recognition, but a steric interaction may have distorted the geometry of the binding. Another argument is to assume that NtdC does, in fact, bind a (d)TDP-sugar, and that the hydrophobic methyl group does not interact favorably with the residues which would normally stabilize the  $\beta$ -phosphate of the diphosphate linkage.

This second argument is countered by the observation that neither  $\alpha$ -D-glucose 1-phosphate nor  $\alpha$ -D-glucose 1,6-bisphosphate are substrates for the reaction catalyzed by NtdC.  $\alpha$ -D-Glucose 1-phosphate was tested and found not to inhibit the oxidation of glucose as catalyzed by NtdC, indicating that  $\alpha$ -D-Glucose 1-phosphate, and by proxy,  $\alpha$ -D-glucose 1,6-bisphosphate, did not bind, even non-productively. As both these compounds mimic the structure of the diphosphate linkage in a (d)TDP-sugar while possessing the same overall charge characteristics, they would be expected to bind with a greater affinity than methyl- $\alpha$ -D-glucose if an NDP-sugar were the natural substrate for the reaction. UDP-glucose, the substrate recognized by RifL, was not a substrate for NtdC, confirming that these two enzymes do not catalyze the same reaction, but this does not confirm that NtdC does not recognize a (d)TDP-sugar; although some NDP-sugar dehydrogenases are capable of catalyzing the reduction of both UDP- and TDP-sugars.<sup>86</sup> RifL, for example, strictly recognizes UDP-sugars,<sup>87</sup> indicating that the two are not necessarily equivocal. However, the inability for NtdC to accommodate small, negatively charged or neutral species at C-1, combined with the fact that it cannot utilize UDP-glucose, favours the assumption that NtdC does not act on a (d)NTP-sugar. There is a possibility that NtdC catalyzes the oxidation of a previously assembled neotrehalose, but this is unlikely given the above and the lack of a reaction in the presence of trehalose.

Neither mannose nor allose, the C-2 and C-3 epimers of glucose respectively, were substrates for the reaction catalyzed by NtdC. This was expected for allose, as the equatorial hydrogen at C-3 would be positioned incorrectly for abstraction by  $NAD^+$ , although allose was not tested as an inhibitor so the possibility that it simply cannot bind to the active site cannot be ruled out. The lack of activity when mannose was tested indicates that the configuration at C-2 is vital for proper binding of the sugar substrate. The configuration at C-4 is less critical, as galactose did act as a substrate. In addition, a pocket may exist near the glucose ring oxygen; both *myo*-inositol and D-*chiro*-inositol, with equatorial and axial hydroxyl groups at inositol C-1, respectively, are substrates, suggesting that there are no significant interactions between the glucose ring oxygen and the protein.

The ability of various inositol compounds to act as substrates suggests that glucose C-6 and its associated hydroxyl group are not critical for binding. It is possible that the hydroxyl group at inositol C-6 partially satisfies a hydrogen bonding interaction formed by the hydroxyl group at glucose C-6, explaining *myo*-inositol's roughly two thirds activity when compared to glucose, but even this is not strictly necessary, as xylose, which lacks a hydroxyl group at xylose C-5, also acted as a substrate. This lack of a mandatory binding requirement at glucose C-6 is surprising given the vastly improved specificity of G6P over glucose (discussed later), but is in agreement with the abilities of IDH and GFOR to accept xylose and AFR to accept xylosone as substrates.<sup>47, 88, 89</sup> Therefore, the ability to accept some variation at this position may be a common feature of the GFOR/IDH/MocA-C superfamily of oxidoreductases. However,

although NtdC readily accepts a small, negatively charged phosphate group at glucose C-6, 4-*O*-derivatives of inositol were not substrates, indicating the presence of a smaller binding pocket than reported for IDH.<sup>47</sup> Testing NtdC's affinity for different 6-*O*-derivatives of glucose will better define the size and shape of this pocket.

Overall, the configurations of glucose C-1, C-2, and C-3 are the most vital for binding, indicating that, as a minimum, the binding substrate should have a small, hydrogen bond accepting group at C-1 and equatorial hydroxyl groups at C-2 and C-3. Whether these groups need be hydrogen bond acceptors or donors can be determined by synthesizing the corresponding deoxyfluorosugars and testing them for substrate/inhibitory activity. Determining NtdC's affinity for substrates varying at glucose C-4, C-5, and at the position occupied by the ring oxygen will provide a better understanding of how crucial each of these positions is to sugar recognition.

### 2.3.4 NtdC is a G6P dehydrogenase

Given the two fold increase in activity seen when moving from glucose to G6P, the first metabolite of glucose in glycolysis and the major form of glucose imported into the *B. subtilis* cells due to the glucose-PTS uptake pathway,<sup>8</sup> it seemed prudent to further investigate G6P as a substrate. As G6P contains an ionic functional group that may affect binding at different pH, a new pH-rate profile needed to be constructed. Figure 2.20 shows the pH rate profile for G6P; the pH optimum has shifted from 9.0 with glucose to 9.5. Due to the sudden drop in activity past this point, further kinetic analysis utilizing G6P was performed at pH 9.2, where greater than 95% of activity was retained. The addition of salts, metal ions, or reducing agents were not investigated, as no discernable effect was observed when using glucose, and GFOR/IDH/MocA family members are not known to bind metal cofactors.



Figure 2.20: NtdC activity using G6P and NAD<sup>+</sup> at varying pH. Conditions: 2 mM glucose, 1 mM NAD<sup>+</sup>, 25 °C, ( $\blacklozenge$ ) 100 mM Tris·HCl, ( $\blacksquare$ ) 100 mM AMP·HCl.

The reactions using glucose and G6P at their respective pH optima were compared in more detail. For the reaction using glucose, initial rates were measured at a constant concentration of 1 mM NAD<sup>+</sup> and varying concentrations of glucose at the pH optimum for the reaction. Leonora<sup>90</sup> was used for the fitting of the initial rates to the Michaelis-Menten model describing a single substrate reaction (Equation 8) (Figure 2.21).<sup>91</sup> Figure 2.22 shows the Hill plot<sup>92</sup> for the same data. Although the slope of the trend line is 0.94, a Hill exponent of h = 1was used for determination of the kinetic constants. This is because the ability to accurately gage  $V_{max}^{app}$  (which determines the mean when estimating the Hill coefficient) was impacted do to an inability to achieve saturating conditions of glucose. For the reaction using G6P, initial rates were measured using the same conditions as for glucose, only at the pH optimum for the oxidation of G6P. Figure 2.23 shows a Lineweaver-Burk representation<sup>93</sup> of this data, along with a trend line representing the best fit of the Michaelis-Menten model to this data when a Hill exponent of h = 1 was used (Equation 9). The data deviate from the model, as the data possess a negative curvature which is indicative of negative cooperativity. A Hill plot of this data (Figure 2.24) more clearly defines this trend, as the slope yields a Hill coefficient of h = 0.7. Figure 2.25 shows the same Lineweaver-Burk representation of the data as above, only using a trend line generated from the Michaelis-Menten equation using a Hill coefficient of h = 0.7 (Equation 9). This model agrees with the recorded data very well. The kinetic constants for the reactions are summarized in Table 2.1. Glucose and G6P show similar turnover numbers suggesting that either both substrates align optimally in the active site for efficient catalysis or that the chemistry of the reaction is not rate determining. However, given NtdC's vastly improved specificity for G6P over D-glucose, G6P was tentatively assigned as the natural sugar substrate for NtdC.



Figure 2.21: Double reciprocal plot of the initial rate of the reaction catalyzed by NtdC versus the concentration of glucose. Conditions: 1 mM NAD<sup>+</sup>, 100 mM Tris·HCl, pH 9.0, 25 °C. The trend line represents the results expected for a single substrate reaction which obeys Michaelis-Menten kinetics where  $V_{max}^{app} = 5.2 \ \mu mol \cdot min^{-1} \cdot mg^{-1}$  and  $K_m^{app} = 360 \ mM$ .



Figure 2.22: Hill plot for the reaction catalyzed by NtdC at varying concentrations of glucose. Conditions: 1 mM NAD<sup>+</sup>, 100 mM Tris·HCl, pH 9.0, 25 °C.  $V_{max}^{app} = 5.2 \ \mu mol \cdot min^{-1} \cdot mg^{-1}$ . Slope = 0.94. The trend line represents a linear, least-squares best-fit of the data.



Figure 2.23: Double reciprocal plot of the initial rate of the reaction catalyzed by NtdC versus the concentration of G6P with a trend line generated using h = 1. Conditions: 1 mM NAD<sup>+</sup>, 100 mM AMP·HCl, pH 9.2, 25 °C. The trend line represents the results expected for a single substrate reaction which obeys Michaelis-Menten kinetics where  $V_{max}^{app} = 5.2 \,\mu\text{mol·min}^{-1} \,\text{mg}^{-1}$  and  $S_{0.5}^{app} = 0.146 \,\text{mM}$ .



Figure 2.24: Hill plot for the reaction catalyzed by NtdC at varying concentrations of G6P.

Conditions: 1 mM NAD<sup>+</sup>, 100 mM AMP·HCl, pH 9.2, 25 °C.  $V_{max}^{app} = 5.8 \ \mu mol \cdot min^{-1} \cdot mg^{-1}$ . Slope = 0.7. The trend line represents a linear, least-squares best-fit of the data.



Figure 2.25: Double reciprocal plot of the initial rate of the reaction catalyzed by NtdC versus the concentration of G6P with a trend line generated using h = 0.7. Conditions: 1 mM NAD<sup>+</sup>, 100 mM AMP·HCl, pH 9.2, 25 °C. The trend line represents the results expected for a single substrate reaction which obeys Michaelis-Menten kinetics and displays cooperativity where  $V_{max}^{app} = 5.8 \ \mu mol \cdot min^{-1} \cdot mg^{-1}$ ,  $S_{0.5}^{app} = 146 \ \mu M$ , and modified using a Hill coefficient of h = 0.7.

Kinetic constant	Substrate			
	Glucose	G6P		
V <sub>max</sub> <sup>app</sup>	$5.2 \pm 0.2 \ \mu mol \cdot min^{-1} \cdot mg^{-1}$	5.8 ± 0.1 μmol·min <sup>-1</sup> ·mg <sup>-1</sup>		
$K_m^{app}$ or $S_{0.5}^{app}$	$K_{m}^{app} = 360 \pm 30 \text{ mM}$	$S_{0.5}^{app} = 0.146 \pm 0.009 \text{ mM}$		
h	0.93 ± 0.02	0.72 ± 0.06		
k <sub>cat</sub> <sup>app</sup>	$3.6 \pm 0.2 \text{ s}^{-1}$	$4.0 \pm 0.1 \text{ s}^{-1}$		
Specificity Constant	$10 \pm 1 \text{ s}^{-1} \cdot \text{M}^{-1}$	$(2.7 \pm 0.2) \times 10^4 \text{ s}^{-1} \cdot \text{M}^{-1}$		

**Table 2.1: Apparent kinetic constants for the reactions catalyzed by NtdC for glucose and G6P a 1 mM NAD<sup>+</sup>.** Conditions for Glucose: 100 mM Tris·HCl, pH 9.0, 25 °C. Conditions for G6P: 100 mM AMP·HCl, pH 9.2, 25 °C.

## 2.3.4.1 The true kinetic constants for NtdC

In order to determine the true kinetic constants for the oxidation of G6P by NAD<sup>+</sup> as catalyzed by NtdC, the initial rates in the absence of products or inhibitors were measured at varying concentrations of both NAD<sup>+</sup> and G6P. Figure 2.26 and Figure 2.27 depict Hill plots for NAD<sup>+</sup> and G6P, respectively. As above, the binding of NAD<sup>+</sup> shows no cooperativity with an average Hill coefficient of  $h = 0.96 \pm 0.09$ . Likewise, the binding of G6P shows the same negative cooperativity as above with an average Hill coefficient of  $h = 0.69 \pm 0.04$  across all concentrations of NAD<sup>+</sup>. Fitting the Michaelis-Menten model for a ternary complex mechanism<sup>91</sup> with NAD<sup>+</sup> as the first substrate and G6P as the negatively cooperative second substrate (Equation 11) resulted in the kinetic constants summarized in Table 2.2. Hill coefficients of  $h_{(NAD)} = 1$  and  $h_{(G6P)} = 0.7$  were used. Figure 2.28 and Figure 2.29 display the Lineweaver-Burk plots when the reciprocal of the concentrations of NAD<sup>+</sup> and G6P are plotted along the abscissa, respectively.



Figure 2.26: Hill plot for the reaction catalyzed by NtdC at varying concentrations of NAD<sup>+</sup> and fixed concentrations of G6P. Conditions: 100 mM AMP·HCl, pH 9.2, 25 °C, (•) 2.0 mM G6P (slope = 0.87), (•) 0.20 mM G6P (slope = 0.96), (•) 0.10 mM G6P (slope = 1.00), (•) 0.060 mM G6P (slope = 1.02), (+) 0.030 mM G6P (slope = 0.94). V<sub>max</sub> = 5.7 µmol·min<sup>-1</sup>·mg<sup>-1</sup>.  $S_{0.5}^{app} = 13 \pm 1 \mu$ M. The trend lines represent linear, least-squares best-fits of the data.

Table 2.2: Kinetic constants for the oxidation of G6P by NAD<sup>+</sup> as catalyzed by NtdC in the absence of inhibitors or products. Conditions: 100 mM AMP·HCl, pH 9.2, 25 °C,  $h_{(NAD)} = 1$ ,  $h_{(G6P)} = 0.7$ .

V <sub>max</sub>	5.7 ± 0.1 $\mu$ mol·min <sup>-1</sup> ·mg <sup>-1</sup>
K <sub>m(NAD)</sub>	9 ± 3 μM
S <sub>0.5(G6P)</sub>	13 ± 1 μM
K <sub>i(NAD)</sub>	370 ± 20 μM
h <sub>(NAD)</sub>	0.96 ± 0.09
h <sub>(G6P)</sub>	0.69 ± 0.04
k <sub>cat</sub>	3.9 ± 0.1 s <sup>-1</sup>
k <sub>cat</sub> /K <sub>m(NAD)</sub>	(4 ± 1) x 10 <sup>5</sup> s <sup>-1</sup> ⋅M <sup>-1</sup>
$k_{cat}/S_{0.5(G6P)}$	$(3.0 \pm 0.2) \times 10^5 \text{ s}^{-1} \cdot \text{M}^{-1}$



Figure 2.27: Hill plot for the reaction catalyzed by NtdC at varying concentrations of G6P and fixed concentrations of NAD<sup>+</sup>. Conditions: 100 mM AMP·HCl, pH 9.2, 25 °C, ( $\diamond$ ) 1.0 mM NAD<sup>+</sup> (slope = 0.70), ( $\blacksquare$ ) 0.20 mM NAD<sup>+</sup> (slope = 0.69), ( $\blacktriangle$ ) 0.10 mM NAD<sup>+</sup> (slope = 0.63), ( $\bullet$ ) 0.060 mM NAD<sup>+</sup> (slope = 0.71), (+) 0.030 mM NAD<sup>+</sup> (slope = 0.65). V<sub>max</sub> = 5.7 µmol·min<sup>-1</sup>·mg<sup>-1</sup>. K<sub>m(NAD)</sub> = 9 ± 3 µM. The trend lines represent linear, least-squares best-fits of the data.



Figure 2.28: Double reciprocal plot of initial rate of the reaction catalyzed by NtdC versus the concentration of NAD<sup>+</sup> at fixed concentrations of G6P. Conditions: 100 mM AMP·HCl, pH 9.2, 25 °C, ( $\blacklozenge$ ) 2000  $\mu$ M G6P, ( $\blacksquare$ ) 200  $\mu$ M G6P, ( $\blacktriangle$ ) 100  $\mu$ M G6P, ( $\blacklozenge$ ) 60  $\mu$ M G6P, (+) 30  $\mu$ M G6P. The trend lines represent the results expected for a two substrate reaction which obeys a ternary-complex mechanism and displays cooperativity for the second substrate where V<sub>max</sub> = 5.7  $\mu$ mol·min<sup>-1</sup>·mg<sup>-1</sup>, S<sub>0.5(G6P)</sub><sup>app</sup> = 13 ± 1  $\mu$ M, K<sub>m(NAD)</sub> = 9 ± 3  $\mu$ M, K<sub>i(NAD)</sub> = 370 ± 20  $\mu$ M, and h<sub>(G6P)</sub> = 0.7.



Figure 2.29: Double reciprocal plot of initial rate of the reaction catalyzed by NtdC versus the concentration of G6P at fixed concentrations of NAD<sup>+</sup>. Conditions: 100 mM AMP·HCl, pH 9.2, 25 °C, (•) 1000  $\mu$ M NAD<sup>+</sup>, (•) 200  $\mu$ M NAD<sup>+</sup>, (•) 100  $\mu$ M NAD<sup>+</sup>, (•) 60  $\mu$ M NAD<sup>+</sup>, (+) 30  $\mu$ M NAD<sup>+</sup>. The trend lines represent the results expected for a two substrate reaction which obeys a ternary-complex mechanism and displays cooperativity for the second substrate where V<sub>max</sub> = 5.7  $\mu$ mol·min<sup>-1</sup>·mg<sup>-1</sup>,  $S_{0.5(G6P)}^{app} = 13 \pm 1 \,\mu$ M, K<sub>m(NAD)</sub> = 9 ± 3  $\mu$ M, K<sub>i(NAD)</sub> = 370 ± 20  $\mu$ M, and  $h_{(G6P)} = 0.7$ .

The initial activity assays performed using 100 mM concentrations of G6P and glucose found an initial rate for the oxidation of G6P that was roughly twice that of the oxidation of glucose, but kinetic analysis of the two reactions at their respective pH optima showed that they possessed similar apparent rates. The earlier result is explained by the difference in concentration of substrates required to reach 50% of the maximal velocity, given by  $S_{0.5}^{app}$  for G6P and  $K_m^{app}$  for glucose. G6P, has a sub-millimolar  $S_{0.5}^{app}$ , and as such, was at a saturating concentration in the initial test. Glucose, however, has a  $K_m^{app}$  of  $360 \pm 30$  mM, indicating that glucose was well below saturating conditions in the initial test. It should be noted that the error in estimating  $V_{max}^{app}$  and  $K_m^{app}$  for glucose should be higher than that calculated statistically by the fitting program, as saturating conditions for the reaction could not be reached. The highest concentration tested was merely 0.5 M, which is far less than the approximately 3.6 M required in order to reach saturating conditions of  $10K_m$ . Although this concentration is physically possible, as the solubility of glucose in water at 25 °C is 5 M, at higher glucose concentrations the physical properties of the reaction solution change, disturbing the reaction. Nevertheless, even a large error in estimating  $V_{max}^{app}$  and  $K_m^{app}$  for glucose is of little consequence, as the apparent specificity constant for G6P is 2,700-fold greater than that for glucose.

As no other sugars were found to be as specific as G6P, G6P was assumed to be the natural substrate for NtdC. This assumption is further supported by analyzing the kinetic constants reported for other members of the GFOR/IDH/MocA-C superfamily of oxidoreductases (Table 2.3). Although the turnover number for the oxidation of G6P when catalyzed by NtdC is two orders of magnitude lower than that for the natural substrates of IDH and AFR, NtdC's specificity constant for G6P is of the same order of magnitude as IDH and AFR's for their natural substrates. This is due to the remarkably low  $S_{0.5}$  value for G6P. This value is comparable to the Michaelis constant for NAD<sup>+</sup>, which is surprising, as members of the GFOR/IDH/MocA-C superfamily of oxidoreductases tend to have a far greater affinity for their adenine dinucleotide cofactor than for their sugar substrate.<sup>39, 45, 81</sup> These comparisons validate the assumption that G6P is the natural substrate for NtdC.

ID	Н	K <sub>m(mvo-Inositol)</sub>	= 18 ± 2 mM	k <sub>cat</sub>	$= 310 \pm 40 \text{ s}^{-1}$	k <sub>cat</sub> /K <sub>m</sub>	$= (1.7 \pm 0.3) \times 10^5 \text{ s}^{-1} \cdot \text{M}^{-1}$
AF	FR*	app K <sub>m(1,5-Anhydrofructose)</sub>	= 6.4 ± 0.01 mM	k <sub>cat</sub> app	= 145 ± 13 s <sup>-1</sup>	k <sub>cat</sub> <sup>app</sup> /K <sub>m</sub> <sup>app</sup>	$= (2.3 \pm 0.2) \times 10^5 \text{ s}^{-1} \cdot \text{M}^{-1}$
Nt	dC	S <sub>0.5(G6P)</sub>	= 13 ± 1 μM	k <sub>cat</sub>	$= 3.9 \pm 0.1 \text{ s}^{-1}$	$k_{cat}/S_{0.5}$	$= (3.0 \pm 0.2) \times 10^5 \text{ s}^{-1} \cdot \text{M}^{-1}$

 Table 2.3: The Kinetic constants of IDH,<sup>47</sup> AFR,<sup>39</sup> and NtdC, members of the GFOR/IDH/MocA-C superfamily of oxidoreductases. \*The values for AFR are reported for the reduction of 1,5-anhydro-D-fructose as opposed to the oxidation of 1,5 

anhydro-D-mannitol. These values are also reported as apparent kinetic constants using a fixed concentration of NADPH.

#### 2.3.5 G6P binds to NtdC with negative cooperativity

The decision to treat G6P binding as a negatively cooperative event was not made lightly, but a negatively cooperative event best describes the patterns observed for the oxidation of G6P. Most NAD(P)<sup>+</sup>/NAD(P)H-dependent oxidoreductases follow a sequentially-ordered mechanism where the oxidized nicotinamide adenine dinucleotide cofactor binds first, followed by the reduced substrate. Following the redox reaction, the oxidized substrate is released, followed by the reduced cofactor (Figure 2.30).<sup>91</sup> It is theoretically possible for G6P to function as a substrate inhibitor; the hydride group transferred to NADH (substrate Q) is small and G6P (substrate B) may be able to bind to the EQ complex to form the non-productive EQB complex. As such, the Michaelis-Menten model describing a ternary complex mechanism, including the term for substrate inhibition by G6P (Equation 12), was fitted to the data presented in Figure 2.29. However, the resulting fit was poor, as Leonora<sup>90</sup> returned a negative substrate inhibition constant for G6P.

As the negative cooperativity was observed across all the experiments using G6P and wild-type NtdC, it can be assured that it is not an artefact of any particular experiment. Negative cooperativity was also observed for different protein preparations, lessening the probability that the protein used was a mixture of fully active and partially damaged or otherwise impaired proteins, which would mimic negative cooperativity when used in kinetics experiments.<sup>92</sup> As the negative cooperativity was only observed for G6P, it would have to be assumed that any damage was solely localized to the C-terminal domain of the enzyme. Enzymes can display different levels of cooperativity with different substrates, such as demonstrated by GFOR binding glucose non-cooperatively while binding xylose with a positive cooperativity of h = 2,<sup>94</sup> which may explain why glucose does not appear to bind cooperatively; interactions with the 6-*O*-phosphate group of G6P may induce conformational changes that are affecting related active sites' affinities for G6P. It is interesting that the H180N mutant seems to abolish this negative cooperativity. Clearly, the binding of G6P is a complicated event, the understanding of which would be greatly improved if a crystal structure of the holo- and apoenzymes were available, but the current evidence all points to a negatively cooperative event.



Figure 2.30: The kinetic mechanism of a sequentially-ordered, bireactant system.

### 2.3.6 NtdC likely obeys a ternary complex mechanism

Having determined that G6P was the natural sugar substrate for NtdC, it was decided that the kinetic mechanism for the reaction should be investigated. First, an experiment was performed in which the concentration of NAD<sup>+</sup> was varied in the presence of different concentrations of its reduced form, NADH. A Hill plot demonstrates that the binding of NAD<sup>+</sup> is not cooperative, as the average slope for the reaction rounds to h = 1 (Figure 2.31). The Michaelis-Menten model including the term for competitive inhibition (Equation 14)<sup>92</sup> matches well with the Lineweaver-Burk plot of the data (Figure 2.32). Inclusion of the term for uncompetitive inhibition (Equation 16) did not improve the quality of the fit. The same experiment was performed again, only varying the concentration of G6P instead of NAD<sup>+</sup>. The Hill plot of the data reveals an average slope of roughly h = 0.7, agreeing with the data from above (Figure 2.33). The Michaelis-Menten model including the term for competitive inhibition agrees with the plotted data when the negative cooperativity of G6P is taken into account (Equation 17) (Figure 2.34). As above, inclusion of the term for uncompetitive inhibition (Equation 19) did not improve the quality of the fit. A summary of the apparent kinetic constants extracted from these two experiments is presented in Table 2.4.



Figure 2.31: Hill plot for the reaction catalyzed by NtdC at varying concentrations of NAD<sup>+</sup> and fixed concentrations of NADH. Conditions: 2 mM G6P, 100 mM AMP·HCl, pH 9.2, 25 °C, ( $\diamond$ ) 0 µM NADH (slope = 0.95), ( $\blacksquare$ ) 15 µM NADH (slope = 1.00), ( $\blacktriangle$ ) 30 µM NADH (slope = 0.89), ( $\bullet$ ) 50 µM NADH (slope = 0.97). V<sub>max</sub><sup>app</sup> = 5.36 µmol·min<sup>-1</sup>·mg<sup>-1</sup>.



Figure 2.32: Double reciprocal plot of initial rate of the reaction catalyzed by NtdC versus the concentration of NAD<sup>+</sup> at fixed concentrations of NADH. Conditions: 2 mM G6P, 100 mM AMP·HCl, pH 9.2, 25 °C, ( $\blacklozenge$ ) 0  $\mu$ M NADH, ( $\blacksquare$ ) 15  $\mu$ M NADH, ( $\blacktriangle$ ) 30  $\mu$ M NADH, ( $\circlearrowright$ ) 50  $\mu$ M NADH. Trend line generated using the Michaelis-Menten equation describing competitive inhibition with V<sub>max</sub><sup>app</sup> = 5.36  $\mu$ mol·min<sup>-1</sup>·mg<sup>-1</sup>, K<sub>m</sub><sup>app</sup> = 24  $\mu$ M, K<sub>ic</sub> = 5.9  $\mu$ M, and h = 0.1.



Figure 2.33: Hill plot for the reaction catalyzed by NtdC at varying concentrations of G6P and fixed concentrations of NADH. Conditions: 200  $\mu$ M NAD<sup>+</sup>, 100 mM AMP·HCl, pH 9.2, 25 °C, ( $\blacklozenge$ ) 0  $\mu$ M NADH (slope = 0.72), ( $\blacksquare$ ) 4  $\mu$ M NADH (slope = 0.72), ( $\blacktriangle$ ) 8  $\mu$ M NADH (slope = 0.72), ( $\circlearrowright$ ) 16  $\mu$ M NADH (slope = 0.72). V<sub>max</sub><sup>app</sup> = 6.0  $\mu$ mol·min<sup>-1</sup>·mg<sup>-1</sup>.



Figure 2.34: Double reciprocal plot of the initial rate of the reaction catalyzed by NtdC versus the concentration of G6P at fixed concentrations of NADH. Conditions: 200  $\mu$ M NAD<sup>+</sup>, 100 mM AMP·HCl, pH 9.2, 25 °C, ( $\blacklozenge$ ) 0  $\mu$ M NADH, ( $\blacksquare$ ) 4  $\mu$ M NADH, ( $\blacktriangle$ ) 8  $\mu$ M NADH, ( $\blacklozenge$ ) 16  $\mu$ M NADH. V<sub>max</sub><sup>app</sup> = 6.0  $\mu$ mol·min<sup>-1</sup>·mg<sup>-1</sup>. Trend line generated using the Michaelis-Menten equation describing competitive inhibition with V<sub>max</sub><sup>app</sup> = 6.0  $\mu$ mol·min<sup>-1</sup>·mg<sup>-1</sup>, S<sub>0.5</sub><sup>app</sup> = 47  $\mu$ M, K<sub>ic</sub> = 6.7  $\mu$ M, and h = 0.7.

Table 2.4: Apparent kinetic constants for the NADH-dependent inhibition of the oxidation of G6P by NAD<sup>+</sup> as catalyzed by NtdC. Conditions: 100 mM AMP·HCl, pH 9.2, 25 °C. For NAD<sup>+</sup>, the concentration of G6P was fixed at 2 mM. For G6P, the concentration of NAD<sup>+</sup> was fixed at 0.2 mM. Hill factors of  $h_{(NAD)} = 1$  and  $h_{(G6P)} = 0.7$  were used.

Kinetic constant	Varied Substrate			
	NAD⁺	G6P		
V <sub>max</sub> <sup>app</sup>	5.36 ± 0.06 $\mu$ mol·min <sup>-1</sup> ·mg <sup>-1</sup>	$6.0 \pm 0.2 \mu mol \cdot min^{-1} \cdot mg^{-1}$		
$K_m^{app}$ or $S_{0.5}^{app}$	$K_m^{app} = 24 \pm 1 \mu M$	$S_{0.5}^{app} = 47 \pm 4 \mu\text{M}$		
h	0.96 ± 0.02	0.71 ± 0.02		
Kic(NADH)	5.9 ± 0.2 μM	6.7 ± 0.4 μM		
k <sub>cat</sub> <sup>app</sup>	$3.70 \pm 0.04 \text{ s}^{-1}$	4.1 ± 0.1 s <sup>-1</sup>		

Determining the kinetic mechanism followed by NtdC was somewhat inconclusive. Cook and Cleland's book on enzyme kinetics gives a thorough listing of the product inhibition patterns possible for all of the bireactant mechanisms.<sup>95</sup> Based on the fact that NADH was found to be a competitive inhibitor with respect to both NAD<sup>+</sup> and G6P, it would appear that NtdC must follow an equilibrium ordered mechanism. However, the Michaelis-Menten model describing an equilibrium ordered mechanism (Equation 13) did not fit the data in Figure 2.29 as well as that describing a ternary complex mechanism, either with or without use of the Hill coefficient.

Ternary complex mechanisms should produce a pattern whereby NADH binds competitively with respect to NAD<sup>+</sup> and non-competitively with respect to G6P. This is purely a mathematical treatment and overlooks one possibility that could explain why a non-competitive inhibition pattern was not observed for NADH with respect to G6P. It is possible that the barrier to G6P binding with the enzyme:NADH complex is high, or that the EQB complex cannot form at all, leading to a  $K_{iu}$  for NADH with respect to G6P much higher than the maximum concentration of NADH tested. Hence, an uncompetitive term could not be detected. Cornish-Bowden acknowledges this possibility in his discussion of the ternary-complex mechanism, but only states the inhibition pattern expected for what would be the reduction of 3-oxo-D-glucose 6-phosphate by NADPH and not the oxidation of G6P by NAD<sup>+</sup>.<sup>91</sup> Unfortunately, a practical consideration prevents the use of much higher concentrations of NADH, as the level of activity quickly falls below the sensitivity of the spectrophotometer. This explanation agrees well with the fact that no substrate inhibition was observed for G6P, as G6P would have to inhibit the reaction through the formation of the EQB dead-end complex.

Validating the assumption that NtdC obeys a ternary-complex mechanism and differentiating between a steady-state ordered and a Theorell-Chance mechanism will require testing all of the substrates and products for their respective inhibition patterns for both the forward and reverse reactions catalyzed by NtdC. However, although NAD<sup>+</sup>, NADH, and G6P are readily available, a reliable method of synthesizing and purifying 3-oxo-D-glucose 6-phosphate has yet to be developed by our lab. Until 3-oxo-D-glucose 6-phosphate can be obtained, there will always be the possibility that NtdC obeys a different kinetic mechanism.

However, as the majority of  $NAD(P)^+/NAD(P)H$ -dependent oxidoreductases obey a ternary complex mechanism,<sup>91</sup> it is reasonable to assume that NtdC does as well.

# 2.4 Mutagenesis of *ntdC*

Insight into the chemical mechanism of NtdC can be gained through mutagenesis experiments. Sequence alignments with other sugar dehydrogenases indicated the presence of a conserved pair of residues, H180 and D176, which were predicted to form a catalytic dyad. The aspartate residue was replaced with an asparagine residue using the QuikChange Site-Directed Mutagenesis Kit and the primers *ntdC*-D172N-Forward and *ntdC*-D176N-Reverse. In addition to encoding the D176N mutation, these primers encoded a silent guanidine to adenine mutation at position 526 of the *ntdC* gene, creating a *Bam*HI restriction site. Figure 2.35 displays *Eco*RV/*Bam*HI restriction maps for wild type pET28b-*ntdC* and mutant pET28b-*ntdC*-D172N. Wild type pET28b-*ntdC* should be cut twice, leaving 2392 bp and 3994 bp fragments, while mutant plasmids should be cut three times, leaving three fragments, one of 540 bp, one of 1852 bp, and one of 3994 bp. Figure 2.36 shows a picture of an ethydium bromide stained agarose gel depicting the successful mutation of the *ntdC* gene; the 540 bp fragment was very faint and could not be captured by the camera. DNA sequencing proved that the mutant *ntdC* gene was free of any PCR errors. Mutants of the histidine residue, H180A, H180Y, and H180N, were constructed by Grey Wilkinson.



Figure 2.35: *EcoRV/Bam*HI restriction map of pET28b-*ntdC* and pET28b-*ntdC*-D176N. Figure generated using NEBcutter V2.<sup>79</sup>



**Figure 2.36: Ethydium bromide stained 1% agarose gel of an** *Eco***RV**/*Bam***HI double-digest of pET28b**-*ntdC*-**D176N.** Lane 1: Invitrogen 1 kbp Plus DNA ladder. Lane 2: pET28b-*ntdC*. Lane 3: pET28b-*ntdC*-D176N.

#### 2.4.1 Disruption of NtdC's proposed catalytic dyad has a large, negative effect

Mutant enzymes were purified in the same manner as wild type NtdC and assayed for activity at 1 mM NAD<sup>+</sup> and 2 mM G6P in 100 mM pH 9.2 AMP·HCl buffer. Initially, the D176N mutant assayed with an initial rate 1% that of wild type NtdC. However, two subsequent batches of enzyme showed no activity under the same conditions. It was assumed that the first batch of enzyme was contaminated with wild type NtdC. With respect to the H180 mutants, H180Y displayed no activity, while H180A displayed roughly 0.1% the activity of wild type NtdC, close to the detection limit of the machine. H180N assayed surprisingly active at roughly 0.5% the activity of wild type NtdC. Figure 2.37 shows the pH-rate profile for this mutant. Overall, Figure 2.37 shows a sharper profile than Figure 2.20 with a slightly lower pH optimum of pH 9.3. As with wild type enzyme, the rate of decrease at higher pH is greater than at lower pH, so further assays were performed at pH 9.2. First, the change in  $S_{0.5}$  for G6P, which relates to the Michaelis constant, as both relate to the concentration of the substrate at one half  $V_{max}$ , was investigated. Figure 2.38 displays the Hill plot for the binding of G6P to the H180N mutant of NtdC in the presence of 4 mM NAD<sup>+</sup>. For this mutant, G6P shows no cooperativity. Figure 2.39 shows the Lineweaver-Burk replot of this data using the apparent kinetic constants summarized in Table 2.5. Figure 2.40 displays the Hill plot for the binding of NAD<sup>+</sup> to the H180N mutant of NtdC in the presence of 150 mM G6P. As expected, no cooperativity is observed. Figure 2.41 shows the Lineweaver-Burk replot of this data using the apparent kinetic constants summarized in Table 2.5.



Figure 2.37: Activity of the H180N mutant of NtdC at varying pH using G6P and NAD<sup>+</sup>. Conditions: 10 mM G6P, 5 mM NAD<sup>+</sup>, 25 °C, (♦) 100 mM Tris·HCl, (■) 100 mM AMP·HCl.



Figure 2.38: Hill plot for the reaction catalyzed by the H180N mutant of NtdC at varying concentrations of G6P. Conditions: 4 mM NAD<sup>+</sup>,

100 mM AMP·HCl, pH 9.2, 25 °C.  $V_{max}^{app} = 0.30$  µmol·min<sup>-1</sup>·mg<sup>-1</sup>. Slope = 1.01. The trend line represents a linear, least-squares best-fit of the data.



Figure 2.39: Double reciprocal plot of the initial rate of the reaction catalyzed by the H180N mutant of NtdC versus the concentration of G6P. Conditions: 4 mM NAD<sup>+</sup>, 100 mM AMP·HCl, pH 9.2, 25 °C. The trend line represents the results expected for a single substrate reaction which obeys Michaelis-Menten kinetics where  $V_{max}^{app} = 0.30 \ \mu mol \cdot min^{-1} \cdot mg^{-1}$  and  $K_m^{app} = 12.8 \ mM$ .



Figure 2.40: Hill plot for the reaction catalyzed by the H180N mutant of NtdC at varying concentrations of NAD<sup>+</sup>. Conditions: 150 mM G6P, 100 mM AMP·HCl, pH 9.2, 25 °C.  $V_{max}^{app} = 0.292$ µmol·min<sup>-1</sup>·mg<sup>-1</sup>. Slope = 1.04. The trend line represents a linear, least-squares best-fit of the data.



Figure 2.41: Double reciprocal plot of the initial rate of the reaction catalyzed by the H180N mutant of NtdC versus the concentration of NAD<sup>+</sup>. Conditions: 150 mM G6P, 100 mM AMP·HCl, pH 9.2, 25 °C. The trend line represents the results expected for a single substrate reaction which obeys Michaelis-Menten kinetics where  $V_{max}^{app} = 0.292$  µmol·min<sup>-1</sup>·mg<sup>-1</sup> and  $K_m^{app} = 0.39$  mM.

Table 2.5: Apparent kinetic constants for the oxidation of G6P by NAD<sup>+</sup> as catalyzed by the H180N mutant of NtdC. Conditions: 100 mM AMP·HCl, pH 9.2, 25 °C. For NAD<sup>+</sup>, the concentration of G6P was fixed at 150 mM. For G6P, the concentration of NAD<sup>+</sup> was fixed at 4 mM. Hill factors of  $h_{(NAD)} = 1$  and  $h_{(G6P)} = 1$  were used.

Kinetic constant	Varied Substrate			
	NAD⁺	G6P		
V <sub>max</sub> <sup>app</sup>	0.292 ± 0.009 μmol·min <sup>-1</sup> ·mg <sup>-1</sup>	0.30 ± 0.01 μmol·min <sup>-1</sup> ·mg <sup>-1</sup>		
K <sub>m</sub> <sup>app</sup>	0.39 ± 0.03 mM	12.8 ± 0.7 mM		
h	1.04 ± 0.05	1.01 ± 0.02		
k <sub>cat</sub> <sup>app</sup>	$0.201 \pm 0.006 \text{ s}^{-1}$	0.20 ± 0.01 s <sup>-1</sup>		
k <sub>cat</sub> <sup>app</sup> /K <sup>app</sup>	$500 \pm 40 \text{ s}^{-1} \cdot \text{M}^{-1}$	$16 \pm 1 \text{ s}^{-1} \cdot \text{M}^{-1}$		

NtdC appears to be more sensitive to manipulation of its Asp-His, catalytic dyad than either AFR or IDH. The conservative mutation of D172 to asparagine in IDH produced an enzyme with 1/30<sup>th</sup> the activity of the wild type enzyme. Along with this decrease in activity came similar increases in the Michaelis constants for the sugar substrate and the dinucleotide (Table 2.6). The less conservative D176A mutation of AFR had a more pronounced effect on

that enzyme's activity, with roughly a 110-fold decrease in the turnover number when compared to the wild type enzyme. Surprisingly, the mutation to alanine, which does not preserve either the steric or hydrogen bonding characteristics of aspartate like asparagine does, resulted in a smaller decrease of the binding constant for the sugar substrate. As a result, despite the much larger decrease in the turnover number for AFR versus IDH, the sugar substrate specificity constant for AFR is impacted much less than that for IDH. Unfortunately, as no activity could be detected for the D176N mutant of NtdC, Michaelis constants could not be compared. However, the lack of activity alone shows that D176 is critical to catalysis in NtdC.

**Table 2.6:** The effect of mutations to the Asp-His catalytic dyads of NtdC, AFR,<sup>39</sup> and IDH.<sup>47</sup> \*The values for AFR are reported for the reduction of 1,5-anhydro-D-fructose as opposed to the oxidation of 1,5-anhydro-D-mannitol. These values are also reported as apparent kinetic constants using a fixed concentration of NADPH. ND = Not Determined. BDL = Below Detection Limit.

NtdC		AFR		IDH	
D176N		D176A		D172N	
K <sub>m(G6P)</sub> <sup>app</sup>	ND	app K <sub>m(1,5</sub> -anhydrofructose)	↑ 6-fold	K <sub>m(inositol)</sub>	↑ 15-fold
$K_{m(NADP)}^{app}$	ND	K <sub>m(NADPH)</sub> <sup>app</sup>	ND	K <sub>m(NAD)</sub>	↑ 14-fold
k <sub>cat</sub> <sup>app</sup>	BDL	k <sub>cat</sub> <sup>app</sup>	↓ 112-fold	k <sub>cat</sub>	↓ 30-fold
k <sub>cat</sub> <sup>app</sup> /K <sub>m(G6P)</sub> <sup>app</sup>	ND	$k_{cat}^{app}/K_{m(1,5-anhydrofructose)}^{app}$	↓ 870-fold	k <sub>cat</sub> /K <sub>m(inositol)</sub>	↓ 1444-fold
H180A		H180A		H176A	
K <sub>m(G6P)</sub> <sup>app</sup>	ND	app K <sub>m(1,5</sub> -anhydrofructose)	↑ 1.4-fold	K <sub>m(inositol)</sub>	↑ 27-fold
$K_{m(NADP)}^{app}$	ND	K <sub>m(NADPH)</sub> <sup>app</sup>	ND	K <sub>m(NAD)</sub>	↑ 4-fold
k <sub>cat</sub> <sup>app</sup>	↓ 1000-fold	k <sub>cat</sub> <sup>app</sup>	↓ 39-fold	k <sub>cat</sub>	↓ 250-fold
$k_{cat}^{app}/K_{m(G6P)}^{app}$	ND	$k_{cat}^{app}/K_{m(1,5-anhydrofructose)}^{app}$	↓ 54-fold	k <sub>cat</sub> /K <sub>m(inositol)</sub>	↓ 6500-fold
H180N					
K <sub>m(G6P)</sub> app	↑ 985-fold				
$K_{m(NADP)}^{app}$	↑ 43-fold				
k <sub>cat</sub> <sup>app</sup>	↓ 19-fold				
K <sub>m(G6P)</sub> <sup>app</sup> /k <sub>cat</sub> <sup>app</sup>	↓ 18,750-fold				
H180Y					
K <sub>m(G6P)</sub> app	ND				
$K_{m(NADP)}^{app}$	ND				
k <sub>cat</sub> <sup>app</sup>	BDL				
k <sub>cat</sub> <sup>app</sup> /K <sub>m(G6P)</sub> <sup>app</sup>	ND				

In the case of the histidine residue, the non-conservative mutation of H176 to alanine in IDH resulted in a 250-fold lowering of the turnover number when compared to wild type IDH. As expected for a residue that directly contacts the sugar substrate but not the dinucleotide, the Michaelis constant for the sugar substrate was more heavily affected than that for the dinucleotide. Overall, the specificity constant was decreased four orders of magnitude. The correlating mutation in AFR had a far less pronounced effect. The turnover number only decreased 39-fold. In addition, the effect on the Michaelis constant for the sugar substrate was minimal, leading to a decrease in the specificity constant of only 54-fold. It is interesting to note that mutation of the aspartate residue is actually more detrimental to catalysis in AFR than mutation of the catalytic base, histidine.

For NtdC, the H180A mutant assayed with only a trace amount of activity. The low activity of wild type NtdC makes the assaying of significantly impaired mutants difficult, so proper kinetic constants for the H180A mutant were not determined. However, the H180N mutant was measurably active. As seen with the H176A mutant of IDH, the Michaelis constant of the sugar substrate was affected more strongly than that of the dinucleotide. The magnitude is what differs, with the mutation in NtdC having a 36-times greater effect than in IDH (and 700 times greater than for AFR). Surprisingly, the rate of catalysis is affected far less than in either IDH or AFR. This is counterintuitive, as the asparagine mutation was predicted to drastically decrease the rate of catalysis, while only marginally decreasing affinity for the sugar substrate. Asparagine is able to act as both a hydrogen bond donor and acceptor like histidine, preserving any hydrogen bonding interactions that histidine may make with the sugar substrate, and is less bulky than histidine, minimizing the possibility of introducing any steric clashes, but normally lacks the ability to act as a base due to the high  $pK_a$  and low  $pK_b$  of its side chain amide group. It could be that the asparagine is coordinating a water molecule that is able to act as a base, but a crystal structure of the holoenzyme with both the cofactor and the sugar substrate bound would be necessary before this could be said for certain. The specificity constant for the H180N mutant was lowered by five orders of magnitude, a larger effect than even for the H176 mutant of IDH. This indicates that even though the H180N mutant of NtdC may still be relatively active compared to wild type NtdC, H180 of NtdC is still more sensitive to manipulation than either H180 in AFR or H176 in IDH.

As mentioned in the introduction, GFOR is theorized to possess a catalytic dyad composed of K129 and T217, corresponding to K91 and H180 in NtdC. In order to test this hypothesis, the H180Y mutant of NtdC was constructed. No activity could be detected for this mutant. Unfortunately, this does not disprove a catalytic dyad between K129 and T217 in GFOR, as the argument can be made that the increased bulk of the tyrosine side chain interferes with the binding of the substrates. A moderately active mutant, however, would have strongly supported this hypothesis.

## 2.4.2 NtdC can accommodate K227 and R227 equally well

In addition, lysine 227 was mutated to the arginine which was encoded by the ntdC sequence published as part of the *B. subtilis* genome.<sup>76</sup> Primers *ntdC*-K227R-Forward and *ntdC*-K227R-Reverse were used to effect this mutation. Unfortunately, this mutation did not alter the restriction map of the pET28b-ntdC plasmid, and no convenient site for a silent mutation that would introduce or remove a restriction site could be located. As a result, restriction digests could not be used to identify mutant plasmids. Instead, DNA sequencing was used to confirm the identity of successful pET28b-ntdC-K227R mutants. An initial rate experiment could not detect a significant change in activity between wild type NtdC and the K227R mutant, so a G6P saturation curve in the presence of 2 mM NAD<sup>+</sup> was constructed. A Hill plot of the data revealed negative cooperativity with a Hill coefficient of roughly h = 0.7 (Figure 2.42). A Lineweaver-Burk plot of the data showed a good agreement between the Michaelis-Menten model and the data, except for at the lowest concentration of G6P where the error in estimating the initial rate is greatest and where the replot magnifies that error (Figure 2.43). The apparent kinetic constants for the reaction are summarized in Table 2.7. The specificity constant for G6P is not significantly different from that presented in Table 2.1, which indicates that this mutation had a trivial affect on NtdC's catalytic activity. Therefore, the fact that the primary structure of the NtdC protein used for this study differed from that reported in the *B. subtilis* genome<sup>80</sup> was not an issue.


Figure 2.42: Hill plot for the reaction catalyzed by the K227R mutant of NtdC at varying concentrations of G6P. Conditions: 2 mM NAD<sup>+</sup>, 100 mM AMP·HCl, pH 9.2, 25 °C.  $V_{max}^{app} = 6.72$ µmol·min<sup>-1</sup>·mg<sup>-1</sup>. Slope = 0.69. The trend line represents a linear, least-squares best-fit of the data.



Figure 2.43: Double reciprocal plot of the initial rate of the reaction catalyzed by the K227R mutant of NtdC versus the concentration of G6P. Conditions: 2 mM NAD<sup>+</sup>, 100 mM AMP·HCl, pH 9.2, 25 °C. The trend line represents the results expected for a single substrate reaction which obeys Michaelis-Menten kinetics and displays cooperativity where  $V_{max}^{app} = 6.72 \,\mu mol \cdot min^{-1} \cdot mg^{-1}$ ,  $S_{0.5}^{app} = 189 \,\mu M$ , and h = 0.7.

Table 2.7: Apparent kinetic constants for the oxidation of G6P by NAD<sup>+</sup> as catalyzed by the K227R mutant of NtdC. Conditions: 2 mM NAD<sup>+</sup>, 100 mM AMP·HCl, pH 9.2, 25 °C,  $h_{(G6P)} = 0.7$ .

V <sub>max</sub> <sup>app</sup>	$6.72 \pm 0.07 \ \mu mol \cdot min^{-1} \cdot mg^{-1}$
S <sub>0.5</sub> app	189 ± 7 μM
h	0.69 ± 0.03
k <sub>cat</sub> <sup>app</sup>	$4.66 \pm 0.05 \text{ s}^{-1}$
k <sub>cat</sub> <sup>app</sup> /S <sub>0.5</sub> <sup>app</sup>	(2.5 ± 0.5) x 10 <sup>4</sup> s <sup>-1</sup> ⋅M <sup>-1</sup>

#### 2.5 Synthesis of kanosamine (7)

As depicted in Scheme 1.2, NtdB may catalyze the condensation between two kanosamine residues in the last step of NTD biosynthesis. Kanosamine was required to test this possibility. The synthesis of kanosamine has been published numerous times. The route used in this thesis is presented in Scheme 2.1. The key step was the oxidation of 1,2:5,6-di-Oisopropylidene- $\alpha$ -D-glucofuranose (1) to 1,2:5,6-di-O-isopropylidene- $\alpha$ -D-ribo-hexofuranos-3ulose (2) by pyridinium dichromate (PDC) in the presence of acetic anhydride, as first demonstrated by Garegg and Samuelsson,<sup>96</sup> followed by a stereospecific reduction at C-3 by sodium borohydride to form 1,2:5,6-di-O-isopropylidene- $\alpha$ -D-allofuranose (3).<sup>97</sup> This inversion of stereochemistry at C-3 allowed for the use of an S<sub>N</sub>2 displacement at C-3, preserving the original configuration of the glucose molecule. Tosylation of the hydroxyl group at C-3 followed by displacement with sodium azide in refluxing DMF resulted in an incomplete conversion of the starting material and a large amount of discolouration. Flash column chromatography was required to remove the colour and separate the newly created compound 5 (88%) from unreacted compound 4 (6%). An alternative method substituting trifluoromethansulfonic anhydride instead of p-toluenesulfonyl chloride has been developed, which although more difficult to perform, resulted in a near quantitative yield of compound 5 from compound **3** and does not require column chromatography. The azide group was converted to an *N*-*t*-butoxycarbonyl (*t*-Boc) protected amine through palladium catalyzed hydrogenolysis in the presence of di-t-butyl dicarbonate. Kanosamine was isolated as the hydrochloride salt (7) after acid catalyzed hydrolysis of compound 6. Although pure by NMR, the product was very hygroscopic. Some water remained, even after several days under high vacuum, and heating discoloured the product. Assuming the last step to be quantitative, kanosamine HCl (7) was isolated in 70% overall yield from the diisopropylidene of glucose (1) with the need to run only two columns.



Scheme 2.1: Synthesis of kanosamine (7). (a) PDC,  $Ac_2O$ , DCM,  $\Delta$ ; (b)  $NaBH_4$ , EtOH/H<sub>2</sub>O (v/v 1:3), 86% (2 steps); (c) TsCl, Pyridine, 94%; (d)  $NaN_3$ , DMF,  $\Delta$ , 88%; (e) H<sub>2</sub>, Pd/C, 'Boc<sub>2</sub>O, 98%; (f) 2 M HCl, quant. 70% overall yield from **1**.

#### 2.6 Synthesis of neotrehalose (15)

The synthesis of NTD and derivatives will be necessary in order to determine if NtdB is capable of catalyzing the hydrolysis of NTD (Scheme 1.2). Before synthesizing NTD, a process which first requires the five step synthesis of kanosamine (**7**) or its immediate precursor, 3-azido-3-deoxy-D-glucose, it seemed prudent to first develop a methodology using the analogous synthesis of neotrehalose (**15**) from glucose (**8**). Neotrehalose has been synthesized before, most often by the Koenigs-Knorr or Helferich methods from 2,3,4,6-tetra-*O*-acetyl- $\alpha$ -D-glucose, logucosyl bromide and 2,3,4,6-tetra-*O*-acetyl-D-glucose.

A different route was chosen in the hopes of improving this synthesis (Scheme 2.2). A thiophenol group was installed at C-1 of glucose, as anomeric thiol groups are relatively stable and are readily converted to any of the common glycosyl donors.<sup>98</sup> The one-pot synthesis of compound **9** did not scale as well as hoped, as only 34% of the product was isolated by

crystallization. Column chromatography is reported to yield much better results,<sup>99</sup> but is not as convenient for gram scale syntheses.

From here, a synthetic route to the perbenzylated derivative of neotrehalose (13) using the trichloroacetimidate of 2,3,4,6-tetra-*O*-benzyl-D-glucopyranoside (12) as the donor sugar was followed.<sup>76</sup> Methanolysis of the acetyl groups in the presence of sodium methoxide, followed by benzylation, produced 2,3,4,6-Tetra-*O*-benzyl-1-*S*-thiophenol- $\beta$ -D-glucopyranose (10) in 82% yield. Selective hydrolysis of the thiophenol group using *N*-bromosuccinimide (NBS) in wet acetone according to Damager *et al.*<sup>100</sup> resulted in 2,3,4,6-tetra-*O*-benzyl-D-glucopyranoside (11) in 91% yield after recrystallization. A portion of compound 11 was converted to the trichloroacetimidate (12) using sodium hydride and trichloroacetonitrile in 87% yield and with a 12:1  $\alpha$ : $\beta$  ratio after column chromatography. This method was later improved upon by substituting Cs<sub>2</sub>CO<sub>3</sub> for sodium hydride, yielding only the  $\alpha$ -anomer in 94% yield after passing through a short column of silica to remove colour.

Compounds **11** and **12** were coupled using silver triflate, producing several compounds visible by UV after TLC. It was hoped perbenzylated neotrehalose (**13**) would prove simpler to isolate than neotrehalose synthesized by more classical methods, but this was not the case. Isolation of pure compound **13** required two chromatographic separations by flash column, followed by two recrystallizations to remove an unidentified, non-sugar contaminant. The vast majority of compound **13** could not be isolated from the other glycosylation products, yielding only 15% of the expected product. Palladium catalyzed hydrogenolysis proceeded smoothly to generate pure neotrehalose (**15**) in quantitative yield. Overall, neotrehalose (**15**) was isolated in 3% overall yield after seven steps, three silica gel columns, and five recrystallizations.







Scheme 2.2: Synthesis of neotrehalose (15). (a) *i*) Ac<sub>2</sub>O, 30% HBr in AcOH, *ii*) Thiophenol, Bu<sub>4</sub>NHSO<sub>4</sub>, Na<sub>2</sub>CO<sub>3</sub>, H<sub>2</sub>O/DCM, 34%;<sup>99</sup> (b) *i*) NaOMe, MeOH,  $\Delta$ , *ii*) NaH, DMF, *iii*) BnBr, 82%; (c) NBS, Acetone/H<sub>2</sub>O (v/v, 9:1), 91%;<sup>100</sup> (d) Trichloroacetonitrile, NaH, DCM, 87%;<sup>100</sup> (e) Silver triflate, DCM, in the dark, 15% perbenzylated neotrehalose isolated yield;<sup>76</sup> (f) 10% Pd/C, H<sub>2</sub>, MeOH, quant. 3% overall yield from glucose.

# 2.7 Synthesis of 3-[<sup>2</sup>H]-D-glucose (20)

In the absence of 3-oxo-D-glucose or 3-oxo-6-O-phospho-D-glucose, the first product of NtdC when assuming a sequentially ordered mechanism, inhibition studies alone are not enough to distinguish between a steady-state ordered mechanism and a Theorell-Chance kinetic mechanism. However, in a Theorell-Chance kinetic mechanism, product release is much slower than chemistry.<sup>101</sup> Therefore, an observable kinetic isotope effect would rule out the possibility of a Theorell-Chance mechanism. In order to study any potential kinetic isotope effects associated with the abstraction of the hydride from C-3, it was necessary to synthesize 3-[<sup>2</sup>H]-D-Glucose (20).  $3 \cdot [^{2}H]$ -D-Glucose (20) was first synthesized by Koch *et al.* in 1970.<sup>102</sup> Since then, this method has been the route of choice for the synthesis of this deuterated sugar.<sup>103, 104</sup> Koch *et* al. leveraged the well known stereospecific reduction of the diisopropylidene of 3-oxo-D-glucose (2) to generate  $3-[^{2}H]-1,2:5,6-di-O-isopropylidene-\alpha-D-allofuranose$  (16).<sup>97</sup> Tosylation of the free hydroxyl group, followed by displacement with sodium benzoate lead to compound 18. Hydrolysis of the benzoyl group resulted in compound **19** in 55% yield from compound **16**. Use of a triflic group instead of a tosyl group improved this process greatly, increasing the yield of these three steps to 87% (Scheme 2.3). The incorporation of the deuteron into the diacetone of glucose (1) was confirmed through <sup>1</sup>H NMR by observing the disappearance of a one proton signal at 4.36 ppm (corresponding to the proton attached to C-3), the reduction of the signal at 4.08 ppm (corresponding to the proton attached to C-4) from a doublet of doublets to just a doublet, and the reduction of the signal at 2.46 ppm (corresponding to the hydroxylic proton at C-3) from a doublet to a singlet.

Dowex 50X-80 (H<sup>+</sup>) was used to mildly hydrolyze the isopropylidene groups, yielding 3- $[^{2}H]$ -D-Glucose (**20**) as a lightly yellow glass in quantitative yield. However, the product retained a small amount of water which could not be removed under high vacuum. Although the NMR appeared pure, the melting point of compound **20** (25-36 °C) was considerably depressed from that of crystalline glucose (162-167 °C). The manufacturer of the crystalline glucose used as a standard reports a melting point of 150-152 °C. Additionally, the UV spectrum of a 100 mM solution of glucose, which had been prepared using sodium borohydride instead of sodium borodeuteride, showed absorbance peaks at 226 nm and 253 nm, with a broad shoulder extending past 400 nm. Crystalline glucose showed a smaller, single peak at 214 nm. Finally, preliminary

kinetics which compared prepared glucose with crystalline glucose at 100 mM concentration and in the presence of 1 mM NAD<sup>+</sup> and 100 Tris·HCl at pH 9.0 showed that prepared glucose exhibited 50% the activity of crystalline glucose in the presence of IDH and only 5% the activity of crystalline glucose when in the presence of NtdC. Some chemical undetected in the NMR must be contaminating glucose prepared in this manner.

Incomplete hydrolysis would result in a small amount of 1,2-*O*-isopropylidene- $\alpha$ -D-glucofuranose. However, this compound did not inhibit the oxidation of crystalline glucose when added to the reaction mixture. The contamination must have happened during hydrolysis, as crystalline glucose, when reacted with acid dowex in the same manner as above, yielded a very lightly yellow solid with a depressed melting point of 43-56 °C, while crystalline glucose that had merely been dissolved in water yielded a crystalline solid with a melting point of 147-154 °C. Alternative methods of hydrolysis using TFA, acetic acid, and HCl were investigated, but these acids proved difficult to remove, leaving the resulting sugar acidic. The 3-[<sup>2</sup>H]-D-Glucose (**20**) synthesized in this manner could not be used for the investigation of a primary kinetic isotope effect for this reason.



Scheme 2.3: Synthesis of 3-[<sup>2</sup>H]-D-glucose (20). (a) NaBD<sub>4</sub>, 25% EtOH<sub>(aq)</sub>, 85%; (b) Tf<sub>2</sub>O, pyridine, DCM, -20°C; (c) Sodium benzoate, DMF; (d) NaOMe, MeOH,  $\Delta$ , 87% (3 steps); (e) Dowex 50X-80, H<sub>2</sub>O, quant. 74% overall yield from **2**.



Figure 2.44: The <sup>1</sup>H NMR spectra of 1,2;5,6-di-*O*-isopropylidene-β-D-allofuranose (1) and 3-[<sup>2</sup>H]- 1,2;5,6-di-*O*-isopropylidene-β-D-allofuranose (16).

# 2.8 Synthesis of 3-deoxy-3-fluoro-D-glucose (23)

It was predicted that 3-deoxy-3-fluoro-D-glucose (23) would act as an inhibitor for the reaction catalyzed by NtdC. One of the earliest syntheses of this compound was reported by

Johansson and Lindberg. They used boron trifluoride in hydrogen fluoride to open the 2,3epoxide of glucose. This method had the disadvantage of producing 2-deoxy-2-fluoro-D-altrose as the major product.<sup>105</sup> This method was improved upon later the same year by Webber *et al.* by reacting the easily obtained tosylate of 1,2:5,6-di-O-isopropylidene-β-D-allofuranose with tetrabutylammonium fluoride.<sup>106</sup> Thanks to the utility of [<sup>18</sup>F]-3-deoxy-3-fluoro-D-glucose in medical imaging, many other syntheses of 3-deoxy-3-fluoro-D-glucose (23) have been reported, with many focusing on short preparation and reaction times in order to maximize activity of this short lived, radioactive isotope of fluorine.<sup>107</sup> Many common preparations employ dimethylaminosulfur triflurode (DAST), with yields of compound 22 as high as 87% when converted directly from compound **3**.<sup>108</sup> However, the availability of cesium fluoride combined with the familiarity of displacing triflates at C-3 made the route developed by Argentini et al. attractive.<sup>109</sup> Scheme 2.4 displays the synthesis of 3-deoxy-3-fluoro-1,2:5,6-di-Oisopropylidene- $\beta$ -D-glucofuranose (22) from 1,2:5,6-di-*O*-isopropylidene- $\beta$ -D-alofuranose (3) via cesium fluoride displacement of the triflated intermediate (21). Compound 22 was obtained in 80% yield after two steps. Although 3-deoxy-3-fluoro-D-glucose (23) should be easy to obtain through hydrolysis of the isopropylidene protecting groups, the difficulty in obtaining fully functioning glucose from the diisopropylidene of glucose (1) cautioned us from performing this reaction until the above problem could be solved.



Scheme 2.4: Synthesis of 3-deoxy-3-fluoro-D-glucose (23). (a)  $Tf_2O$ , pyridine, DCM,  $-20^{\circ}C$ ; (b) CsF, DMF, 60 °C, 80% (2 steps); (c) Dowex 50X-80, H<sub>2</sub>O.

# **Chapter 3: Experimental**

#### **3.1** Reagents and materials

Chemicals and reagents were purchased from EMD (Gibbstown, NJ), TCI America (Portland, OR), Alfa Aesar (Ward Hill, MA), CalBiochem (San Diego, CA), or Sigma-Aldrich (St. Louis, MO). DNA primers were synthesized by either Integrated DNA Technologies, Inc. (Toronto, ON) or Alpha DNA (Montreal, QB). Restriction enzymes were purchased from either Invitrogen Canada Inc. (Burlington, ON) or New England Biolabs Inc. (Pickering, ON). T4 DNA Ligase was purchased from New England Biolabs Inc. *Pfu* Ultra HF was purchased from Invitrogen Canada Inc. as part of the QuikChange Site-Directed Mutagenesis Kit. QIAprep spin miniprep kits, QIAquick gel extraction kits, and QIAquick PCR Purification kits were purchased from GE Healthcare Life Sciences (Baie d'Urfe, QB). Acrodisc 0.45 µm syringe filters were manufactured by Pall Life Sciences (Ville St. Laurent, QB) and purchased from VWR Canlab (Mississauga, ON). For kinetic studies, 3-oxoglucose was synthesized by Hongyan Zheng (Department of Chemistry, University of Saskatchewan, Saskatoon, SK, Canada).

# 3.2 Plasmids, bacterial strains, and growth conditions

The pET-28b and pET-17b plasmids were purchased from Novagen. The pUC18ntdABC plasmid encoding the *ntdA*, *ntdB*, and *ntdC* genes was kindly provided by Dr. Kozo Ochi (Microbial Function Laboratory, National Food Research Institute, Tsukuba, Ibaraki, Japan). The mutant pET28b-ntdC-H180Y, pET28b-ntdC-H180A, and pET28b-ntdC-H180N plasmids were provided by Grey Wilkinson (Department of Chemistry, University of Saskatchewan, Saskatoon, SK, Canada). Competent cells were generated from *E. coli* XL1-Blue and BL21-Gold (DE3) cell lines which were also purchased from Novagen. *E. coli* XL1-Blue was used as the host/storage strain for the recombinant plasmid construction. BL21-Gold (DE3) cells were used for protein overexpression. *E. coli* strains harbouring corresponding plasmids were routinely grown in Luria Bertani (LB) medium containing 50  $\mu$ g/ml ampicillin or 30  $\mu$ g/ml kanamycin overnight at 37 °C and 250 rpm or on LB plates supplemented with the appropriate antibiotic overnight at 37 °C. Cell suspensions were made as aliquots (1 mL) in 15% glycerol, flash frozen in liquid nitrogen, and stored at -80 °C. Plates were stored at 4 °C.

# 3.3 Instrumentation

Flash column chromatography was performed with Merck silica gel 60 (230-400 mesh). The polymerase chain reaction (PCR) was performed using an Eppendorf mastercycler gradient thermocycler. Cells were lysed with the aid of a Virsonic 600 ultrasonic cell disrupter. Spectrophotometric assays were performed using a Beckman DU 640 spectrophotometer. DNA separation and analysis was performed using a BIO-RAD sub-cell GT agarose gel electrophoresis system. Protein separation and analyses was performed using a BIO-RAD miniprotean 3 electrophoresis system. Beckman Coulter microfuge 18 and 22R centrifuges were used for routine DNA purification, cell harvesting, and protein sample preparation. Molecular weights were determined using electro-spray mass spectrometry (Sciex QSTAR hybrid quadrupole-TOF) performed by the Saskatchewan Structural Science Centre (SSSC) (Saskatoon, SK, Canada). DNA sequencing was performed by the DNA Technologies Unit of the National Research Council (NRC), Plant Biotechnology Institute (PBI) (Saskatoon, SK, Canada).

# 3.4 Subcloning of the *ntd* genes

The genes designated *ntdA*, *ntdB*, and *ntdC* were received as part of the pUC18-ntdABC. Each gene was to be subcloned between the *Nde*I and *Bam*HI sites of pET-28b. However, *ntdA* possessed an internal *Nde*I site that needed to be removed before this could be effected.

#### 3.4.1 PCR-based mutagenesis of *ntdA*

Silent, a program from the *EMBOSS* suite<sup>110</sup>, was used to identify a silent T1152C mutation that would remove this site. *MutaPrimer* V1.0<sup>111</sup> was used to design primers encoding this T1152C mutation (Table 3.1). pUC18-ntdABC was isolated from an overnight growth of an *E. coli* DH5 $\alpha$  strain carrying the pUC18-ntdABC plasmid according to the protocol described in the QIAprep spin miniprep kit manual. pUC18-ntdABC, as with all plasmids and gene fragments henceforth, was eluted with deionized water. Mutagenesis was performed using the QuikChange Site-Directed Mutagenesis Kit. The reaction mixture contained a 0.25 mM mixture of dATP, dTTP, dCTP, and dGTP (12.5 nmol each), primer NtdA-T1152C-Forward (125 ng,

9.56 pmol), primer NtdA-T1152C-Reverse (125 ng, 9.35 pmol), plasmid pUC18-ntdABC (400 ng, 100 fmol), and *Pfu* Ultra HF DNA polymerase (2.5 U) in 1x *Pfu* Ultra HF PCR buffer (50  $\mu$ L). The thermocycler was programmed as follows: 1) Denaturation at 95 °C for 30 seconds; 2) Denaturation at 95 °C for 30 seconds; 3) Annealing at 55 °C for 60 seconds; 4) Extension at 68 °C for 360 seconds; 5) Repeat from step 2 for 12 cycles; 6) Hold at 4 °C. *DpnI* (10 U) was added and the reaction mixture allowed to incubate overnight at 37 °C in order to degrade the parental DNA.

**Table 3.1: Primers for subcloning of the** *ntd* genes. Unchanged nucleotides are capitalized. Mutated nucleotides are lower case. Introduced restriction sites are underlined.

Purpose	Primer	Oligonucleotide sequence (5 ' $\rightarrow$ 3')
Subcloning of <i>ntdA</i>	<i>ntdA-Ndel</i> -Forward <i>ntdA-Bam</i> HI-Reverse	GGA GGT ACT G <u>ca tAT G</u> CA AAA ACA G CAA CCG TGG At <u>g gat Cc</u> T TAT ACT CCG ATT TC
Subcloning of <i>ntdB</i>	<i>ntdB-Nde</i> I-Forward <i>ntdB-Bam</i> HI-Reverse	GGA GGG ACT GT <u>c atA TG</u> T TAT TAA G GCC TAT CTT CT <u>g gAT cc</u> T TAT TTC CTC CTC
Subcloning of <i>ntdC</i>	<i>ntdC-Nde</i> l-Forward <i>ntdC-Bam</i> HI-Reverse	CAT GAG GAG GAA <u>cat Atg</u> ATG AAG AAG ATA CG CTT CTC TCT TT <u>G GAT CC</u> C TAG TTG ACT G
Removal of an internal <i>Nde</i> l site from <i>ntdA</i>	ntdA-T1152C-Forward ntdA-T1152C-Reverse	TCC AAT ACT TTC TCA cAT GCA AAA AAC TCC TTT AGT ACA GGA C GTC CTG TAC TAA AGG AGT TTT TTG CAT gTG AGA AAG TAT TGG A

#### 3.4.2 Transformation of competent cells with mutant plasmid

*E. coli* XL1-Blue competent cells were prepared as follows: LB (5 mL) containing tetracycline (20  $\mu$ g/mL) was inoculated with a colony of XL1-Blue cells and allowed to incubate at 37 °C and 250 rpm overnight. The overnight culture (0.5 mL) was used to inoculate a 250 mL Erlenmeyer flask containing LB (50 mL) and tetracycline (20  $\mu$ g/mL). The flask was incubated at 37 °C at 250 rpm shaking until an optical density at 600 nm (OD<sub>600</sub>) of roughly 0.4 was reached. The cell culture was harvested by centrifugation at 11,000 x*g* and 4 °C for ten minutes. The bacterial pellet was resuspended in transformation buffer (25 mM Tris (tris(hydroxymethyl)aminomethane)/HCl, 50 mM CaCl<sub>2</sub>, pH 7.5) (20 mL) and incubated on ice for one hour. The resuspended cells were pelleted as above and resuspended in transformation buffer (1 mL). The competent cells were diluted with sterile, ice-cold 80% glycerol to a final concentration of 15% glycerol. The cell suspension was flash frozen as 100  $\mu$ L aliquots in liquid nitrogen and stored at -80 °C until used.

A frozen aliquot of XL1-Blue competent cells (100  $\mu$ L) was thawed on ice before adding the crude PCR mix from above (1  $\mu$ L). The transformation mixture was incubated on ice for 30 minutes, heat shocked at 42 °C for 45 seconds, and placed back on ice for two additional minutes. The entire transformation mixture was used to inoculate a test tube containing pre-warmed LB (0.5 mL). The cell suspension was incubated at 37 °C with 250 rpm shaking for one hour. The entire sample was used to inoculate two four inch, LB-agar plates containing ampicillin (50  $\mu$ g/mL). These plates were allowed to incubate overnight at 37 °C, yielding several colonies.

#### 3.4.3 Identification of the *ntdA* mutant

Four colonies were picked from the above plates and grown overnight in LB containing ampicillin (50  $\mu$ g/mL). Their plasmids were extracted according to the protocol described in the QIAprep spin miniprep kit manual. Plasmids, diluted in 1x NEB buffer #4 (10  $\mu$ l) containing *Nde*I (10 U), were incubated at 37 °C overnight. Identification of mutated plasmids was carried out by visualizing the products of the *Nde*I digest on an ethydium bromide stained 1% agarose gel that had been developed in Tris-Acetate-EDTA (TAE) buffer at 120 V.

# 3.4.4 PCR-based amplification of the individual *ntd* genes

pUC18-ntdABC-T1152C was isolated from an overnight growth of an *E. coli* XL1-Blue strain carrying the pUC18-ntdABC-T1152C plasmid according to the protocol described in the QIAprep spin miniprep kit manual. The PCR primers for the gene amplification are shown in Table 3.1. The primer sequence designs were generated using *NEBcutter* V2.0<sup>79</sup> and *OligoAnalyzer* V3.1 from Integrated DNA Technologies. The reaction mixture contained a 0.25 mM mixture of dATP, dTTP, dCTP, and dGTP (12.5 nmol each), forward primers (1250 ng, 125-161 pmol), reverse primers (1250 ng, 128-139 pmol), template plasmid pUC18-ntdABC-Mut (400 ng, 100 fmol), and *Pfu* Ultra HF DNA polymerase (2.5 U) in 1x *Pfu* Ultra HF PCR buffer (50  $\mu$ L). The thermocycler was programmed as follows: 1) Denaturation at 95 °C for 60 seconds; 2) Annealing at 54.5 °C for 60 seconds; 3) Extension at 72 °C for 120 s; 4) Repeat 12 for 12 cycles; 5) Hold at 4 °C. *DpnI* (10 U) was added and the reaction mixture allowed to incubate one hour at 37 °C in order to degrade the parental DNA. The PCR product was purified from the reaction mixture using a QIAquick PCR Purification Kit from Qiagen.

# 3.4.5 Restriction digests of pET-28b and the individual *ntd* PCR products

The pET-28b plasmid and the PCR amplified DNA fragments were digested as follows. Plasmids or purified PCR products in 1x NEB buffer #2 (10  $\mu$ l) containing *Nde*I (20 U) were incubated at 37 °C for 48 hours. *Bam*HI (5 U) was added, before incubating an additional one hour. Restriction enzymes and sub-100 bp fragments were removed using the QIAquick PCR Purification Kit.

#### 3.4.6 Ligation of the individual *ntd* genes into pET-28b

The ligation reactions were prepared at both 3:1 and 10:1 (90 fmol:9 fmol) insert to vector ratios. For *ntdC*, the ligation mixture contained the *NdeI/Bam*HI digested PCR product containing the *ntdC* gene (27 or 90 fmol), *NdeI/Bam*HI digested pET-28b (9 fmol), and T4 DNA Ligase (200 U) in 2x T4 DNA Ligase buffer (10  $\mu$ L) containing ATP (20 nmol). The ligation mixture was allowed to incubate at 16 °C for 48 hours. The constructs containing *ntdA* and *ntdB* were prepared in a similar fashion, only using a lower concentration of insert (15 or 50 fmol) and vector (5 fmol) and a shorter incubation time of 24 hours. The ligation mixtures were used to transform XL1-Blue competent cells in the same manner as describe above. These cells were plated on to LB-Agar plates containing kanamycin (30  $\mu$ g/mL).

#### 3.4.7 Identification of successfully subcloned genes

For *ntdC*, the transformed cells from the 10:1 ligation reactions produced the most colonies. The plasmids from six of these colonies were extracted as described above. For *ntdA* and *ntdB*, transformed cells from the 3:1 ligation reactions produced the most colonies. The plasmids from six of these colonies were extracted for both *ntdA* and *ntdB*. Successful ligations were identified by visualizing the products of *Xmn*I, *Eco*RV, and *Bgl*II digests for ntdA, ntdB, and ntdC, respectively, on an ethydium bromide stained 1% agarose gel that had been developed in TAE buffer at 120 V.

#### 3.4.8 Subcloning of *ntdC* into pET-17b

Both pET-17b and pET28b-ntdC were digested with *NdeI* and *Bam*HI as described above. The products of both digests were visualized on an ethydium bromide stained 1% agarose gel that had been developed in TAE buffer at 120 V. The 1059 bp band corresponding to the *ntdC* gene fragment and the 3243 bp band corresponding to empty pET-17b vector were excised and purified using the QIAquick Gel Purification Kit. A 7:1 insert to vector ligation reaction was performed using the *NdeI/Bam*HI digested *ntdC* gene insert (50 fmol), *NdeI/Bam*HI digested pET-17b vector (7 fmol), and T4 DNA Ligase (200 U) in 2x T4 DNA Ligase buffer (10  $\mu$ L) containing ATP (20 nmol). The ligation mixture was allowed to incubate at 16 °C for 48 hours. The ligation mixture (1  $\mu$ L) was used to transform XL1-Blue competent cells as described above.

determined by visualizing the products of a *Bgl*II digest on an ethydium bromide stained 1% agarose gel developed in TAE buffer.

# 3.5 PCR-based mutagenesis of the *ntdC* Gene

A D176N mutation was effected by changing the guanine at position 526 to an adenine. This did not affect the restriction map for plasmid pET28b-ntdC; the program, *Silent*, from the *EMBOSS* suite<sup>110</sup> was used to identify a silent adenine to cytosine change at position 519 which would simultaneously introduce *Bam*HI and *Sty*I restriction sites. A separate K227R mutation was generated by changing an adenine at position 680 to a guanine. No convenient silent mutation could be identified to alter the restriction map. *MutaPrimer* V1.0<sup>111</sup> was used to design primers encoding these D176N and K227R mutants (Table 3.2). The plasmid, pET28b-ntdC, was isolated from an overnight growth of transformed XL1-Blue cells according to the protocol described in the QIAprep spin miniprep kit manual.

Mutagenesis was performed using the QuikChange Site-Directed Mutagenesis Kit from Stratagene. The reaction mixture contained a 0.25 mM mixture of dATP, dTTP, dCTP, and dGTP (12.5 nmol each), forward primer (125 ng, 10.9 pmol), reverse primer (125 ng, 11.1 pmol), plasmid pET28b-ntdC (30 ng, 7.18 fmol), and *Pfu* Ultra HF DNA polymerase (2.5 U) in 1x *Pfu* Ultra HF PCR buffer (50  $\mu$ L). The thermocycler was programmed as follows: 1) Denaturation at 95 °C for 30 seconds; 2) Denaturation at 95 °C for 30 seconds; 3) Annealing at 55 °C for 60 seconds; 4) Extension at 72 °C for 420 seconds; 5) Repeat from step 2 for 15 cycles; 6) Hold at 4 °C. *Dpn*I (10 U) was added and the reaction mixture allowed to incubate overnight at 37 °C in order to degrade the parental DNA. XL1-Blue competent cells were transformed with the PCR mixture as described above. Plasmids were extracted from six of the colonies produced from the D176N reaction. Successful mutants were identified by visualizing the products of a *Bam*HI/*Eco*RV digests on ethydium bromide stained 1% agarose gels developed in TAE buffer at 120 V. For the K227R mutation, successful mutants were identified by sequencing.

**Table 3.2: Primers for mutagenesis of** *ntdC***.** Unchanged nucleotides are capitalized. Mutated nucleotides are lower case. Introduced restriction sites are underlined.

Mutation	Primer	Oligonucleotide sequence (5 ' $\rightarrow$ 3')
K227R	ntdC-K227R-Forward ntdC-K227R-Reverse	GCA GAA ATT TAT GGC CAG CTG AgG AAC AAG GTA TTT G CAA ATA CCT TGT TCc TCA GCT GGC CAT AAA TTT CTG C
D176N	<i>ntd</i> C-D176N-Forward <i>ntd</i> C-D176N-Reverse	GAA GAC GAG TGG AT <u>C cTT GG</u> G GaA TCT GGG TAA TCA C GTG AAT ACC CAG ATt CC <u>C CAA gG</u> A TCC ACT CGT CTT C

# **3.6** Overexpression of the His-tagged *ntdC* gene and its mutants

BL21 Gold (DE3) competent cells, prepared in the same manner as XL1-Blue competent cells only in the absence of tetracycline, were transformed with the newly created pET28b-ntdC plasmid (or the H180N, H180Y, H180A, D176N, or K227R plasmids). The transformed cells were plated onto LB-agar containing kanamycin (30  $\mu$ g/mL). A single colony was used to inoculate a test tube containing LB (5 mL) and kanamycin (30  $\mu$ g/mL). The test tube was incubated overnight at 37 °C with shaking at 250 rpm. The overnight culture (500  $\mu$ L) was used to inoculate a 250 mL Erlenmeyer flask containing LB (50 mL) and kanamycin (30  $\mu$ g/mL). The cell culture was allowed to incubate at 37 °C with 250 rpm of shaking until an OD<sub>600</sub> of roughly 0.6 was reached. Overexpression of the *ntdC* gene was induced by the addition of IPTG (1.0 mM). The cell culture was incubated for an additional three hours.

# **3.7** Purification of NtdC and its mutants

Cell cultures were harvested at 11,000 xg for ten minutes and resuspended in binding buffer (20 mM Tris/HCl, 5 mM imidazole, 500 mM NaCl, 12.5% (v/v) glycerol, pH 7.9) (2.5 mL) supplemented with an additional 20% glycerol to prevent precipitation. Samples were cooled on ice and sonicated ten times (two seconds on and five seconds off) with the sonicator set at level four, and then allowed to cool for five minutes before sonicating again. In total, samples were sonicated 30 times. The sonicated cells were pelleted by centrifugation at 15,000 xg for ten minutes. The supernatants were passed through a 0.45  $\mu$ m filter before being loaded onto HiTrap Chelating HP columns (1 mL) charged with nickel. The columns were washed successively with binding buffer (4 mL) and washing buffer (20 mM Tris/HCl, 60 mM imidazole, 500 mM NaCl, 12.5% (v/v) glycerol, pH 7.9) (8 mL) before eluting the target protein with eluting buffer (20 mM Tris/HCl, 2.5 mM imidazole, 50 mM EDTA, 500 mM NaCl, 12.5% (v/v) glycerol, pH 7.9) (13.5 mL). As with the binding buffer, all buffers were supplemented with an additional 20% glycerol. Nine 1.5 mL fractions of eluting buffer were collected for each sample. The presence of NtdC or one of its mutants was verified by SDS-PAGE. Fractions containing the desired protein were pooled and dialyzed overnight into a pH 6.5 potassium phosphate buffer (50 mM) containing glycerol (50%) and DTT (1 mM). Protein concentrations were determined by UV-spectroscopy using an  $\varepsilon_{280}$  of 1.03 mL·mg<sup>-1</sup>·cm<sup>-1</sup> for wild type NtdC, the H180N mutant, and the H180A mutant, 1.02 mL·mg<sup>-1</sup>·cm<sup>-1</sup> for the D176N and K227R mutants, and 1.06 mL·mg<sup>-1</sup>·cm<sup>-1</sup> for the H180Y mutant. The molar absorptivity values were calculated using the ProtParam tool provided on the ExPASy website.

# 3.8 Overexpression of the non-His-tagged *ntdC* gene

BL21 Gold (DE3) competent cells were transformed with the pET17b-ntdC plasmid. The transformed cells were plated onto LB-agar containing ampicillin (50 µg/mL). A single colony was used to inoculate a test tube containing LB (5 mL) and ampicillin (50 µg/mL). The test tube was incubated overnight at 37 °C with shaking at 200 rpm. The overnight culture (50  $\mu$ L) was used to inoculate four test tubes containing LB (10 mL) and ampicillin (50  $\mu$ g/mL). These test tubes were allowed to incubate for three hours at 37 °C with shaking at 200 rpm. One of the test tubes was removed and placed on ice. One test tube was induced by the addition of IPTG (0.8 mM) and allowed to incubate an additional three hours. Another test tube was induced by the addition of IPTG (0.8 mM) and allowed to incubate overnight. The final test tube was not induced but allowed to incubate overnight as a control. An additional control consisting of BL21 Gold (DE3) competent cells transformed with empty pET-17b and induced with IPTG was also allowed to incubate overnight. The cell cultures were harvested by centrifugation at 11,000 xg for ten minutes and resuspended in binding buffer (20 mM Tris/HCl, 5 mM imidazole, 500 mM NaCl, 12.5% (v/v) glycerol, pH 7.9) (0.5 mL). The cell suspensions were cooled on ice and sonicated ten times (2 seconds on, three seconds off) at power four. The cell suspensions were pelleted at 15,000 xg for ten minutes. The supernatants and pellets were separated and the presence of protein visualized on Coomassie brilliant blue stained SDS-PAGE gels.

#### **3.9** Initial velocity experiments

Reaction progress was recorded using a Beckman DU640 spectrophotometer by measuring the appearance (or disappearance) of NADH at its  $\varepsilon_{max}$  of 340 nm where the absorbance contributed by the enzyme is minimal. The spectrophotometer was connected to a circulating water bath to maintain a constant temperature. The recorded data were converted to a Microsoft Excel file using the DU600/7000 file utility software (version 1.0). A plot of absorbance versus time was generated from the data. An example plot is shown in Figure 3.1. All assays were performed using His-tagged enzymes.



**Figure 3.1: Example plot of absorbance versus time for a typical assay.** Conditions: Conditions: 100 mM AMP·HCl, pH 9.2, 25 °C, 1 mM NAD<sup>+</sup>, 2 mM G6P, 1.6 µg NtdC.

The initial velocity, in the form of the change in absorbance per second, was approximated from the slope of the early, approximately linear part of the curve. Approximately 20-30 seconds elapsed between mixing of the cuvette and reading, which was not adjusted for. The initial rate was converted to micromoles per minute per milligram of NtdC in the following manner:

$$Abs = \varepsilon_{340} \cdot C \cdot L \tag{1}$$

Equation 1 represents Beer's Law, where Abs represents absorbance,  $\varepsilon_{340}$  represents the molar absorptivity of NADH at 340 nM (6,300 mL·mmol<sup>-1</sup>·cm<sup>-1</sup>), C represents the concentration of NADH, and L represents the path length (1 cm). This can be rearranged to form equation 2.

$$C = Abs / (\varepsilon_{340} \cdot L)$$
<sup>(2)</sup>

$$\mathbf{C} = \mathbf{n} / \mathbf{V} \tag{3}$$

Equation 3 relates concentration to the number of moles (n) in a certain volume (V). Equations 2 and 3 can be combined to form equation 4 and rearranged to form equation 5.

$$n / V = Abs / (\varepsilon_{340} \cdot L)$$
<sup>(4)</sup>

$$\mathbf{n} = \mathbf{Abs} \cdot \mathbf{V} / (\mathbf{\epsilon}_{340} \cdot \mathbf{L}) \tag{5}$$

Equation 6 is produced when the initial rate ( $\Delta Abs_{sec}$ ) is substituted for Abs

$$\Delta n_{\text{sec}} = \Delta Abs_{\text{sec}} \cdot V / (\epsilon_{340} \cdot L)$$
(6)

As the volume of the reaction mixture is known (1 mL), equation 6 can be used to determine the amount of NADH produced per second. This value was normalized per milligram of NtdC and the time scale converted to minutes, instead of seconds, so that values could more easily be compared between experiments. The overall conversion is described by equation 7, where W is the mass of protein in milligrams. All rates are reported after conversion in this manner.

$$\Delta n_{\min} = \Delta Abs_{sec} \cdot V / (\epsilon_{340} \cdot L \cdot W) \cdot (60 \text{ s} / 1 \text{ min})$$
<sup>(7)</sup>

# 3.9.1 Optimization of NtdC glucose oxidase activity

Assays were performed in a one millilitre volume containing Tris-HCl (100 mM), or AMP-HCl (100mM), glucose (50 mM), NAD<sup>+</sup> (1 mM), and NtdC (16  $\mu$ g) at 25 °C and varying pH. All assays were performed at least in duplicate.

#### **3.9.2** Assaying NtdC oxidase activity with different sugar substrates

Assays were performed in a one millilitre volume containing Tris-HCl (100 mM) and varying concentrations of NtdC at 25 °C and pH 9.0. NAD<sup>+</sup> concentrations varied between one and five millimoles per litre. Different sugar substrates were added at different concentrations. In some cases, two sugars were added together to test for an inhibitory effect. All assays were performed at least in duplicate.

## **3.9.3** Determining K<sub>m</sub><sup>app</sup> for glucose

Assays were performed in a one millilitre volume containing Tris-HCl (100 mM), NAD<sup>+</sup> (1 mM), and NtdC (1.49  $\mu$ g) at 25 °C and pH 9.0. Glucose concentrations were varied between 20 and 500 millimoles per litre. All assays were performed at least in duplicate with the average reported. The program *Leonora*<sup>90</sup> was used for the non-linear, least-squares fitting of the initial rate data to the equations describing Michaelis-Menten kinetics for a single substrate system (Equation 8):

$$\mathbf{v}^{\mathrm{app}} = \mathbf{V}_{\mathrm{max}}^{\mathrm{APP}} \cdot \mathbf{A} / [\mathbf{K}_{\mathrm{m}}^{\mathrm{app}} + \mathbf{A}]$$
(8)

where  $v^{app}$  represents the apparent initial rate at a given concentration of the varied substrate,  $V_{max}^{app}$  represents the apparent maximum initial rate at saturating concentrations of the varied substrate (in this case, glucose), A represents the concentration of the varied substrate, and  $K_m^{app}$  represents the apparent Michaelis constant for the varied substrate. These are apparent values and not true values, as they are only valid while the concentration of the fixed substrate (in this case, NAD<sup>+</sup>) remains unchanged.

# 3.9.4 Optimization of NtdC G6P oxidase activity

Assays were performed in a one millilitre volume containing Tris-HCl (100 mM), or AMP-HCl (100mM), glucose-6-phosphate (2 mM), NAD<sup>+</sup> (1 mM), and NtdC (0.74  $\mu$ g) at 25 °C and varying pH. The pH of the reaction mixtures were re-evaluated after the assay was complete to account for any pH change over the course of the reaction. All assays were performed at least in duplicate. In addition to fitting the data to Equation 8, the data were also fit to the equation

describing a single substrate reaction which obeys Michaelis-Menten kinetics where the varied substrate binds cooperatively (Equation 9):

$$\mathbf{v}^{\mathrm{app}} = \mathbf{V}_{\mathrm{max}}^{\mathrm{app}} \cdot \mathbf{A}^{h} / \left[ S_{0.5}^{\mathrm{app}} \cdot (1 + \mathbf{A}) \right]$$
(9)

where *h* represents the Hill factor describing the cooperativity of the binding of the substrate and  $S_{0.5}^{app}$  represents the apparent concentration of the varied substrate at which half of  $V_{max}^{app}$  is obtained. The value of the Hill constant was determined in an iterative fashion by calculating  $V_{max}^{app}$  using h = 1, using the calculated  $V_{max}^{app}$  to generate a Hill plot from which a preliminary Hill constant could be derived, then using this Hill constant to calculate a new  $V_{max}^{app}$ . This process was repeated until no additional improvement in the fit could be obtained. The error in estimating the Hill constant was calculated by measuring the slopes of Hill plots constructed using the maximum and minimum values estimated for  $V_{max}$  in order to determine the Hill constant's range.

# **3.9.5** Determination of the true kinetic constants for the oxidation of G6P by NAD<sup>+</sup> as catalyzed by NtdC

Assays were performed in a one millilitre volume containing AMP-HCl (100 mM) and NtdC (1.5  $\mu$ g) at 25 °C and pH 9.2. The concentrations of G6P (30  $\mu$ M-2000  $\mu$ M) and NAD<sup>+</sup> (30  $\mu$ M-1000  $\mu$ M) were varied. All assays were performed at least in duplicate. The program *Leonora*<sup>90</sup> was used for the non-linear, least-squares fitting of the initial rate data to the equation describing Michaelis-Menten kinetics for a ternary complex mechanism with and without the inclusion of the Hill factor for the B term (Equations 10 and 11):

$$\mathbf{v} = \mathbf{V}_{\text{max}} \cdot \mathbf{A} \cdot \mathbf{B} / \left[ \mathbf{K}_{\text{iA}} \cdot \mathbf{K}_{\text{mB}} + \mathbf{K}_{\text{mB}} \cdot \mathbf{A} + \mathbf{K}_{\text{mA}} \cdot \mathbf{B} + \mathbf{A} \cdot \mathbf{B} \right]$$
(10)

$$\mathbf{v} = \mathbf{V}_{\max} \cdot \mathbf{A} \cdot \mathbf{B}^{h} / [\mathbf{K}_{iA} \cdot S_{0.5B} + S_{0.5B} \cdot \mathbf{A} + \mathbf{K}_{mA} \cdot \mathbf{B}^{h} + \mathbf{A} \cdot \mathbf{B}]$$
(11)

where v represents the initial rate at a given concentration of substrates A and B,  $V_{max}$  represents the maximum initial rate at saturating A and B, A represents the concentration of the first substrate to bind, B represents the concentration of the second substrate to bind, *h* represents the

Hill factor,  $K_{iA}$  represents the Michaelis term for the inhibition of the binding of substrate B by a lack of substrate A,  $K_{mB}$  represents the Michaelis constant for the second substrate in a noncooperative system,  $S_{0.5B}$  represents the concentration of B at which half of  $V_{max}$  is obtained in a cooperative system, and  $K_{mA}$  represents the Michaelis constant for the first substrate. The value of the Hill constant was determined iteratively from the slopes of the Hill plots generated from the initial rates.

While searching for an alternative explanation to negative cooperativity, the data were also fitted to the Michaelis-Menten model describing a ternary complex mechanism that exhibits substrate inhibition by the second substrate (Equation 12) and the model describing an equilibrium ordered mechanism (Equation 13):

$$\mathbf{v} = \mathbf{V}_{\text{max}} \cdot \mathbf{A} \cdot \mathbf{B} / \left[ \mathbf{K}_{\text{iA}} \cdot \mathbf{K}_{\text{mB}} + \mathbf{K}_{\text{mB}} \cdot \mathbf{A} + \mathbf{K}_{\text{mA}} \cdot \mathbf{B} + (\mathbf{A} \cdot \mathbf{B}) \cdot (1 + \mathbf{B} / \mathbf{K}_{\text{siB}}) \right]$$
(12)

$$\mathbf{v} = \mathbf{V}_{\text{max}} \cdot \mathbf{A} \cdot \mathbf{B} / \left[ \mathbf{K}_{\text{iA}} \cdot \mathbf{K}_{\text{mB}} + \mathbf{K}_{\text{mB}} \cdot \mathbf{A} + \mathbf{A} \cdot \mathbf{B} \right]$$
(13)

where  $K_{siB}$  represents the Michaelis constant for the inhibition of the reaction by the second substrate.

#### **3.9.6** Determining the inhibitory effect of NADH on G6P and NAD<sup>+</sup> binding

Assays were performed in a one millilitre volume containing AMP-HCl (100 mM) and NtdC (1.49  $\mu$ g) at 25 °C and pH 9.2. In one set of experiments, the concentration of G6P was fixed (2 mM), while the concentrations of NAD<sup>+</sup> and NADH were varied. In another set of experiments, the concentration of NAD<sup>+</sup> was fixed (0.2 mM), while the concentrations of G6P and NADH were varied. All assays were performed at least in duplicate. The program *Leonora*<sup>90</sup> was used for the non-linear, least-squares fitting of the initial rate data to the equations describing Michaelis-Menten kinetics with terms denoting competitive inhibition, uncompetitive inhibition, or both (mixed inhibition) (Equations 14, 15, and 16, respectively):

$$\mathbf{v}^{app} = \mathbf{V}_{max}^{app} \cdot \mathbf{A} / [\mathbf{K}_{m}^{app} \cdot (1 + \mathbf{I} / \mathbf{K}_{ic}^{app}) + \mathbf{A}]$$
(14)

$$\mathbf{v}^{app} = \mathbf{V}_{max}^{app} \cdot \mathbf{A} / [\mathbf{K}_{m}^{app} + \mathbf{A} \cdot (1 + \mathbf{I} / \mathbf{K}_{iu}^{app})]$$
(15)

$$v^{app} = V_{max}^{app} \cdot A / [K_m^{app} \cdot (1 + I / K_{ic}^{app}) + A \cdot (1 + I / K_{iu}^{app})]$$
(16)

where  $v^{app}$  represents the apparent initial rate at a given concentration of the varied substrate and inhibitor,  $V_{max}^{app}$  represents the apparent maximum initial rate at saturating concentrations of the varied substrate, A represents the concentration of the varied substrate, I represents the concentration of inhibitor,  $K_m^{app}$  represents the apparent Michaelis binding constant for the varied substrate,  $K_{ic}^{app}$  represents the apparent Michaelis competitive inhibition constant of the inhibitor, and  $K_{iu}^{app}$  represents the apparent Michaelis uncompetitive inhibition constant of the inhibitor. These are apparent values and not true values, as they are only valid while the concentration of the fixed substrate remains unchanged. For the inhibition of G6P binding by NADH, the initial rate data were fit to the equations describing Michaelis-Menten kinetics with terms denoting competitive inhibition, uncompetitive inhibition, or both (mixed inhibition), with the inclusion of the Hill factor for substrate binding (Equations 17, 18, and 19):

$$\mathbf{v}^{\text{app}} = \mathbf{V}_{\text{max}}^{\text{app}} \cdot \mathbf{A}^{h} / [S_{0.5}^{\text{app}} \cdot (1 + \mathbf{I} / \mathbf{K}_{\text{ic}}^{\text{app}}) + \mathbf{A}^{h}]$$
(17)

$$\mathbf{v}^{\text{app}} = \mathbf{V}_{\text{max}}{}^{\text{app}} \cdot \mathbf{A}^{h} / \left[ S_{0.5}{}^{\text{app}} + \mathbf{A}^{h} \cdot (1 + \mathbf{I} / \mathbf{K}_{\text{iu}}{}^{\text{app}}) \right]$$
(18)

$$\mathbf{v}^{\text{app}} = \mathbf{V}_{\text{max}}^{\text{app}} \cdot \mathbf{A}^{h} / \left[ S_{0.5}^{\text{app}} \cdot (1 + \mathbf{I} / \mathbf{K}_{\text{ic}}^{\text{app}}) + \mathbf{A}^{h} \cdot (1 + \mathbf{I} / \mathbf{K}_{\text{iu}}^{\text{app}}) \right]$$
(19)

#### **3.9.7** Determination of mutant activity

Initial assays were performed in a one millilitre volume containing AMP-HCl (100 mM), G6P (2 mM), and NAD<sup>+</sup> (1 mM) at 25  $^{\circ}$ C and pH 9.2. The concentrations of the H180Y, H180N, H180A, D176N, and K227R mutants ranged from 2.9 µg/mL to 16.4 µg/mL.

For H180N, a pH profile was constructed in a similar manner as above. Assays were performed in a one millilitre volume containing Tris-HCl (100 mM), or AMP-HCl (100mM), G6P (10 mM), NAD<sup>+</sup> (5 mM), and H180N (11.5  $\mu$ g) at 25 °C and varying pH. The pH of the

reaction mixtures were remeasured after the assay was complete to account for any pH change over the course of the reaction. All assays were performed at least in duplicate.

The apparent Michaelis constants were then derived for H180N. Assays were performed in a one millilitre volume containing AMP-HCl (100 mM) and H180N (11.5  $\mu$ g) at 25 °C and pH 9.2. In one set of experiments, the concentration of G6P was fixed (150 mM) while the concentrations of NAD<sup>+</sup> and NADH were varied. In another set of experiments, the concentration of NAD<sup>+</sup> was fixed (4 mM) while the concentrations of G6P and NADH were varied. All assays were performed at least in duplicate.

# 3.10 Synthesis of Kanosamine

The synthesis of kanosamine was performed similar to the method reported with the exception of the installation of a tosyl group instead of a triflate group for nucleophilic displacement and the use of catalytic hydrogenation by Pd/C instead of reduction of the azide by LiAlH<sub>4</sub>.<sup>112</sup>

# **3.10.1 1,2;5,6-Di**-*O*-isopropylidene-β-L-lyxo-hex-3-ulose (2)

A mixture of the diacetone of glucose (1) (9.27 g, 35.7 mmol) and pyridinium dichromate (10.2 g, 26.4 mmol) was dried in an oven at 110 °C for one hour. The mixture was allowed to cool to room temperature under argon before adding freshly distilled DCM (100 mL), followed by acetic anhydride (10.2 mL, 108 mmol). The solution was refluxed with stirring for 3.5 hours. DCM was removed under reduced pressure, yielding a black, oily solid, which was resuspended in ethyl acetate. Large pieces were crushed into smaller particles. The solution was stirred vigorously for 20 minutes in order to extract compound **2** from the solid pyridinium dichromate pieces. The solution was filtered through celite to remove solid pyridinium dichromate, and the solvent removed under reduced pressure, yielding a black oil. This oil was loaded onto a 1.5 inch plug of silica and eluted with ethyl acetate in order to remove colour. Fractions containing compound **2** were pooled and the solvent removed under reduced pressure to yield a pale green oil that smelled faintly of acetic anhydride. Compound **2** was used immediately in the next step.



# **3.10.2 1,2;5,6-Di**-*O*-isopropylidene-β-D-allofuranose (3)

Compound **2** was diluted in 56% ethanol (40 mL). A solution of sodium borohydride (1.67 g, 44.5 mmol) in deionized water (46 mL) was added, slowly, with stirring. The reaction mixture was allowed to stir overnight. Ethanol was removed under reduced pressure to yield a pale blue solution which was extracted three times with ethyl acetate. The organic phase was dried with anhydrous magnesium sulfate and vacuum filtered. Solvent was removed in vacuo, yielding compound **3** as a white solid (7.96 g, 86% from compound **1**). The <sup>1</sup>H NMR spectrum was consistent with that reported.<sup>112</sup>

<sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  5.95 (d, *J* = 3.6 Hz, 1H), 4.54 (d, *J* = 3.6 Hz, 1H), 4.38-4.31 (m, 2H), 4.18 (dd, *J* = 8.6, 6.2 Hz, 1H), 4.08 (dd, *J* = 7.5, 2.7 Hz, 1H), 3.98 (dd, *J* = 8.6, 5.4 Hz, 1H), 2.48 (d, *J* = 3.6 Hz, 1H), 1.50 (s, 3H), 1.45 (s, 3H), 1.37 (s, 3H), 1.32 (s, 3H).



#### **3.10.3** 1,2;5,6-Di-*O*-isopropylidene-3-*O*-tosyl-β-D-allofuranose (4)

Compound **3** (3.44 g, 13.2 mmol) was added to a stirring solution of p-toluenesulfonyl chloride (7.13 g, 37.3 mmol) in pyridine (34 mL). The reaction mixture was allowed to stir under argon for three days. The reaction was quenched by the slow addition of water (20 mL). The solution was allowed to stir for ten minutes, and then slowly poured into a vigorously stirring solution of ice water (400 mL). Vacuum filtration resulted in a white powder contaminated with p-toluenesulfonic acid. The solution. The organic phase was dried with

anhydrous magnesium sulfate and vacuum filtered. Removal of the solvent under reduced pressure yielded compound **4** as a white solid (5.17 g, 94%). The <sup>1</sup>H NMR spectrum was consistent with that of a diisopropylidene of allose possessing a tosyl group.

<sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  7.87 (d, *J* = 8.3 Hz, 2H), 7.34 (d, *J* = 8.1 Hz, 2H), 5.76 (d, *J* = 3.0 Hz, 1H), 4.65 (dd, *J* = 3.1, 2.3 Hz, 2H), 4.25-4.09 (m, 2H), 3.93 (dd, *J* = 8.4, 6.8 Hz, 1H), 3.78 (dd, *J* = 8.4, 6.7 Hz, 1H), 2.45 (s, 3H), 1.53 (s, 3H), 1.32 (s, 3H), 1.30 (s, 3H), 1.28 (s, 3H).



#### **3.10.4 3-Azido-3-deoxy-1,2;5,6-di**-*O*-isopropylidene-β-D-glucofuranose (5)

Oven-dried (4 hours at 110 °C) sodium azide (18.3 g, 281 mmol) was added to a stirring solution of compound **4** (11.6 g, 28.1 mmol) in dry DMF (113 mL) under an atmosphere of argon. The reaction mixture was heated at 150 °C for 24 hours. Solid sodium azide was removed by vacuum filtration. DMF was removed at 80 °C under reduced pressure, yielding a dark sludge which was dissolved in ethyl acetate and washed three times with deionized water and once with brine. The organic phase was dried with anhydrous magnesium sulfate and vacuum filtered. Removal of the solvent under reduced pressure yielded a brown oil containing the starting material **4** and compound **5**. This oil was loaded onto a silica gel column and eluted with a 5:1 mixture of hexanes and ethyl acetate. The starting material, **4**, was isolated (0.74 g, 6.4%) as well as compound **5** (7.03 g, 88%). The <sup>1</sup>H NMR spectrum was consistent with that reported.<sup>112</sup>

<sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  5.85 (d, J = 3.5 Hz, 1H), 4.62 (d, J = 3.6 Hz, 1H), 4.24 (dt, J = 9.6, 5.0 Hz, 1H), 4.14 (dd, J = 8.7, 6.1 Hz, 1H), 4.11-4.06 (m, 2H), 3.98 (dd, J = 8.7, 4.8 Hz, 1H), 1.51 (s, 3H), 1.43 (s, 3H), 1.36 (s, 3H), 1.32 (s, 3H).



# 3.10.5 3-Amino-3-*N*-Boc-3-deoxy-1,2;5,6-di-*O*-isopropylidene-β-D-glucofuranose (6)

Palladium (10% on carbon, 0.25 g) was added to a stirring solution of compound **5** (2.46 g, 8.63 mmol) and di-*t*-butyl dicarbonate (2.26 g, 10.4 mmol) in methanol (20 mL). The reaction vessel was sealed tightly and the atmosphere aspirated and replaced with hydrogen gas three times. The reaction was allowed to proceed for 2 days under an atmosphere of hydrogen gas. The solution was filtered through celite twice in order to remove solid palladium on carbon before the solvent was removed under reduced pressure. The resulting clear oil was loaded onto a silica gel column and eluted with a 4:1 solution of hexanes and ethyl acetate. Fractions containing compound **6** (2.90 g, 98%). The <sup>1</sup>H NMR spectrum was consistent with an *N*-boc protected diisopropylidene of glucose.

<sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>): δ 5.79 (d, *J* = 3.0 Hz, 1H), 5.31 (s, 1H), 4.55 (s, 1H), 4.27 (s, 1H), 4.16-3.99 (m, 3H), 3.80 (s, 1H), 1.44 (s, 3H), 1.38 (m, 12H), 1.29 (s, 3H), 1.24 (s, 3H).



# 3.10.6 Kanosamine·HCl (7)

Compound 6 (2.41 g, 7.02 mmol) was dissolved in an aqueous solution of hydrochloric acid (2M, 100 mL) and stirred overnight. Removal of the solvent under reduced pressure was aided by the addition of toluene as a coevaporant three times. The resulting oil was placed under high vacuum overnight to yield wet compound 7 as a sticky gum (1.67 g). Despite extended

drying under high vacuum, not all of the water could be removed. The <sup>1</sup>H NMR spectrum was consistent with that reported.<sup>112</sup>

<sup>1</sup>H NMR (500 MHz, D<sub>2</sub>O): δ 5.21 (d, J = 3.5 Hz, 1H), 4.67 (d, J = 7.7 Hz, 1H), 3.88-3.77 (m, 4H), 3.76-3.66 (m, 4H), 3.61 (td, J = 9.9, 4.1 Hz, 2H), 3.51 (ddd, J = 9.5, 5.3, 2.1 Hz, 1H), 3.42-3.32 (m, 2H), 3.18 (t, J = 10.3, 1H).

<sup>13</sup>C NMR (500 MHz, D<sub>2</sub>O): δ 96.1 (s), 91.2 (s), 76.8 (s), 71.2 (s), 70.5 (s), 68.1 (s), 66.0 (s), 65.9 (s), 60.2 (s), 60.0 (s), 57.9 (s), 55.1 (s).



#### 3.11 Synthesis of Neotrehalose

Neotrehalose was synthesized using the method described by Ronnow et al.<sup>76</sup>

# **3.11.1** 2,3,4,6-Tetra-*O*-acetyl-1-*S*-thiophenol-β-D-glucopyranose (9)

A process similar to that reported was followed, except the final compound was isolated by crystallization instead of column chromatography.<sup>99</sup> Glucose (4.00 g, 22.2 mmol) was suspended in acetic anhydride (10.9 mL, 113 mmol) under an atmosphere of argon. The stirring suspension was cooled to 0 °C before the addition of one equivalent of hydrobromic acid (33% in acetic acid) (3.9 mL, 22.2 mmol). The solution was allowed to warm to room temperature and react for 20 minutes. The solution was cooled to 0 °C before the dropwise addition of two additional equivalents of hydrobromic acid (33% in acetic acid) (7.8 mL, 44.4 mmol). The reaction mixture was allowed to warm to room temperature and react three hours, resulting in the complete disappearance of solid glucose. The solvent was removed under

reduced pressure, thrice aided by the addition of toluene as a coevaporant. Thiophenol (3.40 mL, 33.3 mmol), TBAHS (1.12 g, 3.3 mmol), an aqueous solution of sodium carbonate (1 M, 117 mL), and DCM (84 mL) were added in series. The reaction mixture was stirred vigorously for one hour and turned from clear, to black, to amber in that time. The reaction mixture was diluted with DCM (100 mL) and washed once with an aqueous solution of sodium hydroxide (1M) and twice with deionized water. The organic phase was dried with anhydrous magnesium sulfate, and then vacuum filtered. The solvent was removed in vacuo, leaving a pungent, yellow oil that formed oily, yellow crystals overnight. These were recrystallized from a hot solution of hexanes and ethyl acetate. Vacuum filtration resulted in pure compound **9** as white, fluffy crystals (3.34 g, 34%).

<sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  7.52-7.48 (m, 2H), 7.35-7.30 (m, 3H), 5.23 (t, *J* = 9.4 Hz, 1H), 5.05 (t, *J* = 9.8 Hz, 1H), 4.98 (t, *J* = 9.7 Hz, 1H), 4.77 (d, *J* = 10.1 Hz, 1H), 4.25-4.16 (m, 2H), 3.73 (ddd, *J* = 9.5, 4.7, 2.2 Hz, 1H), 2.09 (s, 3H), 2.09 (s, 3H), 2.02 (s, 3H), 1.99 (s, 3H).



**3.11.2** 2,3,4,6-Tetra-*O*-benzyl-1-*S*-thiophenol-β-D-glucopyranose (10)

Compound **9** (3.34 g, 7.59 mmol) was dissolved in boiling methanol (30 mL). Sodium metal (0.018 g, 0.76 mmol) was added and the solution allowed to reflux for ten minutes. The reaction mixture was cooled and the solvent removed under reduced pressure, aided by the addition of toluene three times as a coevaporant. The resulting solid was placed under high vacuum for 20 minutes to remove any additional solvent. The intermediate was dissolved in dry DMF (80 mL) under an atmosphere of argon. A 60% w/w dispersion of sodium hydride in mineral oil (1.46 g, 36.4 mmol) was added, causing the reaction mixture to turn cloudy. The reaction mixture was allowed to stir for 30 minutes before the addition of TBAI (1.12 g, 3.03 mmol) and benzyl bromide (4.33 mL, 36.4 mmol). The reaction mixture was allowed to stir overnight and had cleared by morning. TLC showed an incomplete reaction. Additional sodium

hydride (0.73 g, 18.2 mmol) was added; the reaction mixture turned cloudy. The suspension was allowed to stir for 30 min before the addition of benzyl bromide (2.16 mL, 18.2 mmol). The reaction mixture was allowed to stir overnight. The solution was concentrated to a thick oil under reduced pressure. The oil was diluted with ethyl acetate and washed twice with deionized water and once with brine. The organic phase was dried with anhydrous magnesium sulfate, and then vacuum filtered. The solvent was removed under reduced pressure, yielding crude crystals, which were recrystallized from a solution of hot hexanes and ethyl acetate. After vacuum filtration, compound **10** was recovered as yellowish, fluffy crystals (3.73 g, 78%). The mother liquor was reduced in vacuo to leave a yellow oil. The addition of hexanes precipitated additional compound which was recrystallized from hot hexanes to yield small, sharp, yellowish crystals (0.19 g, 4%). In total, 3.92 g of compound **10** were recovered (82%). The <sup>1</sup>H NMR agreed with that reported.<sup>100</sup>

<sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  7.63-7.57 (m, 2H), 7.43-7.17 (m, 23H), 4.98-4.81 (m, 4H), 4.74 (d, *J* = 10.3 Hz, 1H), 4.68 (d, *J* = 9.8 Hz, 1H), 4.65-4.53 (m, 3H), 4.80 (dd, *J* = 10.7, 1.5 Hz, 1H), 3.77-3.69 (m, 2H), 3.66 (t, *J* = 9.3 Hz, 1H), 3.56-3.49 (m, 2H).



10

#### 3.11.3 2,3,4,6-Tetra-*O*-benzyl-D-glucose (11)

A procedure similar to that reported was followed.<sup>100</sup> Compound **10** (2.44 g, 3.85 mmol) was dissolved in a 9:1 solution of acetone and water (40 mL). N-Bromosuccinimide (1.37 g, 7.70 mmol) was added and the solution allowed to stir for five minutes. The reaction was quenched by the addition of sodium bicarbonate (2.78 g). The solution was stirred for ten minutes before removal of the solvent under reduced pressure. The resulting solid was dissolved in ethyl acetate and washed with deionized water until the pH of the washes was neutral. The organic phase was washed with brine, dried with anhydrous magnesium sulfate, and then vacuum filtered. Cottony crystals formed in the vacuum flask. These were recrystallized from a

hot solution of hexanes and ethyl acetate to yield extremely fluffy, white crystals (1.01 g, 49%). The mother liquor was concentrated under reduced pressure to yield more crystals. These were recrystallized as before to yield very slightly yellow, fluffy crystals (0.40 g, 19%). The mother liquor was recrystallized once more to yield identical, light yellow crystals (0.48 g, 23%). In total, 1.89 g of compound **11** were recovered (91%). The <sup>1</sup>H NMR spectrum agreed with that reported.<sup>100</sup>

<sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  7.39-7.26 (m, 18H), 7.17-7.14 (m, 2H), 5.23 (dd, J = 2.8, 2.8 Hz, 1H), 4.94 (d, J = 10.9 Hz, 1H), 4.87-4.75 (m, 3H), 4.69 (d, J = 11.8 Hz, 1H), 4.60 (d, J = 12.2 Hz, 1H), 4.52-4.47 (m, 2H), 4.03 (ddd, J = 10.2, 3.5, 2.1 Hz, 1H), 3.96 (t, J = 9.3 Hz, 1H), 3.71 (dd, J = 10.6, 3.7 Hz, 1H), 3.67-3.57 (m, 3H), 2.88-2.82 (m, 1H).



#### 3.11.4 2,3,4,6-Tetra-O-benzyl-1-O-(trichloroacetamidyl)-D-glucopyranose (12)

Compound 11 was synthesized using the method reported for the synthesis of 2,3,4-tri-*O*-acetyl-l-fucopyranose.<sup>57</sup> Compound **11** (0.50 g, 0.92 mmol) was dissolved in a solution of freshly distilled DCM (5 mL) and trichloroacetonitrile (0.99 mL, 9.20 mmol). A 60% w/w dispersion of sodium hydride in mineral oil (59 mg, 1.5 mmol) was added and the solution allowed to stir overnight under an atmosphere of argon. The reaction mixture was filtered through celite and loaded onto a silica gel column. Compound **12** was eluted with a 3:1 solution of hexanes and ethyl acetate. The solvent was removed under reduced pressure to yield compound **12** as a yellow oil (0.55 g, 87%). Compound **12** was isolated as primarily the alpha anomer with an  $\alpha$ : $\beta$ -ratio of roughly 12:1.

<sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  8.58 (s, 1H), 7.35-7.27 (m, 18H), 7.19-7.11 (m, 2H), 6.53 (d, J = 3.4 Hz, 1H), 4.96 (d, J = 10.9 Hz, 1H), 4.87-4.80 (m, 2H), 4.75 (d, J = 11.7 Hz, 1H), 4.68 (d, J = 11.7 Hz, 1H), 4.61 (d, J = 12.0 Hz, 1H), 4.52 (d, J = 10.6 Hz, 1H), 4.47 (d, J = 12.1 Hz, 1H),

4.05 (t, *J* = 9.4 Hz, 1H), 3.98 (ddd, *J* = 10.2, 3.5, 2.1 Hz, 1H), 3.82-3.74 (m, 3H), 3.67 (dd, *J* = 10.9, 1.7 Hz, 1H).



### 3.11.5 2,2',3,3',4,4',6,6'-Octa-O-benzyl-*neo*-trehalose (13)

The procedure reported by Ronnow et al. was followed.<sup>76</sup> Compound **11** (0.336 g, 0.621 mmol) was dissolved in freshly distilled DCM (17 mL) along with 1.3 equivalents of freshly prepared compound **12** (0.554 g, 0.808 mmol). Silver triflate (0.207 g, 0.805 mmol) was added. The reaction mixture was allowed to stir in the dark under an atmosphere of argon for three days. The solvent was removed under reduced pressure and the resulting oil was loaded onto a long column of silica gel. The oil was eluted with a 19:1 solution of toluene and ethyl acetate. Fractions containing the starting alcohol 11 were pooled and the solvent removed in vacuo to yield a white solid (0.114 g, 15%). Fractions containing mainly 13 were pooled and the solvent removed under reduced pressure to yield a white, crystalline solid (0.153 g). The remaining fractions containing an unseparable mixture of 13, 14, and an unidentified compound were pooled. The solvent was removed in vacuo, and the resulting oil was loaded onto another column of silica gel and eluted with the same solvent system as above. Fractions containing mostly pure 13 were pooled and the solvent removed in vacuo to yield a light yellow solid (27 mg). The remaining 13 was not separable from 14 by silica gel column chromatography, even after repeated attempts. Both fractions of compound 13 were pooled and recrystallized from hot chloroform, yielding the unidentified, primary contaminant (42 mg). The solvent from the mother liquor was removed in vacuo and the resulting solid recrystallized again to yield compound **13** as a white, crystalline precipitate (96 mg, 15%).

<sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  7.35-7.13 (m, 40H), 5.17 (d, J = 3.3 Hz, 1H), 5.11 (d, J = 11.5 Hz, 1H), 4.97 (d, J = 10.9 Hz, 1H), 4.92 (d, J = 10.9 Hz, 1H), 4.88-4.74 (m, 5H), 4.67 (q, J = 12.2 Hz, 2H), 4.59 (d, J = 7.8 Hz, 1H), 4.55 (d, J = 10.8 Hz, 1H), 4.48 (m, 4H), 4.29 (d, J = 12.1

Hz, 1H), 4.16 (ddd, *J* = 10.2, 4.3, 2.1 Hz, 1H), 4.09 (t, *J* = 9.4 Hz, 1H), 3.76 (t, *J* = 9.6 Hz, 1H), 3.68-3.56 (m, 6H), 3.55-3.47 (m, 2H), 3.45 (ddd, *J* = 9.6, 4.1, 2.0 Hz, 1H).



# 3.11.6 Neotrehalose (15)

Compound **13** (20 mg) was dissolved in a suspension of palladium (10% on carbon, 2 mg) in methanol (2 mL). The atmosphere was aspirated and replaced with hydrogen gas three times. The reaction mixture was allowed to stir under an atmosphere of hydrogen gas overnight, and then filtered through celite. The solvent was removed in vacuo to yield compound **15** as a clear syrup (7 mg, 100%).

<sup>1</sup>H NMR (500 MHz, D<sub>2</sub>O):  $\delta$  5.18 (d, J = 3.1 Hz, 1H), 4.59 (d, J = 7.9 Hz, 1H), 3.92-3.61 (m, 6H), 3.58-3.27 (m, 6H).



# **3.12** Synthesis of 3-[<sup>2</sup>H]-D-glucose

The synthesis of 3-[<sup>2</sup>H]-D-glucose was performed according to the published procedure, with the substitution of a trifyl group instead of a tosyl group for nucleophilic displacement.<sup>102</sup>

# **3.12.1 3**-[<sup>2</sup>H]-1,2;5,6-Di-*O*-isopropylidene-β-D-allofuranose (16)

Compound 2 (0.500 g, 1.94 mmol) was dissolved in a 56% solution of ethanol in deionized water (2.2 mL). A solution of sodium borodeuteride (0.10 g, 2.4 mmol) in deionized water (2.5 mL) was added slowly. The reaction mixture was allowed to stir overnight. The ethanol was removed under reduced pressure and the remaining water extracted thrice with ethyl acetate. The organic phase was dried with anhydrous magnesium sulfate, and then vacuum filtered. The solvent was removed in vacuo, leaving compound **16** as a white, crystalline solid (0.431 g, 85%). The <sup>1</sup>H NMR spectrum was identical to that of the diisopropylidene of allose with the omission of the signal for the hydrogen at C-3 and the absence of its corresponding coupling.

<sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  5.82 (d, J = 3.8 Hz, 1H), 4.62 (d, J = 3.8 Hz, 1H), 4.32 (dt, J = 6.5, 5.0 Hz, 1H), 4.09 (dd, J = 8.4, 6.7 Hz, 1H), 4.02 (dd, J = 8.4, 6.7 Hz, 1H), 3.82 (d, J = 4.6 Hz, 1H), 2.52 (s, 1H), 1.58 (s, 3H), 1.47 (s, 3H), 1.39 (s, 3H), 1.38 (s, 3H).



# **3.12.2 3**-[<sup>2</sup>H]-1,2;5,6-Di-*O*-isopropylidene-3-*O*-triflyl-β-D-allofuranose (17)

Compound **16** (2.50 g, 9.59 mmol) was dissolved in freshly distilled DCM (20 mL). Sodium hydroxide dried pyridine (1.6 mL) was added and the reaction mixture cooled to -30 °C. Triflic anhydride (1.93 mL, 11.5 mmol) was added dropwise over ten minutes and the reaction mixture allowed to stir an additional 30 minutes at -30 °C. The reaction was quenched by the addition of methanol (1 mL). The reaction mixture was raised to room temperature and washed twice with deionized water. The organic phase was dried with anhydrous sodium sulfate, and then vacuum filtered. The solvent was removed at room temperature under reduced pressure, yielding an oily, yellow solid. Compound **17** was used immediately in the next step.


## **3.12.3 3**-[<sup>2</sup>H]-**3**-*O*-benzoyl-**1**,2;**5**,6-Di-*O*-isopropylidene-β-D-glucofuranose (18)

Freshly prepared compound **17** was dissolved in dry DMF (100 mL) under an atmosphere of argon. Oven-dried (24 hours at 110 °C) sodium benzoate (6.96 g, 48.0 mmol) was added. The solution was stirred vigorously at 60 °C for eight hours. The sodium benzoate swelled up so that the reaction mixture became nearly solid. DMF was partially removed at 80 °C under reduced pressure. The remaining white solid was dissolved in chloroform and deionized water. The organic phase was washed seven times with deionized water, once with an ice-cold, aqueous solution of sodium hydroxide (1M), once with a saturated sodium bicarbonate solution, and once with brine. The organic phase was dried with anhydrous magnesium sulfate, and then vacuum filtered. The solvent was removed in vacuo to yield a thick, yellow oil. Compound **18** was not purified further before deprotection of the benzoyl group.



## **3.12.4 3**-[<sup>2</sup>H]-1,2;5,6-Di-*O*-isopropylidene-β-D-glucofuranose (19)

Compound **18** was dissolved in refluxing methanol (40 mL). Sodium metal (50 mg, 2 mmol) was added and the reaction mixture allowed to reflux for ten minutes. The solvent was removed under reduced pressure, yielding a yellow solid. This solid was dissolved in an aqueous solution of sodium hydroxide (1M) and extracted four times with chloroform. The organic phase was dried with anhydrous magnesium sulfate, and then vacuum filtered. The solvent was removed under reduced pressure to yield a yellow solid, which was recrystallized from hot ligroin. The resulting white crystals smelled vaguely of benzoic acid. They were dissolved in an

aqueous solution of sodium hydroxide (1M) and extracted four times with chloroform. The organic phase was dried with anhydrous magnesium sulfate, and then vacuum filtered. The solvent was removed under reduced pressure to yield a white solid, which was recrystallized from hot ligroin. Compound **19** was obtained as white, long, fluffy, crystals (2.18g, 87% from compound **16**). The <sup>1</sup>H NMR spectrum was identical to that of the diisopropylidene of glucose with the omission of the signal for the hydrogen at C-3 and the absence of its corresponding coupling.

<sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  5.96 (d, J = 3.6 Hz, 1H), 5.54 (d, J = 3.6 Hz, 1H), 4.35 (dt, J = 7.3, 5.9 Hz, 1H), 4.18 (dd, J = 8.6, 6.2 Hz, 1H), 4.08 (d, J = 7.5 Hz, 1H), 3.98 (dd, J = 8.6, 5.4 Hz, 1H), 2.46 (s, 1H), 1.50 (s, 3H), 1.45 (s, 3H), 1.32 (s, 3H), 1.38 (s, 3H).



#### 3.12.5 3-[<sup>2</sup>H]-D-glucose (20)

Compound **19** (1.00 g, 3.85 mmol) was dissolved in a stirring suspension of acid Dowex 50X-80 (0.500g) in deionized water (50 mL). The reaction mixture was heated at 80 °C for 24 hours. The Dowex was removed by vacuum filtration and the water removed in vacuo over two days to yield compound **20** as a yellowish glass (0.703g). Some water remained, even after additional drying time. The  $\alpha$ : $\beta$  ratio was 1:1.3. Mp = 25-36 °C. The <sup>1</sup>H NMR spectrum was identical to that of glucose with the omission of the hydrogen attached to C-3 and the absence of its corresponding coupling. The <sup>13</sup>C NMR spectrum agreed with that reported where no coupling between the deuteron and C-3 was observed.<sup>102</sup>

<sup>1</sup>H NMR (500 MHz, D<sub>2</sub>O): δ 5.19 (d, J = 3.7 Hz, 1H), 4.60 (d, J = 8.0 Hz, 1.3H), 3.85 (dd, J = 12.3, 2.0 Hz, 1.3H), 3.83-3.76 (m, 2H), 3.76-3.64 (m, 2.3H), 3.49 (d, J = 3.7 Hz, 1H), 3.45-3.39 (m, 1.3H), 3.37 (d, J = 4.9 Hz, 1.3H), 3.35 (d, J = 5.2 Hz, 1H), 3.20 (d, J = 7.9 Hz, 1.3H).

<sup>13</sup>C NMR (500 MHz, D<sub>2</sub>O): δ 95.9 (s), 92.1 (s), 75.9 (s), 74.0 (s), 71.4 (s), 71.4 (s), 69.5 (s), 69.5 (s), 60.7 (s), 60.6 (s).



### 3.13 Synthesis of 3-Deoxy-3-fluoro-D-glucose

The route directed towards the synthesis of 3-deoxy-3-fluoro-D-glucose was similar to that published.<sup>113</sup>

### 3.13.1 1,2;5,6-Di-*O*-isopropylidene-3-*O*-triflyl-β-D-allofuranose (21)

Compound **3** (0.458 g, 1.76 mmol) was dissolved in freshly distilled DCM (5 mL). Sodium hydroxide dried pyridine (0.30 mL) was added and the reaction mixture cooled to -30 °C. Triflic anhydride (0.35 mL, 2.1 mmol) was added dropwise over ten minutes and the reaction mixture allowed to stir an additional 30 minutes at -30 °C. The reaction was quenched by the addition of methanol (1 mL). The reaction mixture was raised to room temperature and washed twice with deionized water. The organic phase was dried with anhydrous sodium sulfate, and then vacuum filtered. The solvent was removed at room temperature under reduced pressure, yielding an oily, yellow solid. Compound **21** was used immediately in the next step.



#### **3.13.2 3-Deoxy-3-fluoro-1,2;5,6-Di**-*O*-isopropylidene-β-D-glucofuranose (22)

Freshly prepared compound **21** was dissolved in dry DMF (20 mL) under an atmosphere of argon. Oven-dried (24 hours at 110  $^{\circ}$ C) cesium fluoride (0.80 g, 5.3 mmol) was

added and the reaction mixture stirred vigorously at 60 °C for five hours. The reaction mixture had turned slightly orange by this point. The bulk of the DMF was removed at 80 °C under reduced pressure. The remaining sludge was partitioned between chloroform and an aqueous solution of hydrochloric acid (0.1M). The aqueous phase was extracted thrice with chloroform. The organic phase was dried with anhydrous magnesium sulfate, and then vacuum filtered. The solvent was removed in vacuo to yield an orange oil. This oil was loaded onto a short column of silica gel and eluted with a 4:1 solution of hexanes and ethyl acetate. Fractions containing compound **22** were pooled and the solvent removed in vacuo to yield a colourless oil (0.368 g, 80%).

<sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  5.95 (d, *J* = 3.7 Hz, 1H), 5.01 (dd, *J* = 49.8, 2.0 Hz, 1H), 4.70 (dd, *J* = 10.7, 3.7 Hz, 1H), 4.29 (dt, *J* = 8.1, 5.7 Hz, 1H), 4.17-4.06 (m, 2H), 4.04 (dd, *J* = 8.8, 4.7 Hz, 1H), 1.50 (s, 3H), 1.45 (s, 3H), 1.37 (s, 3H), 1.33 (s, 3H).



# **Chapter 4: Conclusions**

- *ntdA*, *ntdB*, and *ntdC* were successfully subcloned into pET-28b
- NtdC was expressed in Bl21 (DE3) cells and purified to homogeneity using nickelaffinity chromatography
- NtdC was identified as an NAD<sup>+</sup>-dependent glucose 6-phosphate dehydrogenase
- NtdC was found to have sub-millimolar affinity for its substrates and a specificity constant for its sugar substrate similar to other sugar dehydrogenases
- NtdC likely obeys a ternary complex mechanism, but investigation of the reverse reaction will be necessary to confirm this
- The greatly impaired nature of D176 and H180 mutants support the presence of a catalytic dyad created by these two residues
- Kanosamine has been synthesized
- Neotrehalose has been successfully synthesized
- 3-[<sup>2</sup>H]-D-glucose has been synthesized in high yield, but does not possess the properties of D-glucose
- The diisopropylidene of 3-deoxy-3-fluoro-D-glucose has been synthesized by displacement of a trifyl group with cesium fluoride with a yield comparable to that reported in the best syntheses using DAST.

Ochi *et al.* believed that *ntdA*, *ntdB*, and *ntdC* encoded for the only structural genes involved in NTD biosynthesis.<sup>3</sup> Scheme 1.2 displays the simplest biosynthetic pathway possible if this were the case. However, this route seems too simple given NtdC's identity as a glucose 6phosphate dehydrogenase. Most biological glycosylation reactions require that the glycosyl donor be activated as an NDP-sugar. However, unlike RifL, which is a component of kanosamine biosynthesis in *A. mediterranei*, NtdC does not appear to accept NDP-sugars. At the same time, NtdA is structurally similar to members of the VI<sub>β</sub> subfamily of aminotransferases, which recognize 3-oxo-sugar pyranoses. This suggests that NtdC is only responsible for synthesizing one of the kanosamine subunits of the final NTD molecule. Additionally, NtdB appears to be a phosphatase. It is improbable that it should function as novel type of glycosyltranferase, which would be required for the pathway in Scheme 1.2 to be correct. Instead, it most likely functions to hydrolyze the phosphate introduced by G6P at some point during NTD's synthesis. The resulting possible pathways (Scheme 4.1) are much more complicated than that depicted in Scheme 1.2, requiring at least three additional enzymes to complete. This is exciting, as *E. coli* is not known to possess either kanosamine or NTD biosynthetic pathways. It may be possible that some non-specific reactions are capable of complete pathway for the synthesis of NTD. In either case, this would mean that *E. coli* possesses an enzyme capable of forming an  $\alpha$ , $\beta$ -glycosidic linkage. Confirming the functions of NtdA and NtdB through enzymatic assays will help validate one of the pathways presented below.



Scheme 4.1: Currently proposed biosynthetic pathways for NTD. Dashed lines represent proposed steps.

# **Chapter 5: References**

- (1) Numata, K.; Satoh, F.; Hatori, M.; Miyaki, T.; Kawaguchi, H. J. Antibiot **1986**, *39*, 1346-1348.
- (2) Numata, K.; Yamamoto, H.; Hatori, M.; Miyaki, T.; Kawaguchi, H. J. Antibiot **1986**, *39*, 994-1000.
- (3) Inaoka, T.; Takahashi, K.; Yada, H.; Yoshida, M.; Ochi, K. J. Biol. Chem. 2004, 279, 3885-3892.
- (4) Schito, G. C. Clin. Microbiol. Infect. 2006, 12 (Suppl. 1), 3-8.
- (5) Hiramatsu, K.; Hanaki, H.; Ino, T.; Yabuta, K.; Oguri, T.; Tenover, F. C. J. Antimicrob. *Chemother.* **1997**, *40*, 135-136.
- (6) Sievert, D. M.; Boulton, M. L.; Stoltman, G.; Johnson, D.; Stobierski, M. G.; Downes, F. P.; Somsel, P. A.; Rudrik, J. T.; Brown, W.; Hafeez, W.; Lundstrom, T.; Flanagan, E.; Johnson, R.; Mitchell, J.; Chang, S. *MMWR* **2002**, *51*, 565-567.
- (7) Tsuno, T.; Ikeda, C.; Numata, K.; Tomita, K.; Konishi, M.; Kawaguchi, H. J. Antibiot **1986**, *39*, 1001-1003.
- (8) Inaoka, T.; Ochi, K. J. Bacteriol. 2007, 189, 65-75.
- (9) Inaoka, T.; Satomura, T.; Fujita, Y.; Ochi, K. FEMS Microbiol. Lett. 2009, 291, 151.
- (10) Paulsen, I. T.; Chauvaux, S.; Choi, P.; Saier, M. H. J. Bacteriol 1998, 180, 498-504.
- (11) Cron, M. J.; Fardig, O. B.; Johnson, D. L.; Schmitz, H.; Whitehead, D. F.; Hooper, I. R.; Lemieux, R. U. J. Am. Chem. Soc. **1958**, 80, 2342.
- (12) Umezawa, S.; Umino, K.; Shibahara, S.; Omoto, S. Bull. Chem. Soc. Jpn. **1967**, 40, 2419-2421.
- (13) Fusetani, N.; Ejima, D.; Matsunaga, S.; Hashimoto, K.; Itagaki, K.; Akagi, Y.; Taga, N.; Suzuki, K. *Experientia* **1987**, *43*, 464-465.
- (14) Milner, J. L.; Silo-Suh, L.; Lee, J. C.; Hee, H.; Clardy, J.; Handelsman, J. *Appl. Environ. Microbiol.* **1996**, *62*, 3061-3065.
- (15) Guo, J.; Frost, J. W. Org. Lett. 2004, 6, 1585-1588.

- (16) Floss, H. G.; Yu, T. Chem. Rev. 2005, 105, 621-632.
- (17) Altschul, S. F.; Madden, T. L.; Schaffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. *Nucleic Acids Res.* **1997**, *25*, 3389-3402.
- (18) Voet, D.; Voet, J. D.; Pratt, C. W. In *Amino Acid Deamination;* John Wiley & Sons, Inc.: USA, 1999; pp 615-619.
- (19) van Straaten, K. E.; Langill, D. M.; Palmer, D. R. J.; Sanders, D. A. R. Acta Cryst. 2009, *F65*, 426-429.
- (20) van Straaten, K. E.; Langill, D. M.; Palmer, D. R. J.; Sanders, D. A. R. Unpublished Results, .
- (21) Thompson, J. D.; Higgins, D. G.; Gibson, T. J. Nucleic Acids Res. 1994, 22, 4673-4680.
- (22) Henikoff, S.; Henikoff, J. G. Biochemistry 1992, 89, 10915-10919.
- (23) Gouet, P.; Courcelle, E.; Stuart, D. I.; Metoz, F. Bioinformatics 1999, 15, 305-308.
- (24) Burgie, E. S.; Thorden, J. B.; Holden, H. M. Protein Sci. 2007, 16, 887-896.
- (25) Thoden, J. B.; Schaffer, C.; Messner, P.; Holden, H. M. Biochemistry 2009, 48, 1533-1561.
- (26) Hwang, B.; Lee, H.; Yang, Y.; Joo, H.; Kim, B. Chem. Biol. 2004, 11, 915-925.
- (27) Weinreich, M.; Liang, C.; Stillman, B. Proc. Natl. Acad. Sci. USA 1999, 96, 441-446.
- (28) DeLano, W. L. In *The PyMOL Molecular Graphics System;* DeLano Scientific: Palo Alto, CA, USA, 2000; .
- (29) Allen, K. N.; Dunaway-Mariano, D. Trends Biochem. Sci. 2004, 29, 495-503.
- (30) Fieulaine, S.; Lunn, J. E.; Borel, F.; Ferrer, J. Plant Cell 2005, 17, 2045-2058.
- (31) Nocek, B.; Stein, A.; Wu, R.; Jedrzejczak, R.; Joachimiak, A. Unpublished Results 2009, .
- (32) Collet, J.; Stroobant, V.; Pirard, M.; Delpierre, G.; Van Schaftingen, E. J. Biol. Chem. **1988**, 273, 14107-14112.
- (33) Wang, W.; Kim, R.; Jancarik, J.; Yokota, H.; Kim, S. Structure 2001, 9, 65-71.
- (34) Zhang, G.; Morais, M. C.; Dai, J.; Zhang, W.; Dunaway-Mariano, D.; Allen, K. N. *Biochemistry* **2004**, *43*, 4990-4997.

- (35) Rao, K. N.; Kumaran, D.; Seetharaman, J.; Bonanno, J. B.; Burley, S. K.; Swaminathan, S. *Protein Sci.* **2006**, *15*, 1735-1744.
- (36) Zhao, W., et al Cell Res. 2010, .
- (37) van Straaten, K. E.; Hoffort, A.; Palmer, D. R. J.; Sanders, D. A. R. Acta Cryst. 2008, F64, 98-101.
- (38) van Straaten, K. E.; Hoffort, A.; Palmer, D. R. J.; Sanders, D. A. R. Unpublished Results, .
- (39) Dambe, T. R.; Kuehn, A. M.; Brossette, T.; Giffhorn, F.; Scheidig, A. J. *Biochemistry* **2006**, *45*, 10030-10042.
- (40) Kim, Y.; Arora, M.; Straza, M.; Joachimiak, A. Unpublished Results, .
- (41) Rao, S. T.; Rossmann, M. G. J. Mol. Biol. 1973, 76, 241-250.
- (42) Bellamacina, C. R. FASEB J. 1996, 10, 1257-1269.
- (43) Hanukoglu, I.; Gutfinger, T. Eur. J. Biochem. 1989, 180, 179-484.
- (44) Fujita, Y.; Shindo, K.; Miwa, Y.; Yoshida, K. Gene 1991, 108, 121-125.
- (45) Wiergert, T.; Sahm, H.; Sprenger, G. A. J. Biol. Chem. 1997, 272, 13126-13133.
- (46) Kingston, R. L.; Scopes, R. K.; Baker, E. N. Structure 1996, 2, 1413-1428.
- (47) Daniellou, R.; Zheng, H.; Langill, D. M.; Sanders, D. A. R.; Palmer, D. R. J. *Biochemistry* **2007**, *46*, 7469-7477.
- (48) Nelson, D. L.; Cox, M. M. In *Biosynthesis of glycogen, starch, sucrose, and other carbohydrates;* Morgan, R., Strange, L. and Neal, V., Eds.; Lehninger Principles of Biochemistry; Worth Publishers: 42 Madison Avenue, New York, New York, 2000; pp 735-738.
- (49) Fischer, E. Ber. Dtsch. Chem. Ges. 1893, 26, 2400-2412.
- (50) Koenigs, W.; Knorr, E. Ber. Dtsch. Chem. Ges. 1901, 34, 957-981.
- (51) Alonen, A.; Gartman, M.; Aitio, O.; Finel, M.; Yli-Kauhaluoma, J.; Kostiainen, R. *Eur. J. Pharm. Sci.* **2009**, *37*, 581-587.
- (52) Goldschmid, H. R.; Perlin, A. S. Can. J. Chem. 1961, 39, 2025-2034.
- (53) Helferich, B.; Zirner, J. Chem. Ber. 1962, 95, 2604.

- (54) Kochetkov, N. K.; Bochkov, A. F.; Sokolavskaya, T. A.; Snyatkova, V. J. *Carbohydr. Res.* **1971**, *16*, 17-27.
- (55) Garegg, P. J. Adv. Carbohydr. Chem. Biochem. 1997, 52, 179-205.
- (56) Ramakrishnan, A.; Pornsuriyasak, P.; Demchenko, A. V. J. Carb. Chem. 2005, 24, 649-663.
- (57) Schmidt, R. R.; Michel, J. Angew. Chem. Int. Ed. 1980, 19, 731-732.
- (58) Karasawa, I.; Onishi, R. Nippon Nogei Kagaku Kaishi 1961, 35, 819.
- (59) Pougny, J.; Sinay, P. Tetrahedron Lett. 1976, 17, 4073-4076.
- (60) Hashimoto, S.; Honda, T.; Ikegami, S. J. Chem. Soc. , Chem. Commun. 1989, 1989, 685-687.
- (61) Fraser-Reid, B.; Konradsson, P.; Mootoo, D. R.; Udodong, U. J. Chem. Soc., Chem. Commun. **1988**, 1988, 823-825.
- (62) Nguyen, H. M.; Poole, J. L.; Gin, D. Y. Angew. Chem. Int. Ed. 2001, 40, 414-417.
- (63) Toshima, K.; and Tatsuta, K. Chem. Rev. 1993, 93, 1503-1531.
- (64) Demchenko, A. V. Curr. Org. Chem. 2003, 7, 35-79.
- (65) Low, N. H.; Sporns, P. J. Food Sci. 1988, 53, 558-561.
- (66) Baer, H. H.; Bell, A. J. Can. J. Chem. 1978, 56, 2872-2878.
- (67) Birch, G. G.; Lee, C. K. Carbohydr. Res. 1976, 49, 153-161.
- (68) Baer, H. H.; Bell, A. J. Carbohydr. Res. 1979, 75, 175-184.
- (69) Hui, Y.; Chang, C. T. Org. Lett. 2002, 4, 2245-2248.
- (70) Lemieux, R. U.; Huber, G. J. Am. Chem. Soc. 1953, 75, 4118.
- (71) Sharp, V. E.; Stacey, M. J. Chem. Soc. 1951, 1951, 285-288.
- (72) Demchenko, A. V. Synlett 2003, 2003, 1225-1240.
- (73) Lemieux, R. U.; Hendriks, K. B.; Stick, R. V.; James, K. J. Am. Chem. Soc. **1975**, 97, 4056-4062.
- (74) Hadd, M. J.; Gervay, J. Carbohydr. Res. 1999, 320, 61-69.

- (75) Lemieux, R. U.; Morgan, A. R. Can. J. Chem. 1965, 43, 2190-2197.
- (76) Ronnow, T. E. C. L.; Meldal, M.; Bock, K. Tetrahedron: Asymmetry 1994, 5, 2109-2122.
- (77) Dubendorff, J. W.; Studier, F. W. J. Mol. Biol. 1991, 219, 45-59.
- (78) Hochuli, E.; Bannwarth, W.; Dobeli, H.; Gentz, R.; Stuber, D. *Bio/Technology* **1988**, *6*, 1321-1325.
- (79) Vincze, T.; Posfai, J.; Roberts, R. J. Nucleic Acids Res. 2003, 31, 3688-3691.
- (80) Kunst, F., et al Nature 1997, 390, 249-256.
- (81) Ramaley, R.; Fujita, Y.; Freese, E. J. Biol. Chem. 1979, 254, 7684-7690.
- (82) Yip, V. L.; Varrot, A.; Davies, G. J.; Rajan, S. S.; Yang, X.; Thompson, J.; Anderson, W. F.; Withers, S. G. J. Am. Chem. Soc. 2004, 126, 8354-8355.
- (83) Yip, V. L.; Thompson, J.; Withers, S. G. Biochemistry 2007, 46, 9840-9852.
- (84) Bouma, C. L.; Reizer, J.; Reizer, A.; Robrish, S. A.; Thompson, J. J. Bacteriol. **1997**, *179*, 4129-4137.
- (85) Liljestrom, P. L.; Liljestrom, L. Nucleic Acids Res. 1987, 15, 2213-2220.
- (86) Hung, R.; Chien, H.; Lin, R.; Lin, C.; Vatsyayan, J.; Peng, H.; Chang, H. J. Biol. Chem. **2007**, 282, 17738-17748.
- (87) Arakawa, K.; Muller, R.; Mahmud, T.; Yu, T.; Floss, H. G. J. Am. Chem. Soc. 2002, 124, 10644-10645.
- (88) Zachariou, M.; Scopes, R. K. J. Bacteriol. 1986, 167, 863-869.
- (89) Kuhn, A.; Yu, S.; Giffhorn, F. *Applied and Environmental Microbiology* **2006**, *72*, 1248-1257.
- (90) Cornish-Bowden, A. In *Leonora: A program for robust regression of kinetic data;* Analysis of Enzyme Kinetic Data; Oxford University Press: New York, 1995; pp 133-189.
- (91) Cornish-Bowden, A. In *Reactions of more than one substrate;* Fundamentals of Enzyme Kinetics; Portland Press Ltd.: London, 1995; pp 129-158.
- (92) Cornish-Bowden, A. In *Control of enzyme activity;* Fundamentals of Enzyme Kinetics; Portland Press Ltd.: London, 1995; pp 203-237.

- (93) Cornish-Bowden, A. In *Introduction to enzyme kinetics;* Fundamentals of Enzyme Kinetics; Portland Press Ltd.: London, 1995; pp 19-53.
- (94) Zhang, X.; Chen, G.; Liu, W. FEMS Microbiol. Lett. 2009, 293, 214-219.
- (95) Cook, P. F.; Cleland, W. W. In *Initial velocity studies: Presence of added inhibitors;* Enzyme Kinetics and Mechanism; Garland Science Publishing: New York, 2007; pp 121-204.
- (96) Garegg, P. J.; Samuelsson, B. Carbohydr. Res. 1978, 67, 267-270.
- (97) Theander, O. Acta Chem. Scand. 1963, 17, 1751-1760.
- (98) Li, Z.; Gildersleeve, J. C. J. Am. Chem. Soc. 2006, 128, 11612-11619.
- (99) Kumar, R.; Tiwari, P.; Maulik, P. R.; Misra, A. K. Eur. J. Org. Chem. 2006, 2006, 74-75.
- (100) Damager, I.; Olsen, C. E.; Møller, B. L.; Motawia, M. S. *Carbohydr. Res.* **1999**, *320*, 19-30.
- (101) Cook, P. F.; Cleland, W. W. In *Initial velocity studies: Presence of added inhibitors;* Rogers, R. I., Scholl, S., Eds.; Enzyme Kinetics and Mechanism; Taylor and Francis Group, LLC.: New York, 2007; pp 121-204.
- (102) Koch, H. J.; Perlin, A. S. Carbohydr. Res. 1970, 15, 403-410.
- (103) Gorin, A. J. P. Can. J. Chem. 1974, 52, 458-461.
- (104) Longhi, G.; Zerbi, G.; Paterlini, G.; Ricard, L.; Abbate, S. *Carbohydr. Res.* **1987**, *161*, 1-22.
- (105) Johansson, I.; Lindberg, B. Carbohydr. Res. 1966, 1, 467-473.
- (106) Foster, A. B.; Hems, R. Carbohydr. Res. 1967, 5, 292-301.
- (107) Doboszewski, B.; Hay, G. W.; Szarek, W. A. Can. J. Chem. 1987, 65, 412-419.
- (108) Kovac, P.; Yeh, H. J. C.; Glaudemans, C. P. J. Carbohydr. Res. 1987, 169, 23-34.
- (109) Argentini, M.; Weinreich, R.; Oberti, R.; Ungaretti, L. J. Fluorine Chem. **1986**, *32*, 239-254.
- (110) Rice, P.; Longden, I.; Bleasby, A. Trends Genet. 2000, 16, 276-277.
- (111) PREMIER Biosoft International 2006, 1.0.

(112) Guo, J.; Frost, J. W. J. Am. Chem. Soc. 2002, 124, 10642-10643.

(113) Tewson, T. J.; Welch, M. J. J. Org. Chem. 1978, 43, 1090-1092.