Production of sialic acid analogs in engineered *E. coli*: Characterization of amino sugar recycling

LUIS ROBERTO VILLEGAS-PEÑARANDA

Thesis submitted in partial fulfillment of the requirements
for the Doctorate in Philosophy degree in Chemistry

Department of Chemistry and Biomolecular Sciences Ottawa-Carleton
Chemistry Institute
Faculty of Science
University of Ottawa

Abstract

This research focused on the study of the amino sugar recycling and sialic acid degradation pathway as a possible entry point for *N*-acyl glucosamines for the production of sialic acid analogs. Meeting this objective would allow the development of a bacterial strain capable of producing non-natural nonulosonic acids that could be used in the development of medicines, vaccines or useful compounds for the study of interactions between pathogenic organisms and their host.

The first step was to understand how *N*-acetyl-D-glucosamine-6-phosphate deacetylase reacts to different types of substrates in order to determine its tolerance to the size of acyl groups in acyl amino sugars. This was achieved by studying the enzymatic activity in an in vitro system. We determine that the enzyme has a preference for small and slightly bulky acyl groups. Then, an in silico docking modeling and an in vivo system experiment were carried out. These experiments allowed to confirm the previous results.

The second project was carried out due to the uncertainty of whether the kinase involved in the catabolic pathway would be able to phosphorylate the substrates. By quantifying residual ATP, the high specificity of *N*-acetyl-D-glucosamine kinase could be verified. This result led us to think about the design of an organic synthesis strategy that would allow the phosphorylation of glucosamine in carbon 6. A simple synthetic route was designed based on the protection of the two most reactive moieties of the amino sugars and the reactivity of the hydroxy group on carbon 6. However, we had problems with the purification step of the final product due to its high polarity.

The next stage of this investigation was to confirm the transformation of GlcNAc into ManNAc. For this, an NMR analysis was designed that would detect the presence of

both sugars in the reaction system. The epimerization of ManNAc to GlcNAc was detected successfully. Notwithstanding, the reverse reaction could not be detected. Based on the results obtained in the previous stage, we realized that an error was made in the epimerization reaction since we placed the wrong kinase because we did not take into account its substrate specificity.

Finally, we tried to produce sialic acid analogs in a fermentative system using different genetic variants of *Escherichia coli*. Two of the expected analogs, Neu5Pr and Neu5nBu, were obtained. In addition, NagA activity towards substrates with small acyl groups was confirmed.

Resumen

Esta investigación se enfocó en el estudio de la ruta de reciclaje de amino azúcares y degradación de ácido siálico como posible punto de entrada de *N*-acil glucosaminas para la producción de análogos del ácido siálico. El cumplir con este objetivo permitiría el desarrollo de un microorganismo capaz de producir ácidos nonulosónicos no naturales que podrían ser utilizados en el desarrollo de medicamentos, vacunas o compuestos útiles para el estudio de las interacciones entre organismos patógenos y su huésped.

El primer paso fue entender como reacciona NagA ante diferentes tipos de sustratos para poder determinar su rango de tolerancia al tamaño de los grupos acilo en el amino azúcar. Esto se realizó mediante el estudio de la actividad enzimática en un sistema in vitro que permitió determinar que dicha enzima tiene preferencia por grupos acilo pequeños y poco voluminosos. Luego se llevó a cabo un modelado de acoplamiento in silico y un sistema in vivo que permitió confirmar los resultados de la primera etapa.

El segundo proyecto se realizó debido a la incertidumbre de si la quinasa involucrada en la ruta catabólica sería capaz de fosforilar los sustratos. Mediante la cuantificación de ATP residual se pudo comprobar la alta especificidad de NanK y NagK. Este resultado nos llevó a pensar en el diseño de una estrategia de síntesis orgánica que permitiera la fosforilación de los azúcares en el carbono 6. Se diseñó una ruta sintética sencilla que se basa en la protección los dos grupos más reactivos de los amino azúcares utilizados y la reactividad del hidroxilo del carbono 6. Sin embargo, se presentaron problemas de purificación del producto final debido a la alta polaridad de este.

La siguiente etapa de esta investigación consistió en confirmar la transformación de GlcNAc en ManNAc. Para esto se diseñó un análisis mediante NMR que permitiría

detectar la presencia de ambos azúcares en el sistema de reacción. La epimerización de ManNAc a GlcNAc fue detectada exitosamente. Sin embargo, la reacción inversa no se pudo detectar. Basados en los resultados obtenidos en la etapa anterior nos dimos cuenta de que se cometió un error en la reacción de epimerización ya que colocamos la quinasa equivocada ya que no tomamos en cuenta la especificidad de esta.

Por último, se trató de producir análogos de ácido siálico en un sistema fermentativo utilizando diferentes variantes genéticas de *Escherichia coli*. Se obtuvieron dos de los análogos esperados, el Neu5Pr y el Neu5nBu, Por otro lado, se confirmó la actividad de NagA hacia sustratos con grupos acilo pequeños.

Acknowledgments

First of all, I want to thank my family for their total and unconditional support during all the years I was away from home and during the writing of this document, without them this would not have been possible.

I want to thank Chris for his patience and enthusiasm, that enthusiasm he showed every morning when, first thing in the morning he asked, “how science is going?”. As well as all the support he gave me during the preparation of this manuscript. Also, to all my colleagues and friends from the Boddy Group; I learned from everyone things that have made me a better professional, better researcher and teacher.

To all those people I met throughout these years at the University and to the staff of the University of Ottawa, you guys make the University a special place to be and learn. To all the Costa Ricans I met in Ottawa, thank you guys, you allowed me to make my stay in Canada even more enjoyable.

Of course, I cannot fail to thank the Universidad Nacional and the Ministerio de Ciencia y Tecnología de Costa Rica for financing my scholarship

“Perseverance, secret of all triumphs.”

Victor Hugo

Table of Contents

List of acronyms and abbreviations

ADP	Adenosine diphosphate
AMP	Adenosine monophosphate
Asp	Aspartic acid
ATP	Adenosine triphosphate
CBZ	benzyl chloroformate
DCM	Dichloromethane
DMAP	4-Dimethylaminopyridine
DNA	Deoxyribonucleic acid
E. coli	Escherichia coli
Econf	conformational strain
Eint	energy of interaction
EtOAc	Ethylacetate
GalNAc	N-acetylgalactosamine
GlcNAc	N-acetylglucosamine
GlcNAc-1-P	N-acetylglucosamine-1-phosphate
GlcNAz	N-azidoglucosamine
GlcNBz	N-benzoylglucosamine
$GlcNH_2$	Glucosamine
$GlcNH_2$-1-P	Glucosamine-1-phosphate
$GlcNH_2$-6-P	Glucosamine-6-phosphate
GlcNHpty	N-heptynoylglucosamine

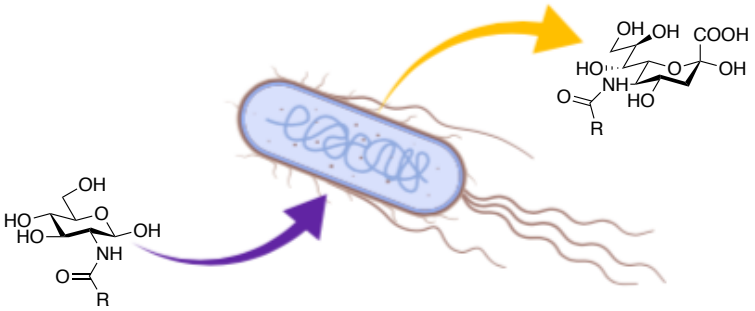
GlcNibu	N-isobutylglucosamine
GlcNiVa	N-isovaleroylglucosamine
GlcNnBu	N-butylglucosamine
GlcNnHp	N-heptanoylglucosamine
GlcNPiv	N-pivaloylglucosamine
GlcNPr	N-propylglucosamine
GlcNVa	N-valeroylglucosamine
GlmM	Phosphoglucosamine mutase
GlmS	$GlcNH_2$-6-P synthase
GlmU	Bifunctional protein GlmU
GNE	UDP-GlcNAc 2-epimerase/ManNac Kinase
HEF	Hemagglutinin-esterase-fusion
His	Histidine
IPTG	Isopropyl-β-D-thiogalactoside
kDa	kilo Dalton
Kdn	2-keto-3-3deoxy-D-glycero-D-galacto-nonulosonic acid
LB	Lysogeny broth
Leg5,7Ac2	legionaminic acid
MALDI-TOF	Matrix-assisted laser desorption/ionization- time of flight
ManNAc	N-acetylmannosamine
ManNAc-6-P	N-acetylmannosamine-6-phosphate
MeOH	Methanol
NA	Neuraminidase

NADH	nicotinamide adenine dinucleotide plus hydrogen
NagA	N-acetylmannosamine-6-phosphate deacetylase
NagB	Glucosamine-6-phosphate deaminase
NagK	N-acetyl-D-glucosamine kinase
NanA	N-acetylneuraminate lyase
NanE	N-acetylmannosamine-6-phosphate 2-epimerase
NANP	N-acetylneuraminic acid phosphatase
NANS	N-acetylneuraminic acid synthase
NanT	Sialic acid transporter
Neu5Ac	N- acetylneuraminic acid
Neu5Gc	N-glycolyneuraminic acid
NeuB	N-acetylneuraminic acid synthetase
NeuC	UDP-GlcNAc 2-epimerase
NHS	N-Hydroxysuccinimide
PCR	Polymerase chain reaction
PTS	phosphoenolpyruvate-carbohydrate phosphotransferase system
UDP-GlcNAc	Uridine diphosphate N-acetylglucosamine
UTP	Uridine triphosphate

List of Figures

List of tables

Chapter 1. Introduction



Carbohydrates are the principal source of energy in living organisms. These compounds are chemically defined as polyhydroxy-aldehydes and polyhydroxy-ketones that can form rings (hemiacetals) and polymers under physiological conditions[1]. The relative position of the hydroxyl groups to the plane of the cycle confers chemical individuality to the molecule, for example, glucose and mannose differs in the stereochemical configuration of the OH on carbon two. Another important structural feature is the anomeric freedom of the glycosidic bonds. This configuration allows even more structural diversity due to the α and β linkages[1].

Although, carbohydrates are used as primary energy source, they play wide variety of roles in living organisms such as receptor recognition, structural, cell signaling, post-translational modification of proteins, etc[2].

Due to the large variety of roles the glycans have, organizing them becomes a difficult task. In the broadest sense, glycans can be grouped into three large categories. The first is glycans that have a structural role or that modulate the function of a biomolecule they are conjugated to. The other two groups involve specific glycan recognition interactions. If the interactions are between cells of the same kind, it is called intrinsic recognition, and if the interaction is with cells of a different type it is known as extrinsic

recognition (figure 1.1). A special type of extrinsic recognition is mimicry. In this case, a microorganism –virus, bacteria, etc.– or a toxin, can emulates the glycans in the cell surface to be recognized as a harmless particle. Table 1 shows some of the roles glycans can have in the cell[1].

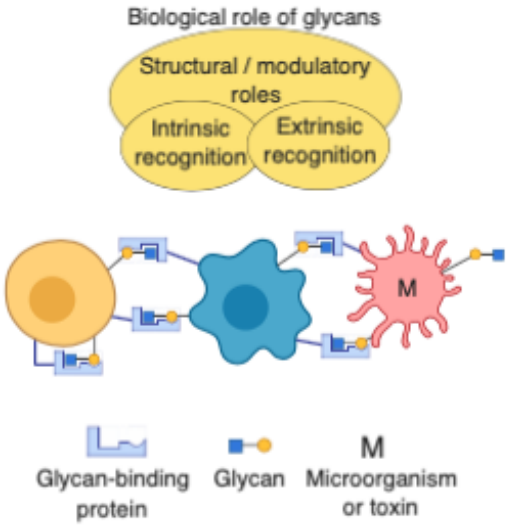


Figure 1. 1. General classification of glycans. Adapted from Varki, 2017[2].

More than half of all proteins and many lipids in biological systems are glycosylated[3,4]. There are three posttranslational protein modifications that involve the addition of carbohydrates. These modifications are N- and O- glycosylation and glycosyl phosphatidyl inositol anchors[3,4].

[1] For a more detailed list, check table 1 in Biological roles of glycans[2].

Table 1. 1. Biological role of glycans[2]

Structural and modulatory roles	Extrinsic (interspecies) recognition of glycans	Intrinsic (intraspecies) recognition of glycans	Molecular mimicry of host glycans
Physical structure,	Bacterial, fungal and parasite adhesins	Intracellular glycoprotein folding and degradation	Convergent evolution of host-like glycans
Physical protection and tissue elasticity,	Viral agglutinins	Intracellular glycoprotein trafficking	Appropriation of host glycans
Water solubility of macromolecules,	Bacterial and plant toxins	Triggering of endocytosis and phagocytosis	
Physical expulsion of pathogens,	Soluble host proteins that recognize pathogens	Intercellular signaling	
Diffusion barriers,	Pathogen glycosidases	Intercellular adhesion	
Protection from proteases	Host decoys	Cell–matrix interactions	
Modulation of membrane receptor signaling	Herd immunity	Fertilization and reproduction	
Membrane organization	Pathogen-associated molecular patterns	Clearance of damaged glycoconjugates and cells	
Antiadhesive action	Immune modulation of host by symbiont/parasite	Glycans as clearance receptors	
Protection from immune recognition	Antigen recognition, uptake and processing	Danger-associated molecular patterns	
Cell surface glycan:lectin-based lattices	Bacteriophage recognition of glycan targets	Self-associated molecular patterns	
Masking or modification of ligands for glycan-binding proteins		Antigenic epitopes	
		Xeno-autoantigens	

1.1 Sialic acids play a key role in biology

Nonulosonic acid is the generic name of a compound family with more than 40 different natural compounds characterized by the nine carbon keto sugar acid skeleton[5,6]. The most representative compound of this family is the sialic acid (fig. 1.2), or 2-keto-3-deoxy-5-acetamido-D-glycero-D-galacto-nonulosonic acid (*N*-acetylneuraminic acid [Neu5Ac])[5].

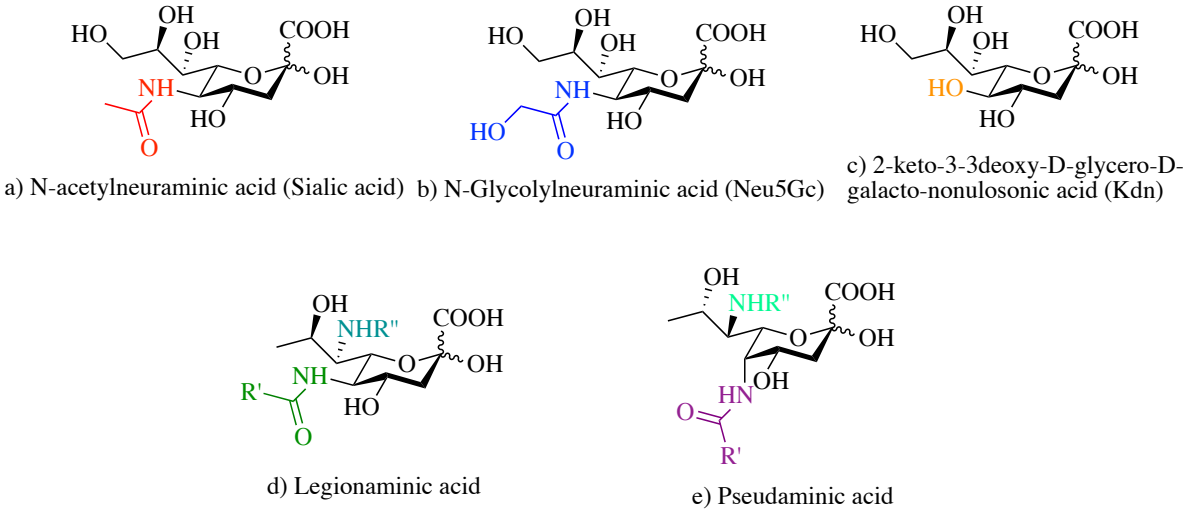


Figure 1. 2. Structure of natural occurring nonulosonic acids. A-C are part of the sialic acid family, D-E are other common nonulosonic acids[7,8].

Sialic acids are, typically, the most common monosaccharide in the outer region of glycolipids and glycoproteins associated with the cell membrane and are often part of the mechanism of recognition of pathogens[5,9,10].

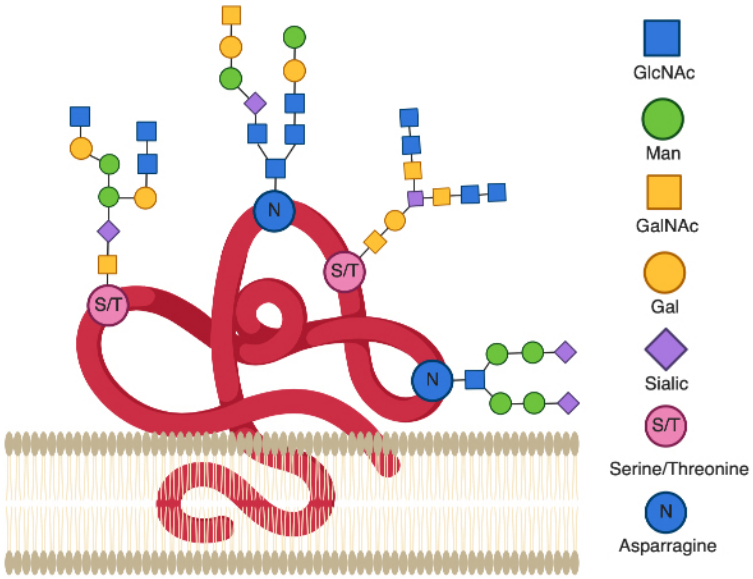


Figure 1. 3. Representation of glycosylation process in a cell protein. Adapted from Van den Steen *et al.*, 2008[3].

Sialic acids are the most common terminal sugars in the membrane proteins on the cell surface of eukaryotic cells[1], and are the only nine carbon sugars present in prokaryotes[5]. In eukaryotes, sialic acids are crucial for several biological processes, such as cellular adhesion[11], stabilization of the structure of glycoproteins, cell-cell interactions, chemical signaling, regulating transmembrane receptor function, affecting membrane transport, controlling the half-lives of circulating glycoproteins and cells,[5] and cell surface sialylation has been correlated with the tumorigenicity and metastatic behavior of malignant cells[11]. On prokaryotes, microbial sialic acid has been identified as a virulence determinant in a range of infectious diseases[5].

There are some sialic acids that are present on eukaryotes exclusively, others are present in bacteria and there are others, like Neu5Ac, that can be found both eukaryote and prokaryote organisms[5]. Human pathogens, such as *Escherichia coli*, *Pseudomonas pneumonia*, take advantage of a Neu5Ac-based capsule to mimic the host cells and avoid detection by the immune system[12]. For this reason, in pathogenic bacteria, sialic acid is correlated with the pathogenicity. Pathogenic bacteria can produce sialic acid through two different ways; *de novo* synthesis and scavenging from the environment. These two methods provide opportunities to scientists to find compounds capable of inhibiting sialidases (enzymes that cleave sialic acids from host cell wall) and inhibitors of sialic acid anabolic pathway.

1.2 Sialic acid analogs as tools to study physiology and pathology

In line with their vital role in many physiological processes, several lines of evidence imply that aberrant expression of sialic acids confers major advantages to tumor

cells, ranging from inhibition of apoptosis to resistance to cancer therapy. [1314–16]. Taking advantage of this observation Bertozzi's group has developed several methods to incorporate chemically modified sialic acids onto the cell surface of tumors by feeding precursor *N*-acylmannosamines[14,17,18]. With this technology, they have been able to label different tissues in vivo with fluorescent tags and use these for detection of tumors in living mice.

Sialic acid analogues are valuable tools in analyzing the interactions between glycoconjugates involved in different processes. Sialic acid is recognized as a receptor by more viruses than any other determinant known[19,20]. Substituents at C-9 of sialic acid were used to study the activity of the influenza virus C glycoprotein HEF[19]. Herrler *et al*. concluded that the structure of the substituent in a sialic acid analog is critical for its recognition for influenza virus C[19]. Their experiments show that C-9 acetylated sialic acid is a valuable tool to study the importance of the receptor destroying enzyme, also, it will be used to obtain information about the role of the enzyme in the maturation process of the virus[19]. Virus receptors have been defined as structures on the cell surface to which virus attachment produces a biological response such as infection of the cell[20].

1.3 Biosynthesis of sialic acids

The biosynthesis of sialic acid in eukaryotes (fig 1.4) is a high regulated process and begins with fructose-6-phosphate, which is directed to the hexosamines synthesis by glutamine fructose-6-phosphate amidotransferase[7,21]. Following the acetylation of the free amine group of glucosamine-6-phosphate to form *N*-acetylglucosamine-6-phosphate. This sugar undergoes an epimerization reaction to produce GlcNAc-1-P which reacts with UTP

to produce UDP-GlcNAc. Subsequently, the action of the bifunctional enzyme, UDP-GlcNAc 2-epimerase/ManNac Kinase (GNE) commits UDP-GlcNAc to the sialic acid biosynthesis pathway[21]. GNE catalyzes two reactions, first the hydrolysis of UDP-GlcNAc followed by an epimerization to form ManNAc. Then, GNE catalyze the phosphorylation of ManNAc to produce ManNAc-6-phosphate.

Free sialic acid is produced in two steps. The first reaction is the condensation of ManNAc-6-phosphate with phosphoenol pyruvate, mediated by *N*-acetylneuraminic acid synthase (NANS) to produce Neu5Ac-9-phosphate. The second reaction, *N*-acetylneuraminic acid phosphatase (NANP) removes the phosphate group to produce Neu5Ac[7,21].

Bacterial sialic acid biosynthesis is highly similar to the eukarotic pathway with key differences. The first reaction involves the transformation of fructose-6-phosphate into $GlcNH_2$-6-phosphate by GlmS ($GlcNH_2$-6-P synthase). $GlcNH_2$-6-P is epimerized by phosphoglucosamine mutase (GlmM) into $GlcNH_2$-1-P. After the epimerization, the bifunctional protein GlmU catalyze two reactions, first, the acetylation of $GlcNH_2$-1-P to generate GlcNAc-1-P, second, GlcNAc-1-P reacts with UTP to produce UDP-GlcNAc. UDP-GlcNAc 2-epimerase (NeuC) then transform the common cell wall precursor into ManNAc. The elimination reaction catalyzed by NeuC is thought to involve the formation of a 2-acetamidoglucal intermediate followed by the irreversible epimerization of this intermediate to ManNAc, as happens in mammalian systems[5,22]. ManNAc reacts with phosphoenolpyruvate by a condensation reaction catalyzed by N-acetylneuraminic acid synthetase (NeuB) to produce Neu5Ac[5].

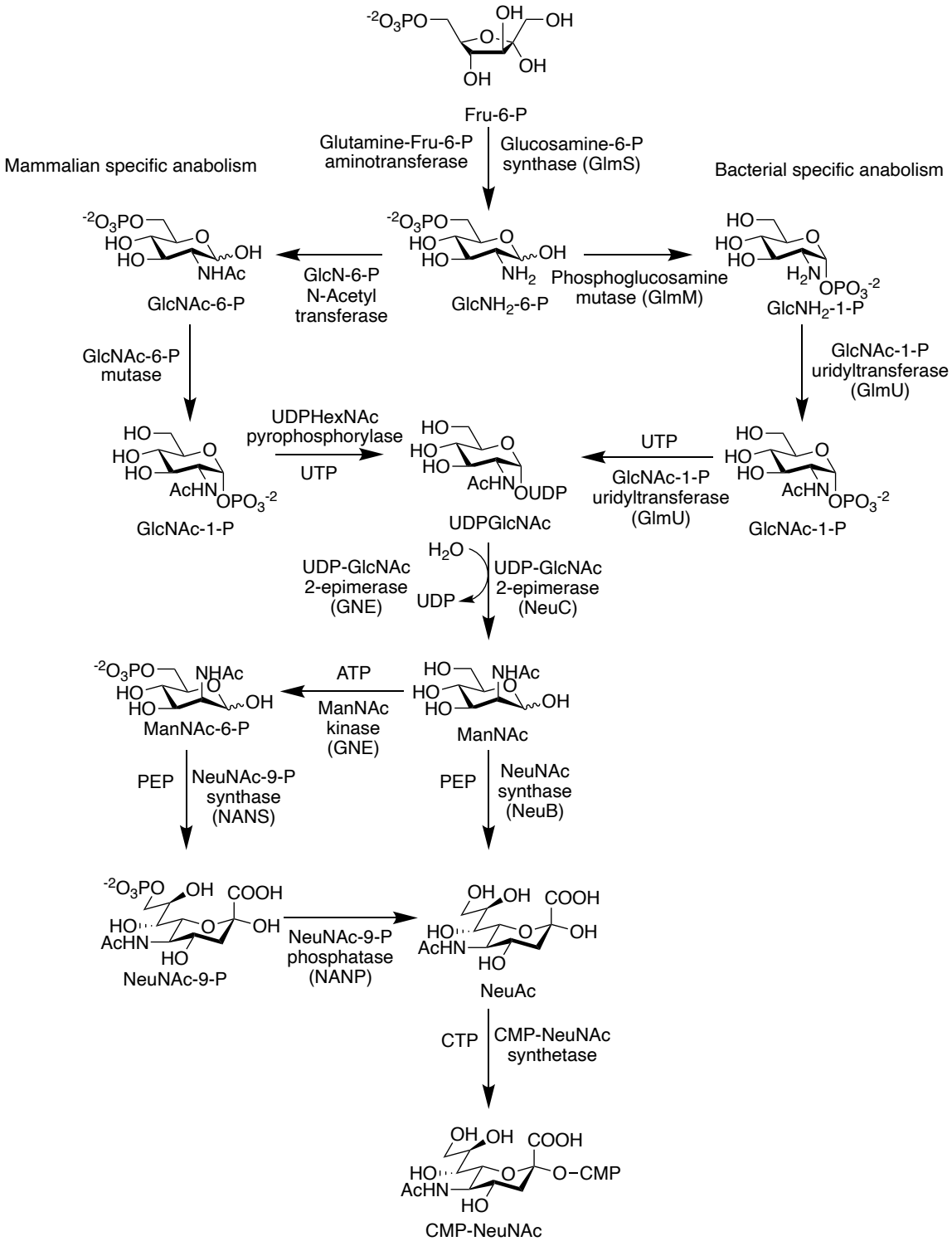


Figure 1. 4. Comparison of sialic acid biosynthetic routes for eukaryotes and prokaryotes organisms[5,6,23,24].

1.4 Sialic acid production

As mentioned before, nonulosonic acids play several and very different roles in living organisms, ranging from neural plasticity to pathogen interactions. Moreover, the structural and functional diversity of this compound family and derivatives make these metabolites important targets to be synthesized for biological research and drug development[8,9,25]. The methodologies to synthesize these compounds are challenging and are related to the 9 carbon backbone of the sialic acids, as well as the relatively acid liability of its glycosidic bonds, and the instability of some of its modifications.[26]

The chemical synthesis of sialic acid is very challenging. It involves many steps, complex functional group protections, and coupling strategies for the control of both the stereochemistry and regiochemistry of bond formation[25]. For this reason, many research groups have tried to find an alternative route to produce these molecules. One alternative is the development of engineered bacterial strains that overproduce sialic acid and releases the product into the culture media by the inactivation of the recovery or absorption mechanism.

Lundgren and coworkers were able to develop an *E. coli* strain capable to produce 1.7 g/L of sialic acid. To achieve this, they removed the endogenous transferase (NanT) and aldolase (NanA) and introduced two enzymes from *N. meningitidis* –NeuB and NeuC. They overexpressed GlmS in order to use inexpensive carbon sources such as glucose, fructose and glycerol(fig. 1.5)[27].

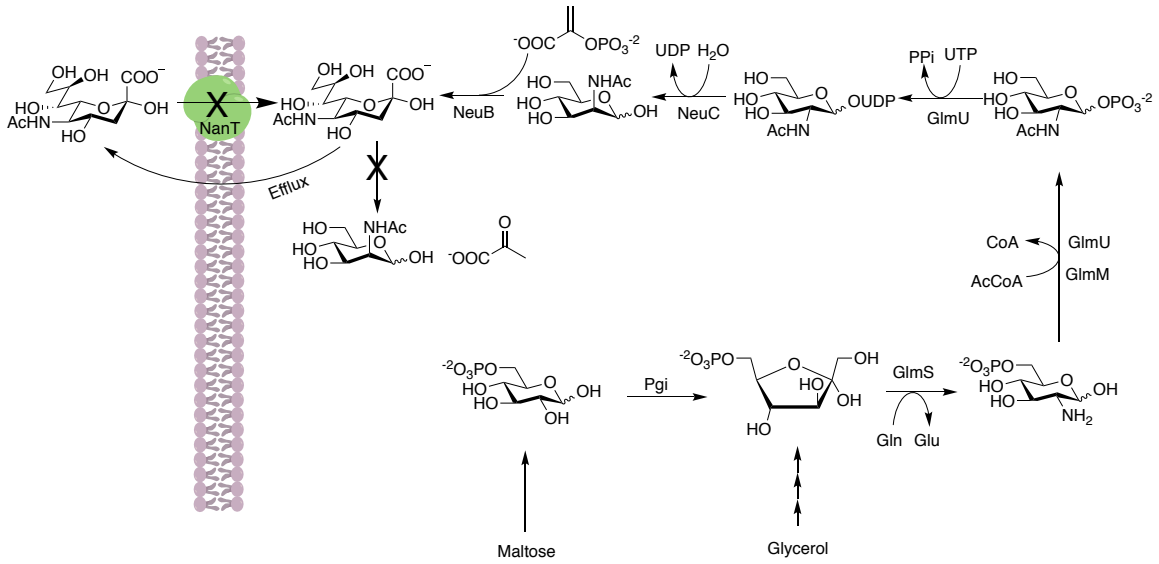


Figure 1. 5. Sialic acid production by metabolically engineered *E. coli*[27]

Another successful example to produce a nonulosonic acid is shown by Hassan *et al*. In this case, the group installed genes from *Saccharomyces cerevisiae*, *Campylobacter jejuni* and *Legionella pneumophila* in *E. coli* (fig. 1.6). This engineered strain can produce close to 120 mg/L of legionaminic acid (Leg5,7Ac$_2$)[28].

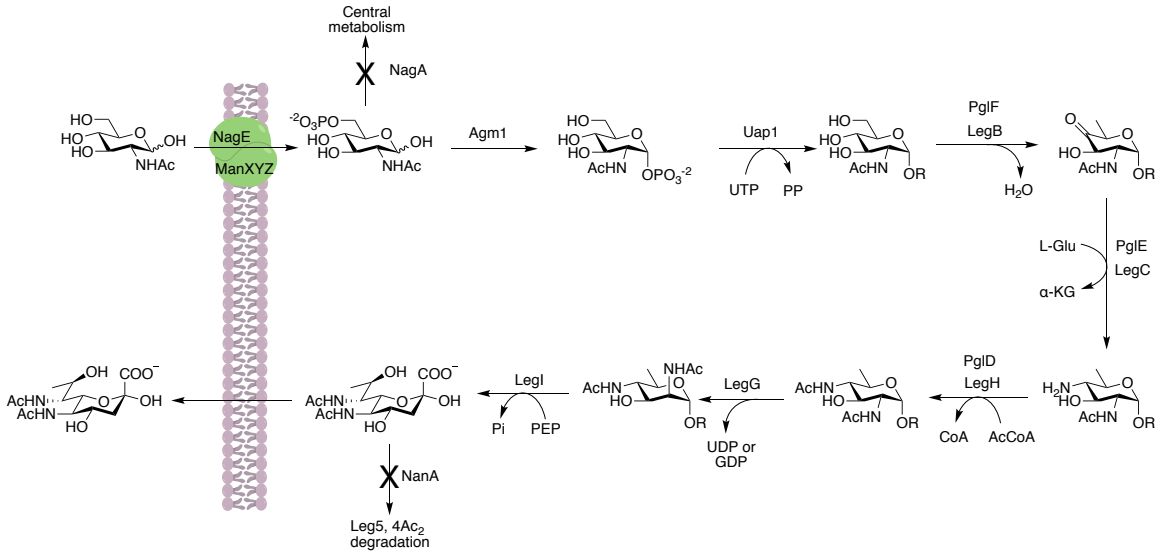


Figure 1. 6. Legionaminic acid production by metabolically engineered *E. coli*[28].

Figure 1.7 shows the metabolic pathways involved in sialic acid production and degradation in *E. coli*. This figure shows different entry points where the bacteria can utilize several compounds as precursors for the biosynthesis of sialic acid. Murein recycling produces GlcNAc-6-P that can be deacetylated by NagA and produce glucosamine-6-P that can be transformed by NagB into fructose-6-P. $GlcNH_2$-6-P can be incorporated by the cell via the PTS system which takes the extracellular $GlcNH_2$ and incorporate it into the cell with the phosphorylation on C6.

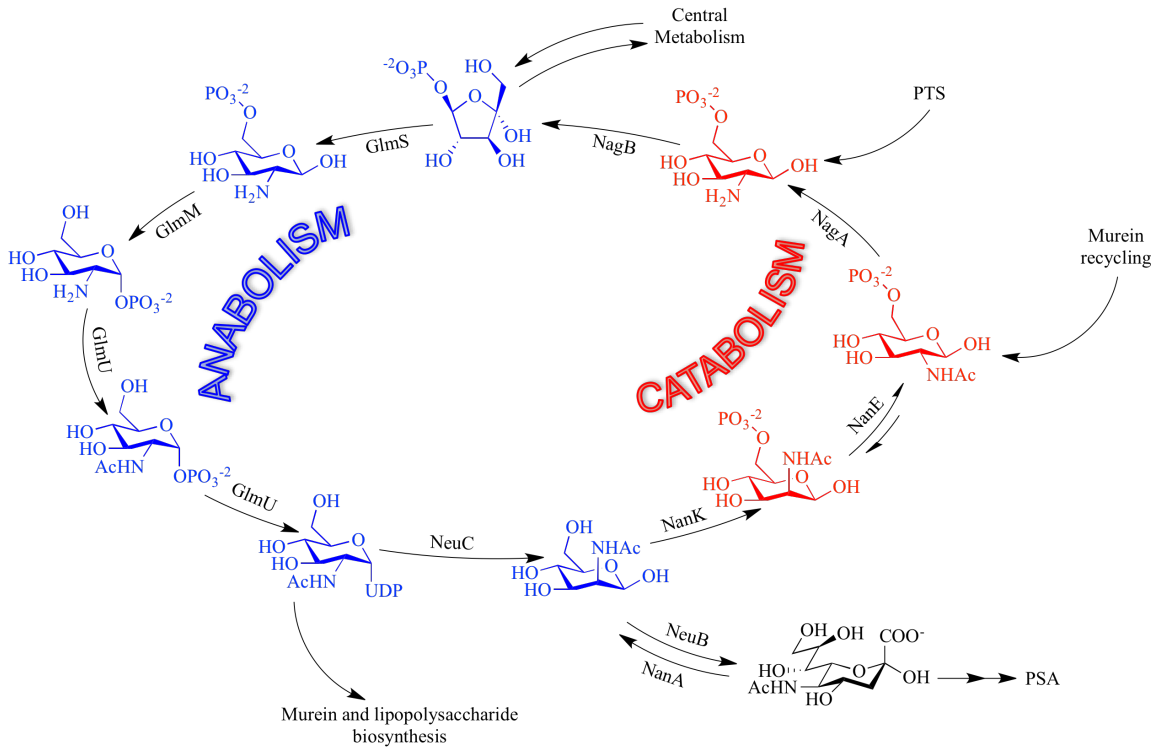


Figure 1. 7. Sialic acid metabolism[5,29,30].

1.5 Production of sialic acid analogs

Several research groups have shown the importance and the applications of non-natural nonulosonic acid derivatives. These unnatural compounds can be used as cancer

treatments[31–33] or vaccine development[34]. However, their synthesis and over-production is also highly challenging.

As mentioned previously, chemical synthesis of Sialic acids is difficult, and the use of metabolic engineered bacteria has shown very promising results at producing high levels of natural nonulosonic acids. A reasonable strategy could thus be the use of engineered bacteria to produce the unnatural sialic acids. In order to achieve this goal, the bacteria could be fed with an analog of the sugar used as precursor.

However, this straightforward approach is not immediately viable due to the fact that GlmM and GlmU show high substrate specificity[27,35,36]. The glucosamine mutase (GlmM) is an enzyme responsible to transform glucosamine-6-phosphate ($GlcNH_2$-6-P) to glucosamine-1-P ($GlcNH_2$-1-P). On the other hand, GlmU is an enzyme with dual activity. First, GlmU catalyze the acetylation on the amino group in $GlcNH_2$-1-P to produce GlcNAc-1-P, then, it catalyzes the installation of an uridyl group in C1, producing GlcNAc-1-UDP. These two enzymes prevent the production of N-acyl sialic analogs by fermentation directly from N-acylglucosamines.

1.6 Thesis Objectives and Hypotheses

In this thesis, we developed an *E. coli*-based system to produce N-acyl sialic acid analogs. This route relies on NanE, an epimerase that converts N-acylglucosamines into N-acylmannosamine, which then can intercept the synthesis pathway to produce sialic acid analogs. Since N-acetylglucosmaine is first deacetylated to glucosamine by NagA, we were concern that the N-acylglucosamine substrates may not prove stable enough to use. Thus, we first determined the substrate scope of recombinant purified NagA to understand its

selectivity towards the synthetic substrates. During this process we analyzed the action of the kinase involved in the first step of the aminosugar recycling pathway what led us to develop a short and simple synthesis to get 6-phosphoamino hexopyranoses. Finally, we tested our hypothesis by designing an in vivo experiment where an *E. coli* strain was modified to over produce the two enzymes involved in our proposed route, NanE and NeuB. This experiment helped us to determine feasibility of using bacteria to convert a non-natural N-acylglucosamine into the correspondent sialic acid.

References

1. Gerardy-Schahn, R., Delannoy, P. & von Itzstein, M. *SialoGlyco Chemistry and Biology I*. **366**, (2015).
2. Varki, A. Biological roles of glycans. *Glycobiology* **27**, 3–49 (2017).
3. Steen, P. Van den, Rudd, P. M., Dwek, R. A. & Opdenakker, G. Concepts and Principles of O-Linked Glycosylation. *Crit. Rev. Biochem. Mol. Biol.* **33**, 151–208 (2008).
4. Corfield, A. P. Structure/function of O -glycans. in *Encyclopedia of Genetics, Genomics, Proteomics and Bioinformatics* (John Wiley & Sons, Ltd, 2004). doi:10.1002/047001153X.g305205
5. Vimr, E. & Kalivoda, K. Diversity of microbial sialic acid metabolism. *Microbiol. Mol. Biol. Rev.* **68**, 132–153 (2004).
6. Du, J. *et al.* Metabolic glycoengineering: sialic acid and beyond. *Glycobiology* **19**, 1382–401 (2009).
7. Varki, A. & Schauer, R. Sialic Acids. in *Essentials of Glycobiology* 1–13 (Cold

Spring Harbor Laboratory Press, 2009).

8. Chen, X. & Varki, A. Advances in the biology and chemistry of sialic acids. *ACS Chem. Biol.* **5**, 163–176 (2010).

9. Varki, N. M. & Varki, A. Diversity in cell surface sialic acid presentations: implications for biology and disease. *Lab Invest* **87**, 851–857 (2007).

10. Sharmila, D. J. S., Blessia, Tf. & Rapheal, V. Molecular dynamics of sialic acid analogues and their interaction with influenza hemagglutinin. *Indian J. Pharm. Sci.* **72**, 449 (2010).

11. Mantey, L. R., Keppler, O. T., Pawlita, M., Reutter, W. & Hinderlich, S. Efficient biochemical engineering of cellular sialic acids using an unphysiological sialic acid precursor in cells lacking UDP-N-acetylglucosamine 2-epimerase. *FEBS Lett.* **503**, 80–84 (2001).

12. Vimr, E. & Lichtensteiger, C. To sialylate, or not to sialylate: That is the question. *Trends Microbiol.* **10**, 254–257 (2002).

13. Pearce, O. M. T. & Läubli, H. Sialic acids in cancer biology and immunity. *Glycobiology* **26**, 111–128 (2015).

14. Laughlin, S. T. & Bertozzi, C. R. Imaging the glycome. *Pnas* **106**, 12–17 (2009).

15. Fuster, M. M. & Esko, J. D. The sweet and sour of cancer: glycans as novel therapeutic targets. *Nat. Rev. Cancer* **5**, 526 (2005).

16. Dennis, J. W., Granovsky, M. & Warren, C. E. Glycoprotein glycosylation and cancer progression. *Biochim. Biophys. Acta - Gen. Subj.* **1473**, 21–34 (1999).

17. Bertozzi, C. R., Hang, H. C., Hanover, J. A., Kim, E.-J. & Vocadlo, D. J. A chemical approach for identifying O-GlcNAc-modified proteins in cells. *Proc. Natl. Acad.*

Sci. **100**, 9116–9121 (2003).

18. Prescher, J. A., Dube, D. H. & Bertozzi, C. R. Chemical remodelling of cell surfaces in living animals. *Nature* **430**, 873–877 (2004).

19. Herrler, G. *et al.* A synthetic sialic acid analogue is recognized by influenza C virus as a receptor determinant but is resistant to the receptor-destroying enzyme. *J. Biol. Chem.* **267**, 12501–12505 (1992).

20. Lentz, T. L. The recognition event between virus and host cell receptor: A target for antiviral agents. *J. Gen. Virol.* **71**, 751–766 (1990).

21. Bhide, G. P. & Colley, K. J. Sialylation of N-glycans: mechanism, cellular compartmentalization and function. *Histochem. Cell Biol.* **147**, 149–174 (2017).

22. Chou, W. & Hinderlich, S. Sialic acid biosynthesis: stereochemistry and mechanism of the reaction catalyzed by the mammalian UDP-N-acetylglucosamine 2-epimerase. *J. Am. ...* **125**, 2455–2461 (2003).

23. Tanner, M. E. The enzymes of sialic acid biosynthesis. *Bioorg. Chem.* **33**, 216–28 (2005).

24. Nagel, A. K. & Ball, L. E. *Intracellular protein O-GlcNAc modification integrates nutrient status with transcriptional and metabolic regulation*. *Advances in Cancer Research* **126**, (Elsevier Inc., 2015).

25. Lv, X. *et al.* Synthesis of Sialic Acids, Their Derivatives, and Analogs by Using a Whole-Cell Catalyst. *Chem. - A Eur. J.* **23**, 15143–15149 (2017).

26. Ang, T. N., Ngoh, G. C., Seak, A. & Chua, M. A quantitative method for fungal ligninolytic enzyme. *Screening* 589–595 (2011). doi:10.1002/apj

27. Lundgren, B. R. & Boddy, C. N. Sialic acid and N-acyl sialic acid analog production

by fermentation of metabolically and genetically engineered Escherichia coli †. *Org. Biomol. Chem.* 1903–1909 (2007). doi:10.1039/b703519e

28. Hassan, M. I. *et al.* Total Biosynthesis of Legionaminic Acid, a Bacterial Sialic Acid Analogue. *Angew. Chemie - Int. Ed.* **55**, 12018–12021 (2016).

29. Vimr, E. R. & Troy, F. a. Regulation of sialic acid metabolism in Escherichia coli: role of N-acylneuraminate pyruvate-lyase. *J. Bacteriol.* **164**, 854–60 (1985).

30. Ringenberg, M., Lichtensteiger, C. & Vimr, E. Redirection of sialic acid metabolism in genetically engineered Escherichia coli. *Glycobiology* **11**, 533–539 (2001).

31. Kiefel, M. J. & von Itzstein, M. Recent Advances in the Synthesis of Sialic Acid Derivatives and Sialylmimetics as Biological Probes. *Chem. Rev.* **102**, 471–490 (2002).

32. Wang, H. *et al.* Selective in vivo metabolic cell-labeling-mediated cancer targeting. *Nat. Chem. Biol.* **13**, 415–424 (2017).

33. Büll, C. *et al.* Metabolic sialic acid blockade lowers the activation threshold of moDCs for TLR stimulation. *Immunol. Cell Biol.* **95**, 408–415 (2017).

34. Micoli, F., Costantino, P. & Adamo, R. Potential targets for next generation antimicrobial glycoconjugate vaccines. *FEMS Microbiol. Rev.* **42**, 388–423 (2018).

35. Gehring, A. M., Lees, W. J., Mindiola, D. J., Walsh, C. T. & Brown, E. D. Acetyltransfer Precedes Uridylyltransfer in the Formation of UDP- N -acetylglucosamine in Separable Active Sites of the Bifunctional GlmU Protein of Escherichia coli † . *Biochemistry* **35**, 579–585 (2002).

36. Mengin-Lecreulx, D. *et al.* Reaction mechanism of phosphoglucosamine mutase from Escherichia coli. *Eur. J. Biochem.* **262**, 202–210 (2003).

Chapter 2. Biochemical study of *N*-acetyl-D-glucosamine-6-phosphate deacetylase (NagA)

2.1 Introduction

N-acetyl-D-glucosamine-6-phosphate deacetylase (NagA) catalyzes the hydrolysis of *N*-acetylglucosamine-6-phosphate (GlcNAc-6-P) into glucosamine-6-phosphate ($GlcNH_2$-6-P) and acetate (fig. 2.1). Figure 2.2 shows the reaction, which plays an essential role in cell wall recycling in bacteria. Peptidoglycan is constantly being turned over by the cell and converted into GlcNAc, which is phosphorylated by *N*-acetyl-D-glucosamine kinase (NagK) to produce intracellular GlcNAc-6-P. NagA hydrolyze the acetate group and the resultant glucosamine-6-P ($GlcNH_2$-6-P) can be readily converted into fructose-6-P, which supplies central metabolism with both energy and biosynthetic intermediates or back into uridine 5'-diphosphate-N-acetylglucosamine (UDP-GlcNAc) to support cell wall biosynthesis.

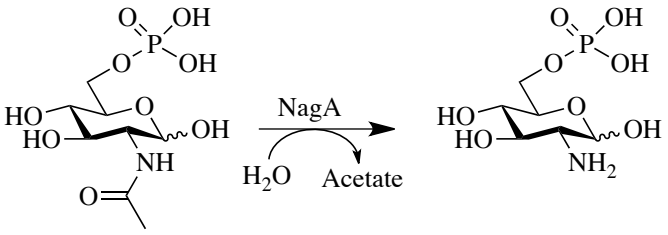


Figure 2. 1. Deacetylation reaction catalyzed by NagA.

In the context of the *N*-acetyl neuraminic acid overproducing *E. coli* strain BRL02 developed previously by the Boddy lab, NagA plays an essential role converting intracellular GlcNAc-6-P, generated from exogenously fed GlcNAc by the phosphotransferase system, into $GlcNH_2$-6-P. The *Glm* operon is then responsible for converting $GlcNH_2$-6-P into UDP-GlcNAc, which can be used in peptidoglycan or

lipopolysaccharide biosynthesis and by the sialic acid biosynthetic enzymes NeuA and NeuB to produce NeuAc[1] (fig 2.2).

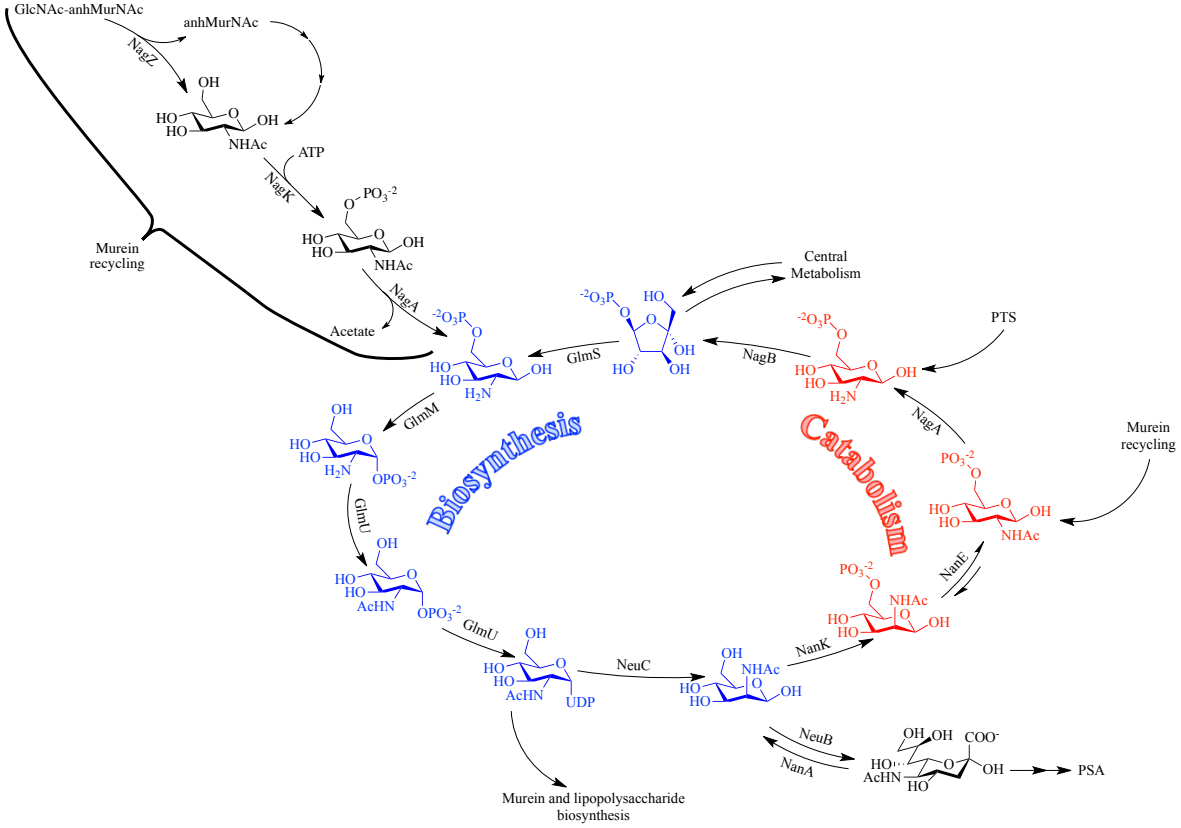


Figure 2. 2. Relationship between sialic acid metabolism and murein degradation that shows how NagA is involved in the recycling of amino sugars. Adapted from Vimr *et al.* and Park J[1,4].

While NagA can play a role in enabling access to NeuAc, we anticipated that it would be incompatible with the production of *N*-acylneuraminic acids if it was able to hydrolyze acyl groups from *N*-acylglucosamine-6-Ps. In this chapter we thus investigate the ability of NagA to hydrolyze acyl chains from *N*-acylglucosamine-6-Ps.

N-acetyl-D-glucosamine-6-phosphate deacetylase belongs to the amidohydrolase superfamily. This superfamily is comprised of a remarkable number of enzymes that hydrolyzes diverse substrates containing amides or esters on a carbon or phosphorous center[2,3]. The most notable structural feature of this superfamily is the presence of one or

two metal ions in the active site. This metal center has a dual functionality in the catalytic activity. The substrate must be activated to make it a better electrophile before bond cleavage and the water must be made more nucleophilic by deprotonation. In all cases, water is activated by complexation with the metal center. In the binuclear metal centers, the carbonyl and phosphoryl groups of the substrates are polarized through Lewis acid catalysis via complexation with the metal ion. In the mononuclear metal centers, the substrate is activated by a proton transfer from the active site[3].

Figure 2.3 shows the reaction mechanism of NagA from *E. coli* proposed by Hall *et al*[5]. Water coordinates with the metal making it acidic. Asp 237 then deprotonates the water with concomitant attack on the carbonyl of GlcNAc-6-P to generate the tetrahedral intermediate. The amine leaving group departs with protonation by Asp 237, thus hydrolyzing the amide bond.

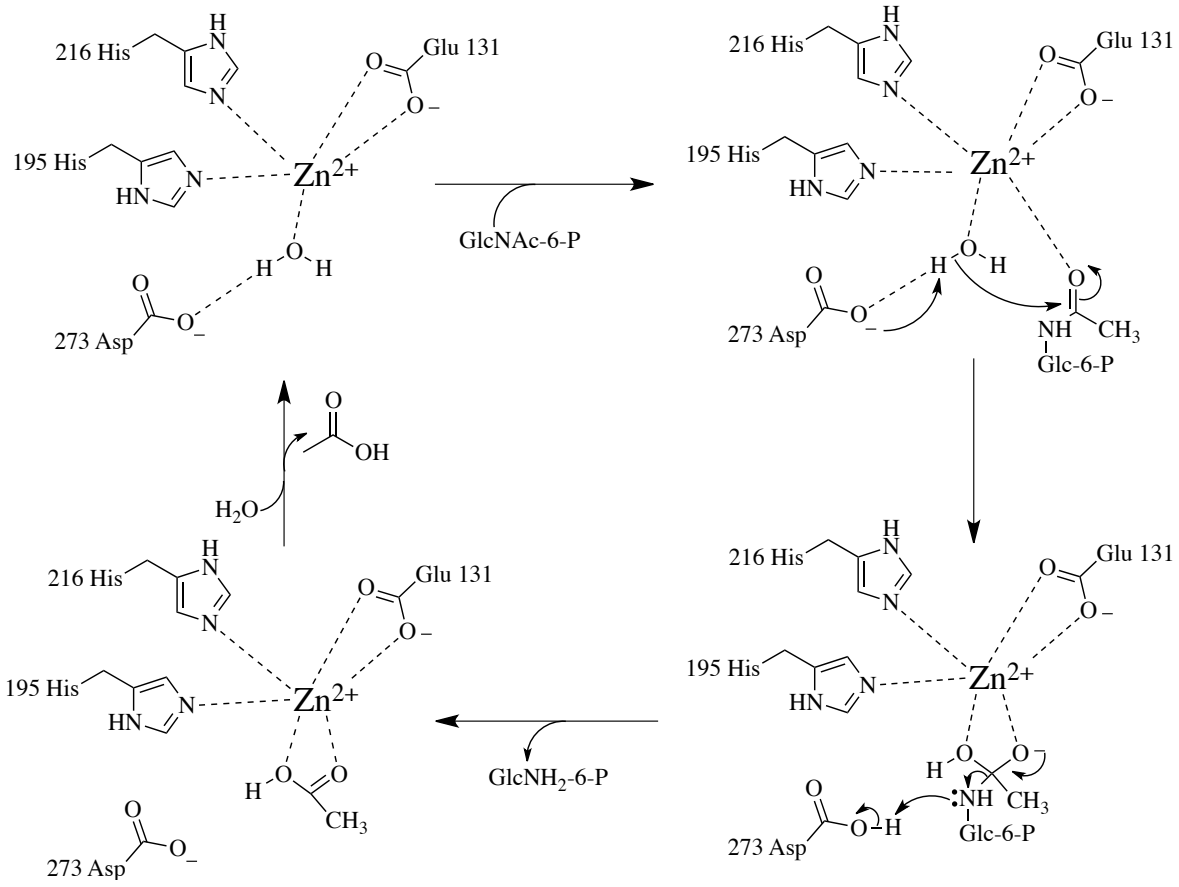


Figure 2. 3. N-acetyl-D-glucosamine-6-phosphate deacetylase hydrolysis reaction mechanism. Adapted from Hall *et al*[5].

In this chapter we characterize the substrate scope of *E. coli* NagA by probing its reactivity in vitro and in vivo with N-acylglucosamine derivatives. In addition, we carry out an in-silico analysis of the substrate selectivity to support our in vitro and in vivo results. While *E. coli* NagA substrate selectivity has been investigated in the past[5–7], no study has explicitly examined the effect of modifying the acetyl group of GlcNAc-6-P. This work thus represents an important addition to the understanding of NagA catalysis in bacterial amino sugar metabolism.

2.2 Results and Discussion

2.2.1 Heterologous Expression of NagA

To obtain the desired gene, DNA from *E. coli* BL21 was isolated using the Wizard gDNA Isolation kit from Promega. Once the genomic DNA was isolated it was used as a template for the amplification of *nagA*. The gene was amplified using a touchdown PCR protocol and was purified by agarose gel electrophoresis.

Touchdown and hotstart PCR were performed for all amplifications, and the thermocycler conditions were as follows: one cycle of 98 °C for 5 min, 14 cycles of 20 s at 98 °C, 20 s at 70 °C (-1 °C per cycle), and 60 s at 68 °C; 16 cycles of 20 s at 98 °C, 20 s at 55 °C, and 60 s at 72 °C, and one cycle of 10 min at 72 °C. Figure 2.4 shows the electrophoresis analysis of the PCR. It shows a band between 1000 and 1500 base pairs what corresponds with the size of the gene; moreover, the intensity of the band indicates the success of the amplification. On the other hand, the gel shows a big band under 250 bp, this corresponds to the presence of primer dimers.

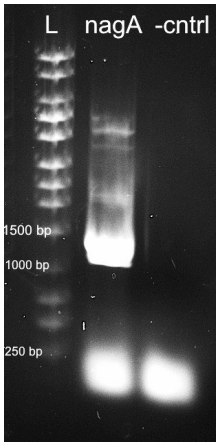


Figure 2. 4. Touchdown PCR of *nagA* analyzed in 0.7% agarose gel.

After the gene was purified, it was inserted into PCR-Blunt II-TOPO vector (Invitrogen, Life Technologies Corporation) and the ligation reaction was used to transform a *E. coli* BL21 strain and plated on LB plates with kanamycin as resistance marker. Six clones were selected, plasmids extracted and digested with EcoRI. The analysis of the electrophoresis shows that *nagA* is present in clones 1, 3 and 5 what gives a transformation efficiency of approximately 50%.

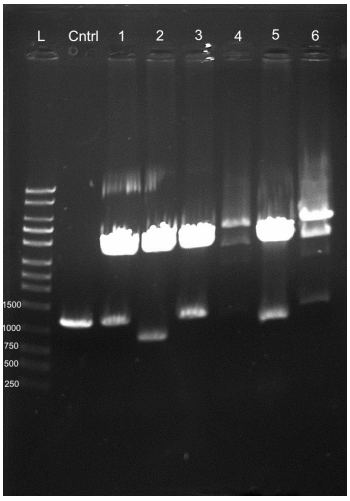


Figure 2. 5. Comparison of six clones from the transformation of *E. coli* Top 10 with the ligation reaction of *nagA* in the blunt vector.

Clone one was selected to be pLRPV02 (appendix) and passed to the next step. EcoRI and NdeI digestion provided the linearized gene, which was inserted into NdeI/EcoRI linearized pET28 to give pLRVP03 (map annex 4). Ligation into pET28 was confirmed by the digestion of the plasmids obtained from 4 clones with EcoRI/NdeI. Figure 2.6 shows the result of the restriction enzyme reaction and a faint band around 1000 bp corresponding to *nagA* and a much brighter band between 5 and 6 kb corresponding to the pET28 backbone. This confirms the success in the construction of the expression system.

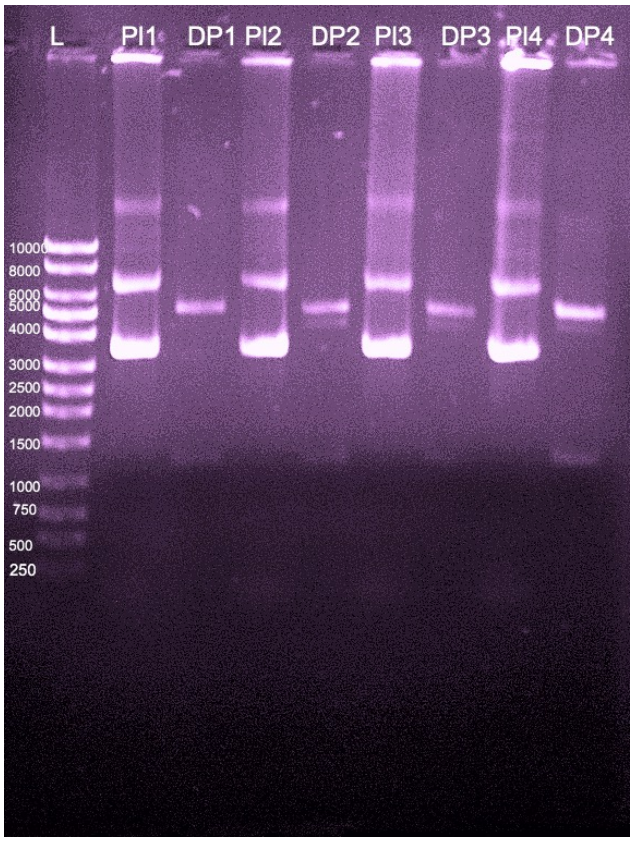


Figure 2. 6. EcoRI/NdeI digestion of pLRVP03. Pl: plasmid; DP: digested plasmid.

Escherichia coli BL21 DE3 was successfully transformed with pLRVP03 to overproduce NagA. The enzyme begins to be produced in detectable amounts 4 hours after induction (fig.2.7). A large band can be seen in fraction 5 between 40 and 50 kDa. The flow through (line F) shows unbound NagA; this indicates that the amount of resin was insufficient due to the high level of overexpression of the enzyme. The highest concentration of enzyme is eluted in the first fraction of 250 mM imidazole, and the molecular weight corresponds with the theoretical value. In addition, a MALDI-TOF was performed to provide a higher resolution analysis of the molecular weight. MALDI-TOF reveals a molecular weight of 42805.610 Da and the calculated mass of NagA+His tag is 43.2 kDa.

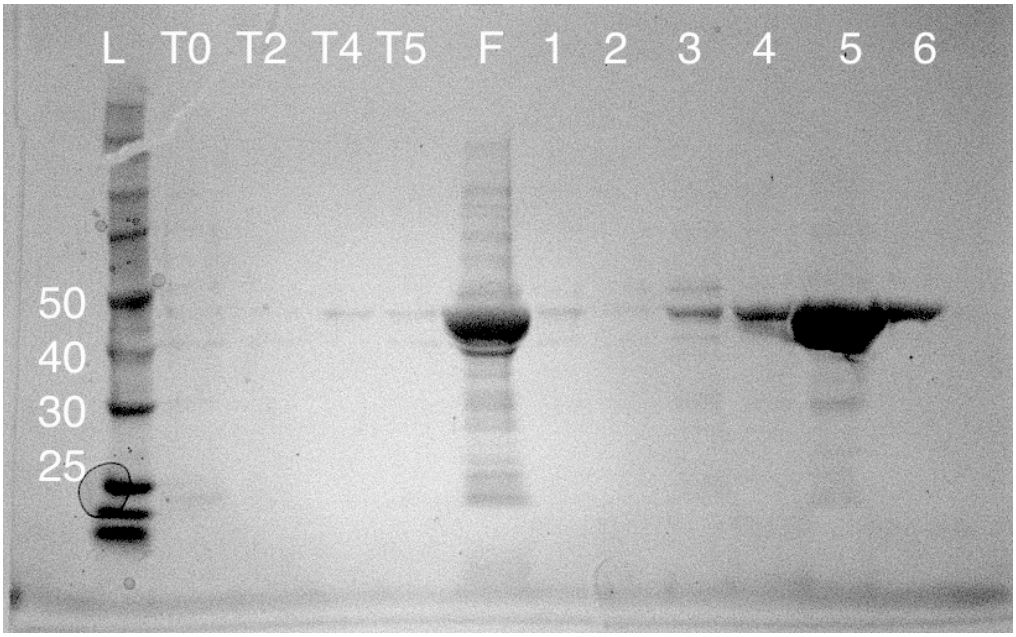


Figure 2. 7. SDS-PAGE of the bacterial lysate. L, ladder (kDa); Tn, time course; F, flow through; 1, 0 mM imidazole; 2, 20 mM imidazole; 3-4, 100 mM imidazole; 5-6, 250 mM imidazole.

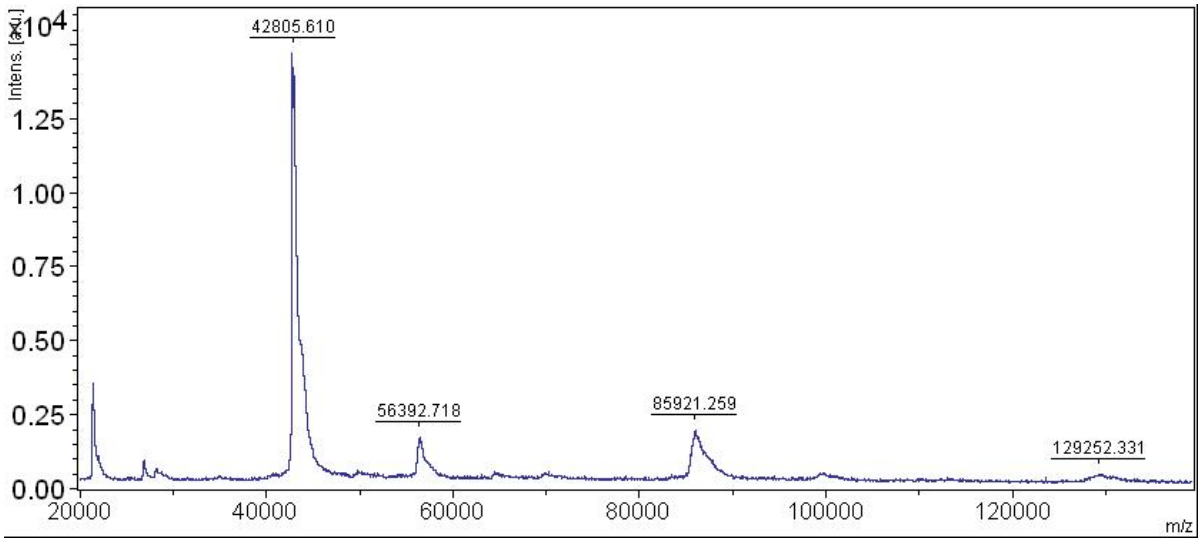


Figure 2. 8. MALDI-TOF analysis of NagA solution using sinapic acid as matrix.

The mass spectrum shows two more peaks, one at 89.9 kDa and another at 129.2 kDa, corresponding to the dimer and trimer respectively. There is a third peak at 56.4 kDa that likely correspond to a contaminant protein that co-eluted with NagA.

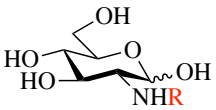
N-acetyl-D-glucosamine-6-phosphate deacetylase shows a good behavior towards the transformation and overexpression in *E. coli*. After the enzyme was overexpressed and characterized, it was used to determine the kinetic parameters with the native substrate and compare it with the GlcNAc analogs.

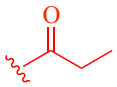
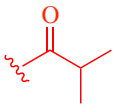
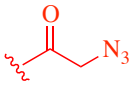
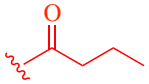
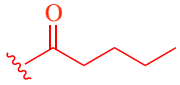
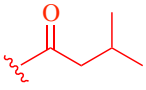
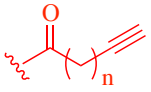
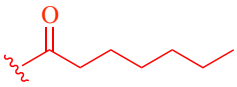
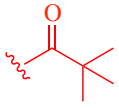
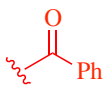
2.2.2 Biochemical characterization of NagA

2.2.2.1 Kinetics

A serious concern regarding the production of sialic acid analogs from *N*-acylglucosamines is the possibility of hydrolysis of the substrate by NagA. For this reason, kinetic analysis was performed with different *N*-acylglucosamines (table 2.1 and table 2.2)

Table 2. 1. *N*-acylglucosamine substrates used in kinetic analysis



R = 	propanoyl (Pr)	R = 	isobutanoyl (iBu)	R = 	azide (Az)
R = 	butanoyl (Bu)				
R = 	valeroyl (Va)	R = 	isovaleroyl (iVa)	R = 	alkyne (Al)
R = 	heptanoyl (Hp)	R = 	pivaloyl (Piv)	R = 	benzoyl (Bz)

The deacetylation reaction was measured using the Ninhydrin assay[9]. Optimal time, pH and concentration were determined in order to set the parameters of the experiment (data not shown).

The aminosugars were dissolved in phosphate buffer with sodium chloride to emulate cellular ionic strength and NagK was added along with 1.2 equivalents of ATP to phosphorylate the sugars.

NagA was added and the enzymatic reaction was placed in a water bath at 37 °C in triplicate. Samples were taken at 0, 2, 5, 8 and 10 minutes. The aliquots were mixed with the ninhydrin reagent and placed in a heating block set at 95 °C for 10 minutes to generate color. After the incubation time, samples were placed on ice to stop the reaction and the absorbance was measured at 575 nm against a reactive blank. Figure 2.9 represents the experimental approach to study the effect of the acyl group in the enzymatic activity of NagA.

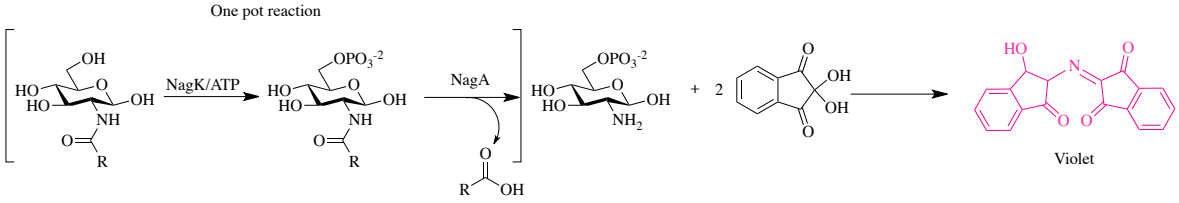


Figure 2. 9. Reaction scheme of the kinetic study. Acylglucosamine is first phosphorylated by NagK and then is deacetylated by NagA. Ninhydrin assay allows to study the kinetic parameters.

Absorbance was plotted against time and linear regression was calculated to get the slope (V_0), which corresponds to the initial velocity of the reaction. V_0 data was plotted against substrate concentration and fit to the Michaelis-Menten enzyme kinetics non-linear model (fig 2.10 and 2.11) using **GraphPad Prism** version 5.00 for Mac OS X, (**GraphPad Software**, San Diego California USA, www.**graphpad**.com). K_m and k_{cat} were calculated for every substrate.

Figure 2.10 compares all the substrates analyzed. *N*-acetylglucosamine, the natural substrate of NagA shows the highest activity of all the sugars, followed by GlcNPr, where

the activity decreases 50%. The reduction on the enzymatic activity is due to the increment in one carbon in the acyl chain of the sugar. The rest of the sugars shows a very low activity and are found at the bottom of the graph.

When the bottom range of figure 2.10 is expanded, we can see the different activities shown by the rest of the sugars. GlcNiBu shows the highest activity among the rest of the analogs, this may be due, to the reduction in length of the side chain similar to GlcNPr, even though carbon 2 is substituted.

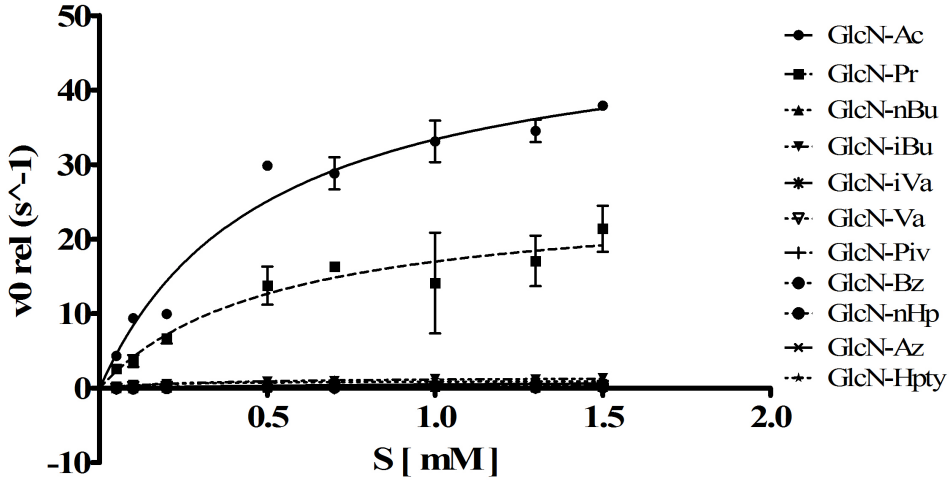


Figure 2. 10. Michaelis-Menten model for kinetic analysis of GlcNAc and all substrates

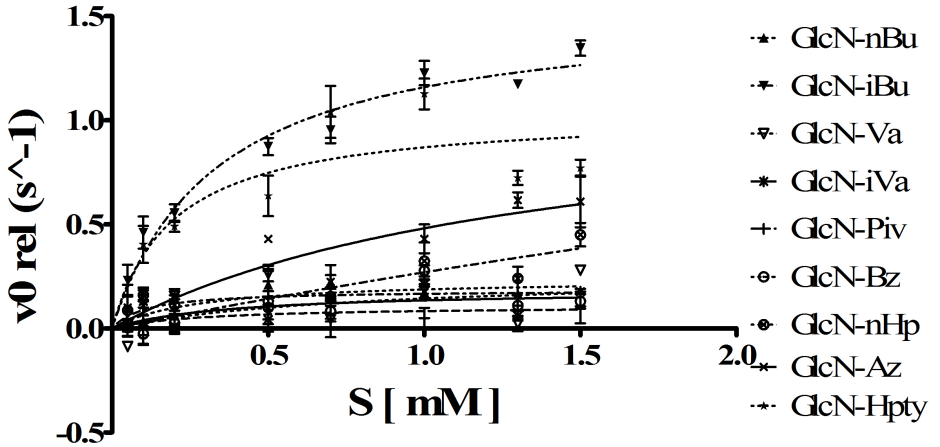


Figure 2. 11. Zoom of the kinetic analysis for the substrates with a low enzymatic activity.

Table 2.2 shows the kinetic parameters for NagA with different GlcNAc analogs. When NagA is incubated with GlcNAc, the K_m value is 0.5±0.1 mM at 37 °C and pH 7.4. According with White and Pasternak[10], when the enzymatic reaction is performed at 37 °C and pH 8.0, K_m value is 0.8 mM and Souza *et al*.[7] report that at 30 °C and pH 7.5 K_M is 0.3±0.02 mM. Thus, the K_m determined in this study is well in accordance with literature values.

GlcNAc is the fastest substrate to by hydrolyzed followed by GlcNPr, however, the maximal reaction velocity in this case decreases 50%, as seen in figure 2.10. The rest of the substrates present a very low maximal velocity based on the k_{cat} values.

The K_m values across the panel of substrates do not vary substantially. All the values are similar in the high micromolar to low millimolar range. Many have large error ranges making it impossible to draw any detailed conclusions on relative affinity for the enzyme. Thus, in general NagA appears to have similar affinity for the acylated aminosugars used in this study.

However, the efficiency of NagA shows a very clear behavior, the value of the specificity constant (k_{cat}/K_m) decreases with the size of the acyl group. NagA's natural substrate presents the highest value. Previous works[5,7,11] reports a different values (1.3 x 10^6 $mM^{-1}s^{-1}$) for NagA with GlcNAc, which is slightly greater than the error range for the value determined in this, however the differences are likely due to a change in assay conditions between our assays and those in the literature. It is important to highlight that Hall *et al*. over expressed NagA and after purification, he used EDTA to eliminate the metal ion from the enzyme. Once NagA was inactivated, they incubated it with 1 eq of Zn^{+2} to be sure that the enzyme had only this ion. In addition, these studies were done at

30 °C. In our case, we over expressed NagA without any Zn^{2+} supplement and worked at 37 °C. Thus, we expect our data to be somewhat different that the data presented in Hall's study, though of similar order of magnitude.

The specificity constant, $k_{cat}/K_{m,}$ shows a very clear tendency to decrease as the acyl chain increases in size and bulk. As seen before, and corresponding with the previous results, GlcNiBu provides a larger specificity constant that the other substrates. this could be due to the better fit of the acyl group in the cavity of the enzyme resulting in a better positioning of the sugar, however, the efficiency value is much lower than the GlcNAc and even the GlcNPr.

In summary, the aminosugars appear to be able to bind NagA, however, when as the acyl group increases in size, the reaction velocity (turnover number) goes down and the efficiency of conversion into the hydrolyzed product decreases. For this reason, from a kinetic point of view, the substrates used for the biosynthesis of their corresponding sialic acids should have an acyl group around four carbon atoms in length, however, we have to consider the possibility that substituted acyl groups could be hydrolyzed but a much lower rate.

Table 2. 2. Kinetic parameters of NagA using different *N*-acyl-D-glucosamine substrates.

Substrate	k_{cat} (s^{-1})	K_m (mM^{-1})	k_{cat}/K_m ($mM^{-1}s^{-1}$) x 10^2	R^2
GlcN-Ac	50±4	0.5±0.1	1010±220	0.94
GlcN-Pr	26±6	0.5±0.3	497±310	0.64
GlcN-nBu	0.24±0.08	0.3±0.3	8.47±10.40	0.26
GlcN-iBu	1.55±0.09	0.3±0.1	45.9±8.6	0.92
GlcN-Va	0.18±0.05	0.1±0.1	17.1±22.5	0.17
GlcN-iVa	0.18±0.06	0.4±0.4	4.89±5.29	0.24
GlcN-Piv	0.11±0.05	0.3±0.5	3.58±6.24	0.17
GlcN-Bz	0.3±0.2	0.8±1.0	3.23±6.25	0.28
GlcN-nHp	3±9	9±34	3.00±15.7	0.36
GlcN-Az	1.1±0.5	1±1	8.28±6.42	0.69

2.2.2.2 In silico docking experiment

The results of the kinetic study are very promising. To gain a better understanding of how the acyl group impacts the conversion of the enzyme-substrate complex into product we collaborated with James Davey from the Chica laboratory (U. Ottawa). He developed a docking experiment that allowed us to model different substrates in the enzyme active site and measuring several parameters such as dihedral angle, distance to the zinc ion, among others.

Based on the crystal structure found in the Protein Data Bank website (http://www.rcsb.org/pdb/home/home.do) and the information reported by Ferreira *et al.*[11] and Hall *et al.*[12], the active site cavity was modeled. For this, the amino acid sequence

related to the active site was used to outline the cavity. Then, the rotamers of the amino acids are calculated to give the total volume of the cavity (fig 2.12 A, B, C).

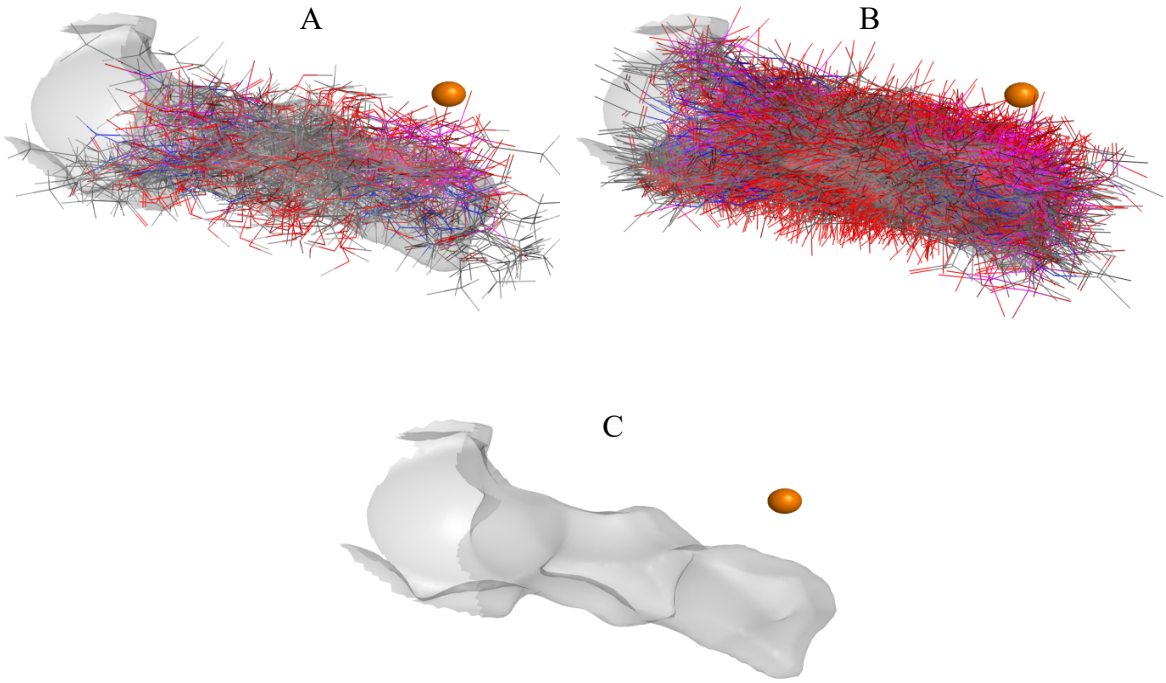


Figure 2. 12. Modeling of NagA active site cavity. A; amino acids backbones involved in the active site. B; spatial distribution of amino acid rotamers of the active site. C; active site model.

Once the binding site was modeled, the different substrates were docked, and distances and angles were measured. Table 2.3 shows the values of the angles between the Zn^{2+} ion and the carbonyl group and the dihedral angle between the plane formed by the Zn-O-C and the plane formed by O-C-N (fig 2.13).

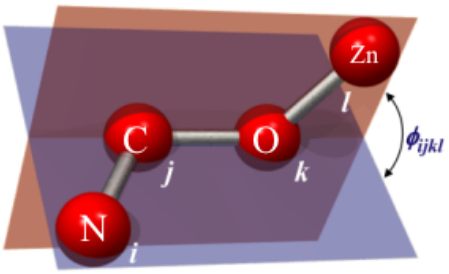


Figure 2. 13. Representation of the dihedral angle formed between the sugar molecule and the Zn^{2+} ion. Image taken and adapted from http://cbio.bmt.tue.nl/pumma/index.php/Theory/Potentials.

Table 2. 3. Geometry parameters found between the carbonyl group of the substrates and the Zn^{2+} ion in NagA active site

Substrate	Zn-C=O distance (Å)	Zn-C=O angle	Dihedral angle
GlcNAc	2.58	123.6	86.1
GlcNPr	2.61	118.4	89.4
GlcNnBu	2.44	127.8	102.9
GlcNiBu	2.65	119.3	91.4
GlcNPiv	2.46	162.7	138.9
GlcNBz	2.75	127.7	171.7
GlcNAz	2.52	107.4	82.8

The theoretical values for the geometry in the active site are 90 degrees for the dihedral angle and 120 degrees for the angle between the Zn and the carbonyl group. For GlcNAc the distance between Zn^{2+} and the oxygen of the carbonyl is 2.58 Å, the angle formed between the carbonyl and the Zn^{2+} has a value of 123.6, and the dihedral angle is 86.1. These values are very comparable to theory. When the acyl group increases by one methylene group (GlcNPr) the distance to the Zn^{2+} slightly increases, this corresponds to the data from in the kinetic assay, where the GlcNPr leads to a 50% decrease in enzymatic activity. Besides the change in distance, the angles change too. The dihedral angle gets closer to 90 degrees, however the Zn^{2+} angle decreases. The difference in these values is small suggesting that the enzyme should be active with this substrate, as is observed.

When we compare the GlcNiBu the distance to the Zn^{2+} increases, nevertheless the angles get closer to the theoretical value. The kinetic analysis shows a considerably reduction in the enzymatic activity with this substrate, the in-silico model could explain

this due to the fact that the distance to the Zn^{2+} increases although the angles match the theoretical values. This suggests that for NagA to have certain activity the angles should be close to 90 degrees for the dihedral and 120 degrees for the carbonyl-Zn^{2+} and the distance to the Zn^{2+} should be the determinant factor in the activity. This can be confirmed when we compare the rest of the substrates. In those cases, the angles values are too far from the theoretical values, for this reason, the enzymatic reaction does not take place.

An additional docking experiment was performed to analyze the potential energy of interaction and conformational strain in two binding modes, the correct binding mode (carbonyl towards the Zn^{2+} ion) and flipped binding mode (carbonyl towards the entrance of the cavity), for the docked ligands. Table 2.4 shows the results of this analysis.

For this experiment, ligands were bound in all possible conformations in the active site and the energy was calculated for each conformation. Values were ranked and the table presents the best-ranked conformer and the energy associated with it.

Table 2. 4. Rank, relative potential energy of interaction (Eint) and relative conformational strain (Econf) for correct and flipped binding modes of docked ligands. Values are normalized against the natural substrate (GlcNAc) energy values

Ligand	Reactive Binding Mode			Flipped Binding Mode		
	Rank	Eint	Econf	Rank	Eint	Econf
inhibitor	1/179	1.26	0.89	5/179	1.20	0.86
acetyl	1/178	1.00	1.00	2/178	1.00	1.00
propyl	1/173	1.03	1.02	2/173	0.99	0.95
t-butyl	79/168	0.81	1.42	1/168	0.99	1.30
azide	73/188	0.91	1.24	1/188	1.07	1.12
n-butyl	48/179	0.87	1.07	1/179	1.00	0.92
i-butyl	2/180	1.02	1.12	1/180	1.01	1.05
benzyl	134/156	0.62	1.59	1/156	1.03	1.49

Rank indicates the occurrence of the specified ligand-binding mode in the set of all binding modes ranked by a score value calculated as the sum of Eint and Econf. Eint and Econf, calculated using the MMFF94s force field and a general born implicit solvation model, reported with units kcal/mol.

Table 2.4 shows the interaction and configuration energies of the different substrates when they are normalized against the calculated energy of GlcNAc, the natural substrate. This comparison allows having a clear idea of which substrates can be hydrolyzed by NagA. When the relative energy of interaction is analyzed, we can see that GlcNPr and GlcNiBu present relative energies similar to the natural substrate. These results confirm the findings obtained in the kinetics experiment. On the other hand, large and voluminous acyl groups present relative energies that distant from the normalized value for GlcNAc.

Figure 2.14 shows the interaction of the acylated amino sugar in the binding pocket. Figure A represents GlcNAc inside the active site and we can see the position of the sugar in the correct docking mode. When we move to figure 2.14. B, the full color structure is GlcNPr and in purple we can see the original position of GlcNAc. In this case, the methyl group of the acyl chain can rotate and fits in the far side of the active site. This extra carbon displaces the carbonyl from the original GlcNAc location. GlcNiBu (fig 2.14 C) acyl group, like GlcNPr, can accommodate the two methyl groups in the active site, however, the carbonyl of the amide group moves further from the Zn^{2+}. In the last case (figure D), the benzene ring is so big that locates the carbonyl in a completely different position. This figure exemplified why NagA can hydrolyze the small substrates, while the bigger substrates result in a loss of the activity of the enzyme.

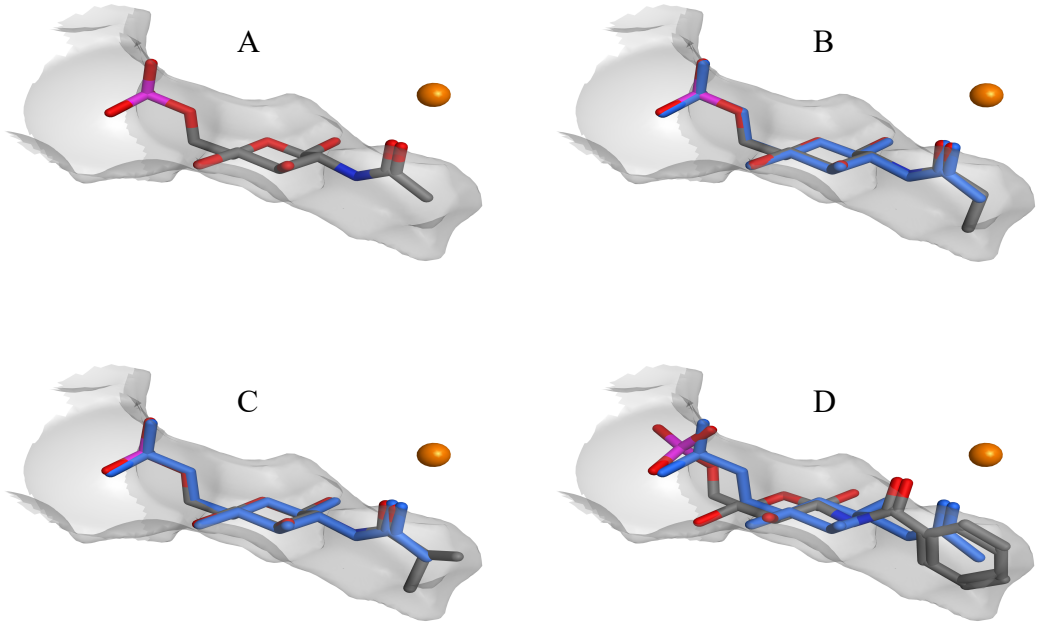


Figure 2. 14. Docking modeling of acyl-aminosugars in NagA's active site. A, GlcNAc in correct binding mode; B, GlcNPr is showed in full color and GlcNAc is represented in purple; C, GlcNiBu docking model; D GlcNBz docking model.

In summary, in silico docking experiments suggest that NagA can hydrolyze on N-acyl aminosugars with a small acyl groups and when these groups become bulkier, the sugar binds to the active site in a flipped orientation that prevents hydrolysis from occurring. These findings confirm the kinetic characterization of NagA, suggesting at this point that NagA could only hydrolyze N-acyl aminosugars with a small acyl group.

2.2.2.3 In vivo assay

Within vitro and in silico experiments both suggesting that NagA could hydrolyze N-acyl aminosugars that contain small acyl groups, we focused on collecting corresponding in vivo data. If NagA is able to hydrolyze the amide bond of N-acylglucosamines, then the glucosamine liberated can enter into central metabolism and potential provide carbon source for bacterial survival. Thus, we hypothesized that *Escherichia coli* could only survive on N-acylglucosamines as the sole carbon source if NagA was competent to hydrolyze them.

For the *in vivo* experiment, three strains were transformed. *E. coli* BL21 (DE3) was transformed with pLRVP08 (*nagA* on pET21). This strain overexpresses NagA what increase the concentration of the enzyme in the cell. This trial allows us to study the reaction under high concentrations of enzyme. *E. coli* BL21 (DE3) transformed with pET21 as test subject (normal levels of NagA) and BRL04[13,14] [(λ(DE3) *nanT::Tn5KAN-I-SceI nanA::TetR* Δ*nagA*)] transformed with pET21 as negative control. Three colonies of each strain were selected and suspended in 100 μL sterile water.

F1 minimum media supplemented with 50 μg/mL of ampicillin were made to test the activity *in vivo*. The minimum media was enriched with trace metal solution, IPTG 0.5 mM final concentration, and *N*-acylglucosamine (analog) 0.1% as carbon source.

The plate was divided in 9 parts where 5 μL of each repetition was stroke. The plates were incubated at 37˚C for 72 hours. Pictures were taken at 24, 48 and 72 h.

Figure 2.15 shows one of the control plates, where the samples were inoculated on a LB medium plate to corroborate the viability of the samples. As is demonstrated in the figure, all the samples can grow in a rich media.

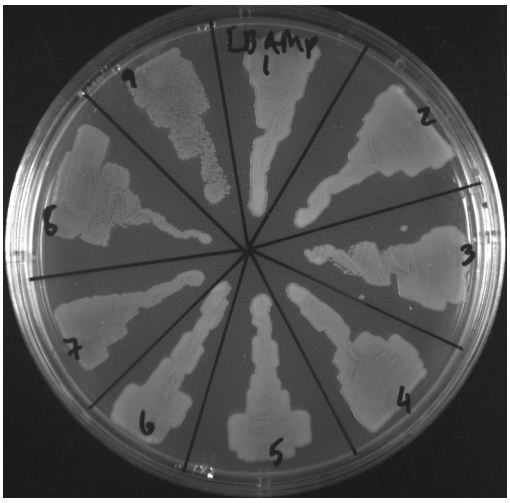


Figure 2. 15. LB plate inoculated with *E. coli* BL21 transformed with pLRVP08 (1-3), *E. coli* BL21 transformed with pET21 (4-6) and *E. coli* BRL04 transformed with pET21 (7-9).

To confirm that hydrolyzed GlcNAc ($GlcNH_2$) could be used as a carbon source, the strains were inoculated in a plate with $GlcNH_2$ as carbon source. Figure 2.16 shows the growth of the samples on $GlcNH_2$ plate confirming that BL21 transformants could readily grow. As high concentrations of $GlcNH_2$ are known to be growth inhibitor, long incubation times were required to observe significant growth. The strain BRL04, with its multiple gene deletions is further growth impaired and shows little to no growth even after 72h.

This was not considered a significant problem, since we anticipated that NagA mediated hydrolysis would slowly release $GlcNH_2$, limiting its growth inhibitory effect.

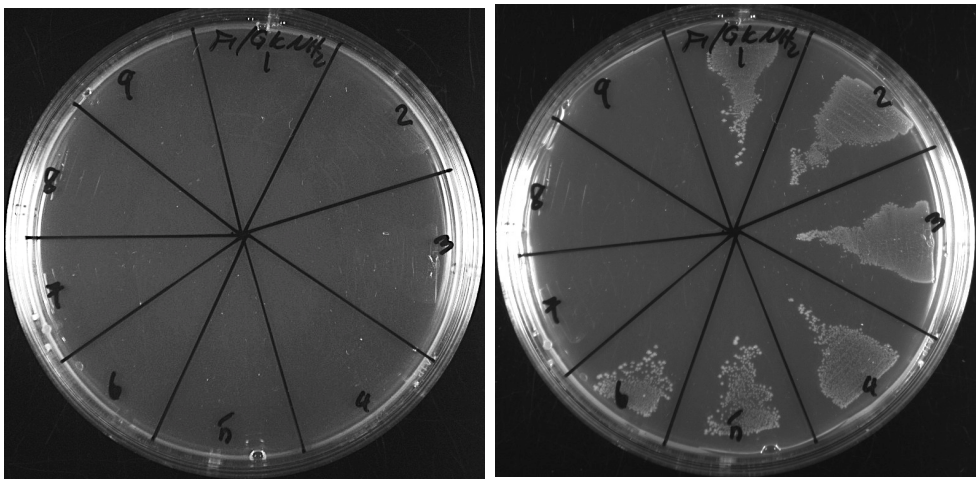


Figure 2. 16. F1 minimum media enriched with 0.1% of glucosamine as carbon source. Left; 24 hours after inoculation. Right; 72 hours after inoculation.

In order to corroborate the results of the in-silico experiment, the feeding test was performed with the same N-acyl aminosugars used in the computer modeling. Table 2.5 summarizes the results of this assay.

When GlcNAc is used as carbon source, the *E. coli* with pET21 has moderate growth, going from a low growth after 24 hours to substantial growth at 72 hours. As expected, the strain overexpressing NagA showed greater growth.

NagA hydrolyzes the amide bond in N-acetylglucosamine-6-phosphate giving glucosamine-6-phosphate as primary product. This glucosamine-6-phosphate undergoes a deamination and an isomerization. The final product of these reactions is fructose-6-phosphate can be used in central metabolism as a carbon source. The assay with GlcNAc shows that in normal conditions, *E. coli* can deacetylate GlcNAc and use it to grow.

When GlcNPr growth is analyzed, we can see a subtle decrease in the colony growth. The first 24 hours do not show any colony growth in the test triplicate, however,

at 48 hours, the bacterial growth develops to an excellent point. The positive control in this case, shows a faster development, as the growth begins at 24 hours.

Butyl analogs present an interesting behavior, both substrates produce no growth in replicates 4 to 6, meaning that NagA cannot hydrolyze them to produce glucosamine, thus, no carbon source is available for cell growth. Nevertheless, when we compare the growth of the positive control, we can see a moderate growth after 48 hours that remain constant until 72 hours of incubation. This growth can be due to the high concentration of NagA in the cells. It is very important to notice that GlcN*i*Bu presents poor growth at 72 hours (fig 2.17), something that does not happen with its lineal counterpart. This result confirms our findings in previous experiments, where small and less bulky acyl groups present some degree of hydrolysis. The same result is obtained when GlcNAz is used as carbon source in the plates.

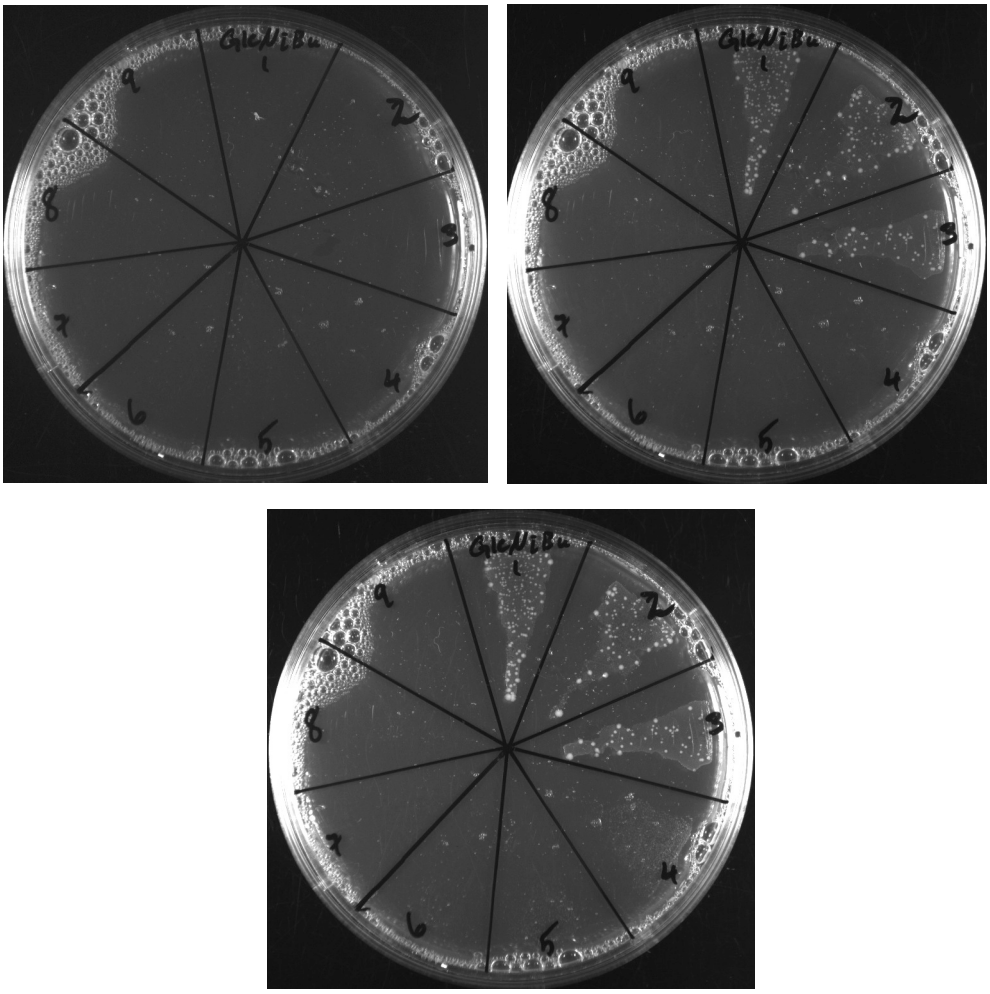


Figure 2. 17. Results of the feeding experiment using N-*i*-butylglucosamine as carbon source. Top left, 24 h; top right, 48 h; bottom, 72 h.

Moreover, when the phenyl analog is used, the result of the test is negative in all trials confirming our previous findings.

Table 2. 5. Qualitative results of feeding experiment. BL21* is the positive control, BL21 is the test and BRL04 is the negative control. The scale of measurement is: - no growth; + poor growth; ++ moderate growth; +++ excellent growth.

Carbon source	Strain	24 h	48 h	72 h
GlcNAc	BL21*	++	+++	+++
	BL21	+	++	+++
	BRL04	-	-	-
GlcNPr	BL21*	+	+++	+++
	BL21	-	+++	+++
	BRL04	-	-	-
GlcN*n*Bu	BL21*	-	++	++
	BL21	-	-	-
	BRL04	-	-	-
GlcN*i*Bu	BL21*	-	++	++
	BL21	-	-	+
	BRL04	-	-	-
GlcNBz	BL21*	-	-	-
	BL21	-	-	-
	BRL04	-	-	-
GlcNAz	BL21*	-	+	+
	BL21	-	-	+
	BRL04	-	-	-
LB	BL21*	+++	+++	+++
	BL21	+++	+++	+++
	BRL04	+++	+++	+++

Figure 2.1 shows that NagA works on phosphorylated GlcNAc. In the kinetic study, we used a kinase (NagK) in order to phosphorylate the glucosamine analogs, although, we have no certainty that the sugars are totally phosphorylated, thus, the importance of this in vivo assay.

Escherichia coli incorporate sugars via Phosphoenolpyruvate-carbohydrate Phosphotransferase System (PTS)[15]. This transport system allows the incorporation and simultaneous phosphorylation of the sugars. Under this premise, we hypothesize that the sugar would be incorporated and phosphorylated by the PTS into the cell and then can be used by *E. coli*. If NagA hydrolyze the acyl group, the resulting glucosamine can be used in central metabolism and the bacteria could grow[16].

The phosphorylation process mediated by PTS begins with the phosphorylation of EI by PEP. Phosphorylated EI transfer the phosphate moiety to HPr protein that will transfer it to EII complex (this protein complex is formed by EIIA, EIIB and EIIC) and EIIC is the membrane protein responsible for the translocation and later phosphorylation of the sugar[17]. There are several types of PTS that can transport different sugars. The glucose PTS family can translocate glucose, glucosamine, N-Acetylglucosamine, among others. However, there is not enough information about the specificity of these membrane proteins[17–19].

Our assumption is that the EIIC protein is promiscuous enough to translocate the different glucosamine analogs and phosphorylate them. If we look close to the results obtained with GlcNnBu we can see how the test replicates (4 to 6) present no growth after 72 hours of incubation. This result could mean two things, first one, that the sugar is not incorporated by PTS, thus the absence of colonies; second, that the sugar is incorporated

by PTS, but the bacteria cannot grow because NagA is unable to hydrolyze the substrate and subsequently cannot be incorporated to central metabolism.

Nonetheless, when GlcNiBu was used as carbon source, colonies begin to grow at 72 hours (fig 2.17). The fact that bacterial colonies grow in this test confirms that the sugar analog is incorporated into the cells, not necessarily by PTS, but if the bacteria can grow is because is using the sugar in the culture media. Besides, the other reason why *E. coli* can grow in this test is because NagA can hydrolyze the i-butyl group and generate fructose that can enter central metabolism. Moreover, based on figure 2.1, the only way in which NagA can hydrolyze a substrate, is if the substrate is phosphorylated, for these reasons we can conclude that the N-acyl aminosugars are, not only incorporated into the cell, but phosphorylated too. Furthermore, when positive controls are analyzed, the bacterial growth begins at 48 hours. Considering that the positive control overexpresses NagA, thus the enzyme concentration is much higher, there is a faster hydrolysis reaction due to the increase of enzyme availability.

Additionally, the growth rate of the bacterial colonies corresponds with the evidence found in the in vitro and in silico experiments. This correspondence allows us to conclude that NagA can hydrolyze N-acylglucosamines, as long as acyl groups are small. Substrates that can undergo hydrolysis, react very slowly, taking between 48 to 72 hours to be incorporated in central metabolism. This is important information regarding sialic acid analogs because the bacterial culture should be fed with the amino sugar analog every 48 hours to assure minimal decomposition of the substrate and a higher yield of the final product.

References

1. Park, J. T. Identification of a Dedicated Recycling Pathway for Anhydro- N -Acetylmuramic Acid and N -Acetylglucosamine Derived from Escherichia coli Cell Wall Murein. *Society* **183**, 3842–3847 (2001).
2. Sugrue, E. *et al.* Evolutionary expansion of the amidohydrolase superfamily in bacteria in response to the synthetic compounds molinate and diuron. *Appl. Environ. Microbiol.* **81**, 2612–2624 (2015).
3. Seibert, C. M. & Raushel, F. M. Structural and catalytic diversity within the amidohydrolase superfamily. *Biochemistry* **44**, 6383–6391 (2005).
4. Vimr, E. & Kalivoda, K. Diversity of microbial sialic acid metabolism. *Microbiol. Mol. Biol. Rev.* **68**, 132–153 (2004).
5. Hall, R. S., Xiang, D. F., Xu, C. & Raushel, F. M. Articles N Acetyl D glucosamine 6 phosphate Deacetylase Substrate Activation via a Single Divalent Metal Ion. *Biochemistry* **46**, 7942–7952 (2007).
6. Xu, C., Hall, R., Cummings, J. & Raushel, F. M. Tight binding inhibitors of N-acyl amino sugar and N-acyl amino acid deacetylases. *J. Am. Chem. Soc.* **128**, 4244–5 (2006).
7. Souza, J. M., Plumbridge, J. a & Calcagno, M. L. N-acetylglucosamine-6-phosphate deacetylase from Escherichia coli: purification and molecular and kinetic characterization. *Arch. Biochem. Biophys.* **340**, 338–46 (1997).
8. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–54 (1976).

9. Starcher, B. A ninhydrin-based assay to quantitate the total protein content of tissue samples. *Anal. Biochem.* **292**, 125–9 (2001).

10. White, R. J. & Pasternak, C. a. The purification and properties of N-acetylglucosamine 6-phosphate deacetylase from Escherichia coli. *Biochem. J.* **105**, 121–5 (1967).

11. Ferreira, F. M. *et al.* Structural analysis of N-acetylglucosamine-6-phosphate deacetylase apoenzyme from Escherichia coli. *J. Mol. Biol.* **359**, 308–21 (2006).

12. Hall, R. S. *et al.* Structural diversity within the mononuclear and binuclear active sites of N-acetyl-D-glucosamine-6-phosphate deacetylase. *Biochemistry* **46**, 7953–7962 (2007).

13. Lundgren, B. R. & Boddy, C. N. Sialic acid and N-acyl sialic acid analog production by fermentation of metabolically and genetically engineered Escherichia coli †. *Org. Biomol. Chem.* 1903–1909 (2007). doi:10.1039/b703519e

14. Horsman, M. E., Lundgren, B. R. & Boddy, C. N. N-Acetylneuraminic Acid Production in Escherichia coli Lacking N-Acetylglucosamine Catabolic Machinery. *Chem. Eng. Commun.* **203**, 1326–1335 (2016).

15. Escalante, A., Cervantes, A. S., Gosset, G. & Bolívar, F. Current knowledge of the Escherichia coli phosphoenolpyruvate-carbohydrate phosphotransferase system: Peculiarities of regulation and impact on growth and product formation. *Applied Microbiology and Biotechnology* **94**, 1483–1494 (2012).

16. Barnhart, M. M., Lynem, J. & Chapman, M. R. GlcNAc-6P levels modulate the expression of Curli fibers by Escherichia coli. *J. Bacteriol.* **188**, 5212–9 (2006).

17. Kotrba, P., Inui, M. & Yukawa, H. Bacterial Phosphotransferase System (ITS) in

Carbohydrate Uptake and Control of Carbon Metabolism. *J. biocience Bioeng.* **92**, 502–517 (2001).

18. Tchieu, J. H., Norris, V., Edwards, J. S. & Saier, M. H. The complete phosphotransferase system in Escherichia coli. *J. Mol. Microbiol. Biotechnol.* **3**, 329–346 (2001).

19. Siebold, C., Flükiger, K., Beutler, R. & Erni, B. Carbohydrate transporters of the bacterial phosphoenolpyruvate: Sugar phosphotransferase system (PTS). in *FEBS Letters* **504**, 104–111 (2001).

2.3 Experimental Section

2.3.1 Plasmid Production

Two strains of chemical competent *E. coli* BL21 were transformed with plasmid pLRVP03 that carries *nagA*. The strain was placed on a glass test tube and incubated on ice for 5 min. 5 µL of plasmid solution were placed in the bacterial suspension and incubated for 30 min on ice.

After the incubation period the test tube was placed in a 42 °C water bath for 45 s and returned to the ice bath for 2 min. 1 mL of LB broth was added to the tube and it was incubated at 37 °C for half an hour; then, the bacterial suspension was inoculated on LB plates supplemented with kanamycin and incubated overnight at 37 °C.

Single colonies were taken and inoculated in 5 mL of LB broth supplemented with kanamycin (50 µg/mL) and incubated for 18 h. After the incubation time, the cells were spun down and the plasmids were extracted using the P1, P2, P3 protocol.

2.3.2 Plasmid Mini-Preps via P1, P2, & P3

1 mL of culture was centrifuged at 16,000 × g in a 1.5 ml Eppendorf tube to pellet cells. Supernatant was discarded and the procedure was repeated 4 more times to collect the cells contained in 5 mL of culture. The pellet was resuspended in 250 µL of P1[2] (250 µL per 5 mL of culture). The cell suspension was mixed thoroughly until no cell clumps were visible in suspension. 250 µL of P2[3] were added (250 µL per 5 mL of culture) and mixed gently

[2] P1: 6.1 g Tris, 3.7 g EDTA-2H2O pH 8.0 w/ HCl/1 liter, add 100 µg/ml RNAse A as needed, usually 10 mg RNase in 100 ml batches, store 4 degrees.
[3] P2: 8.0 g NaOH in 900 ml H2O plus 100 ml of 10 % SDS/1 liter, store R.T.

by inverting the tube 4-6 times to mix[4]. After mixing, 350 µL of P3[5] (350 µL per 5 mL of culture) were added and the tube was inverted immediately, but gently, 4-6 times to mix. Tubes were centrifuged at 16,000 × g for 10 minutes to pellet debris and 700 µL of the supernatant were transferred to a tube containing 700 µL of cold isopropanol and put on ice for 15 min. After the incubation time the tubes were centrifuged at 16,000 × g for 30 minutes to pellet plasmid DNA and the supernatant was discarded.

500 µL of chilled 70% ethanol were added to the plasmid DNA pellet and centrifuged at 16,000 × g for 10 minutes. The supernatant was discarded, and the pellet was dried via air or speed-vac for 20 minutes. The plasmid DNA pellet was resuspended in elution buffer[6].

2.3.2 Gene expression and protein extraction from E. coli

The T7 expression system, e.g., pET (Novagen), is commonly used to generate heterologous (recombinant) proteins. This system is based on the T7 promoter and its cognate T7 RNA polymerase. Gene expression is turned on in the presence of the inducer isopropyl-β-thiogalactoside or IPTG. IPTG induces the transcription of the T7 RNA polymerase gene from the host's chromosome, e.g., BL21 (DE3) *E. coli*. As T7 RNA polymerase becomes more prevalent, transcription from the T7 promoter of the recombinant vector increases. Consequently, the concentration of heterologous protein rises.

---

[4] Solution should become viscous and slightly clear. DO NOT let lysis proceed for more than 5 minutes.
[5] P3: 294 g KAcetate in 500 ml H2O pH to 5.5 with Acetic Acid (~110 ml), bring to 1 liter, store R.T.
[6] EB: 10 mM Tris, pH 8.5

To produce the seed culture one single colony of the transformants was used to inoculate 10 mL of LB broth supplemented with kanamycin (50 μg/mL). The culture was incubated at 37°C and 200 rpm overnight. Next day, 400 mL of fresh LB were supplemented with 400 μL of 1000 × kanamycin solution (final concentration 50 μg/mL) and inoculated with the seed culture. The bacterial culture was incubated at 37 °C with shaking until OD_{600} reaches 0.3-0.6.

When the proper optical density was reached, IPTG 1M was added to a final concentration of 0.5 mM and the culture was placed in the incubator at 30 °C at 200 rpm.

After 12 to 18 hours of incubation, the culture was centrifuged at 5000 × g for 20 min. at 4 °C. Cells were resuspended in 50 mL of lysis buffer[7] and placed in an ice bath. The cells were lysed by sonication using 5 pulses of 20 s with 30 s intervals of cooling.

The cell suspension was centrifuged at 9000×g for 1 h and the supernatant was transferred to a clean falcon tube.

The recombinant proteins were purified by Ni^{2+}-NTA affinity chromatography.

2.3.3 Nickel purification for 6X-His tagged proteins

For our lab purposes, proteins are generally expressed using pET-vectors (Novagen). These proteins are equipped with a 6X-His tag, which adopts a conformation that is optimal for coordination to divalent nickel ions, Ni^{2+}. We use Ni-NTA superflow (Qiagen) for the resin, and the amount of this resin to the lysate containing the tagged protein should be determined empirically.

---

[7] Lysis buffer = 100 mM sodium phosphate, 300 mM NaCl, 10% (v/v) glycerol, 1 mg/mL lysozyme, 1 μg/mL pepstatin A, 1—2 μg/mL leupeptin, pH 8.0

The lysate from the protein production was incubated with nickel resin, Ni-NTA superflow, for 1-4 hours at 4 °C with gentle rocking/shaking. After the incubation, the mixture was poured into a column and the flow-through was collected and identified. The column was washed with 10 × resin volume using elution buffer[8] 0 mM imidazole and the fraction was collected and labeled 0 mM. Then the column was wash/eluted with increasing concentration of imidazole (10× 20 mM, 5× 100 mM twice and 5× 250 mM twice). The fractions were kept in ice to prevent protein denaturation and the fractions were analyzed by SDS-PAGE.

2.3.4 Deacetylation reaction conditions and ninhydrin assay[9] for kinetic analysis

Reagents

4 N sodium acetate buffer: Dissolve 27.9 g of sodium acetate 3 H_2O in 3.4 mL of glacial acetic acid and add water up to 100 mL

Stannous chloride solution: dissolve 50 mg of $SnCl_2$ 2 H_2O in 500 µL of ethylene glycol.

Ninhydrin reagent: Dissolve 200 mg of ninhydrin in a mixture of 7.5 mL of ethylene glycol and 2.5 mL of 4 N acetate buffer. Add 250 µL of stannous chloride solution and mix well. The reagent should be pale red in color.

Assay: 10 µL of sample or standard are mixed with 100 µL of ninhydrin reagent. Put the reaction mixture in a boiling water bath for 10 minutes and read absorbance at 575 nm.

Deacetylation reaction conditions:

---

[8] Solutions: 100 mM Tris, 300 mM NaCl, X mM imidiazole (X = 0, 20, 100, 250), pH 7.4

Adjust to pH 7.4 after the addition of imidiazole.

Total volume 200 µL, 50 mM phosphate buffer pH 7.5, 50 mM NaCl, 1.2 mM ATP, substrate concentrations from 0 to 1.5 mM, NagA 0.04 µM, 100 µg of NagK. Reactions were done in triplicate and incubated in a water bath at 37 °C. 10 µL samples were taken at 0, 2, 5, 8 and 10 min for each concentration.

Annex

***nagA* sequence**

ATGTATGCATTAACCCAGGGCCGGATCTTTACCGGCCACGAATTTCTTGATGA

CCACGCGGTTGTTATCGCTGATGGCCTGATTAAAAGCGTCTGTCCGGTAGCGG

AACTGCCGCCAGAGATCGAACAACGTTCACTGAACGGGGCCATTCTCTCCCC

CGGTTTTATCGATGTGCAGTTAAACGGCTGCGGCGGCGTACAGTTTAACGAC

ACCGCTGAAGCGGTCAGCGTGGAAACGCTGGAAATCATGCAGAAAGCCAAT

GAGAAATCAGGCTGTACTAACTATCTGCCGACGCTTATCACCACCAGCGATG

AGCTGATGAAACAGGGCGTGCGCGTTATGCGCGAGTACCTGGCAAAACATCC

GAATCAGGCGTTAGGTCTGCATCTGGAAGGTCCGTGGCTGAATCTGGTAAAA

AAAGGCACCCATAATCCGAATTTTGTGCGTAAGCCTGATGCCGCGCTGGTCG

ATTTCCTGTGTGAAAACGCCGACGTCATTACCAAAGTGACCCTGGCACCGGA

AATGGTTCCTGCGGAAGTCATCAGCAAACTGGCAAATGCCGGGATTGTGGTT

TCTGCCGGTCACTCCAACGCGACGTTGAAAGAAGCAAAAGCCGGTTTCCGCG

CGGGGATTACCTTTGCCACCCATCTGTACAACGCGATGCCGTATATTACCGGT

CGTGAACCTGGCCTGGCGGGCGCGATCCTCGACGAAGCTGACATTTATTGCG

GTATTATTGCTGATGGCCTGCATGTTGATTACGCCAACAT

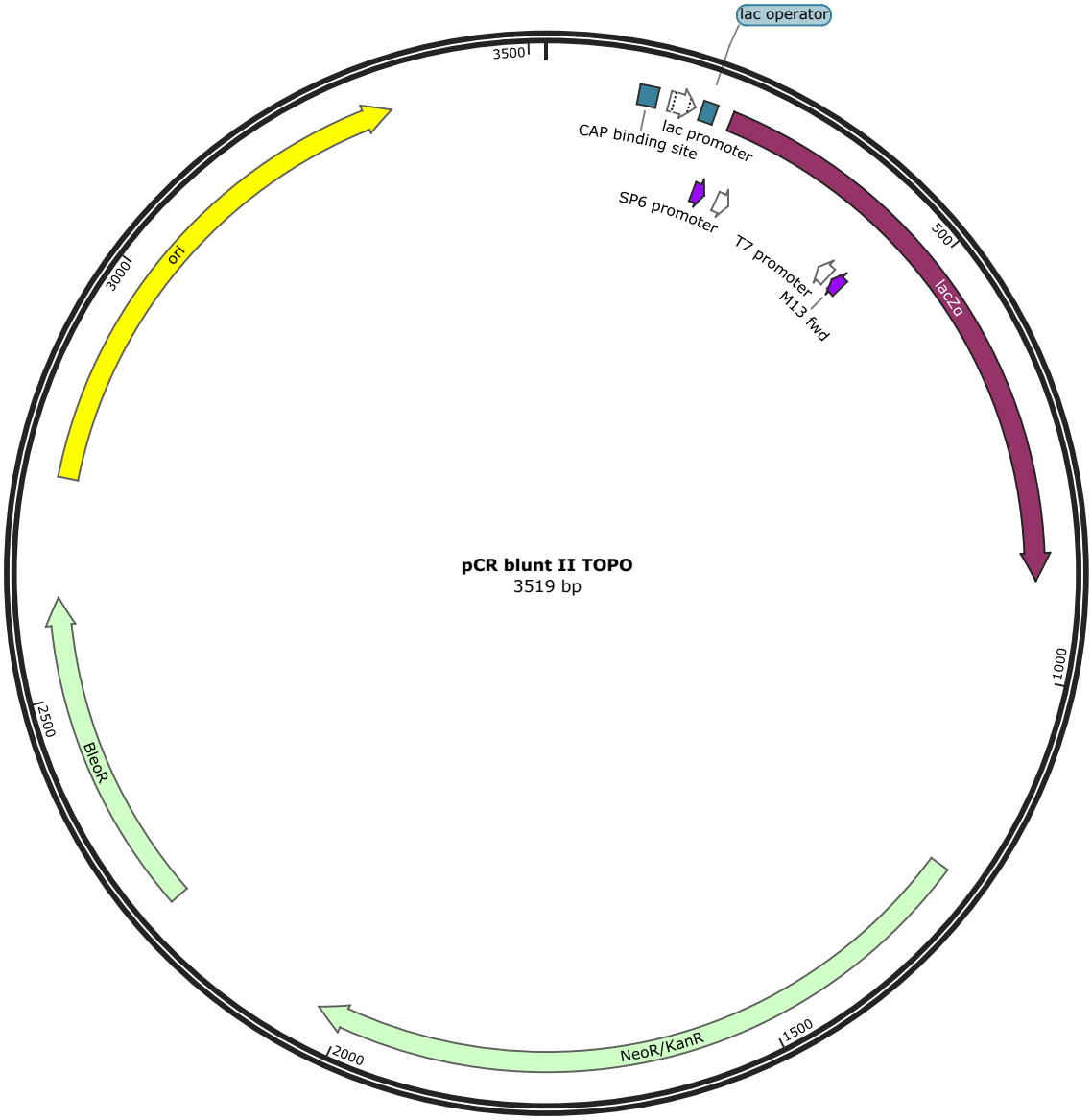
Primers for nagA amplification:

forward *C AGT CAT ATG TAT GCA TTA ACC CAG GGC*,

reverse *C AGT GAA TTC TTA TTG AGT TAC GAC CTC GTT A*.

Plasmids maps

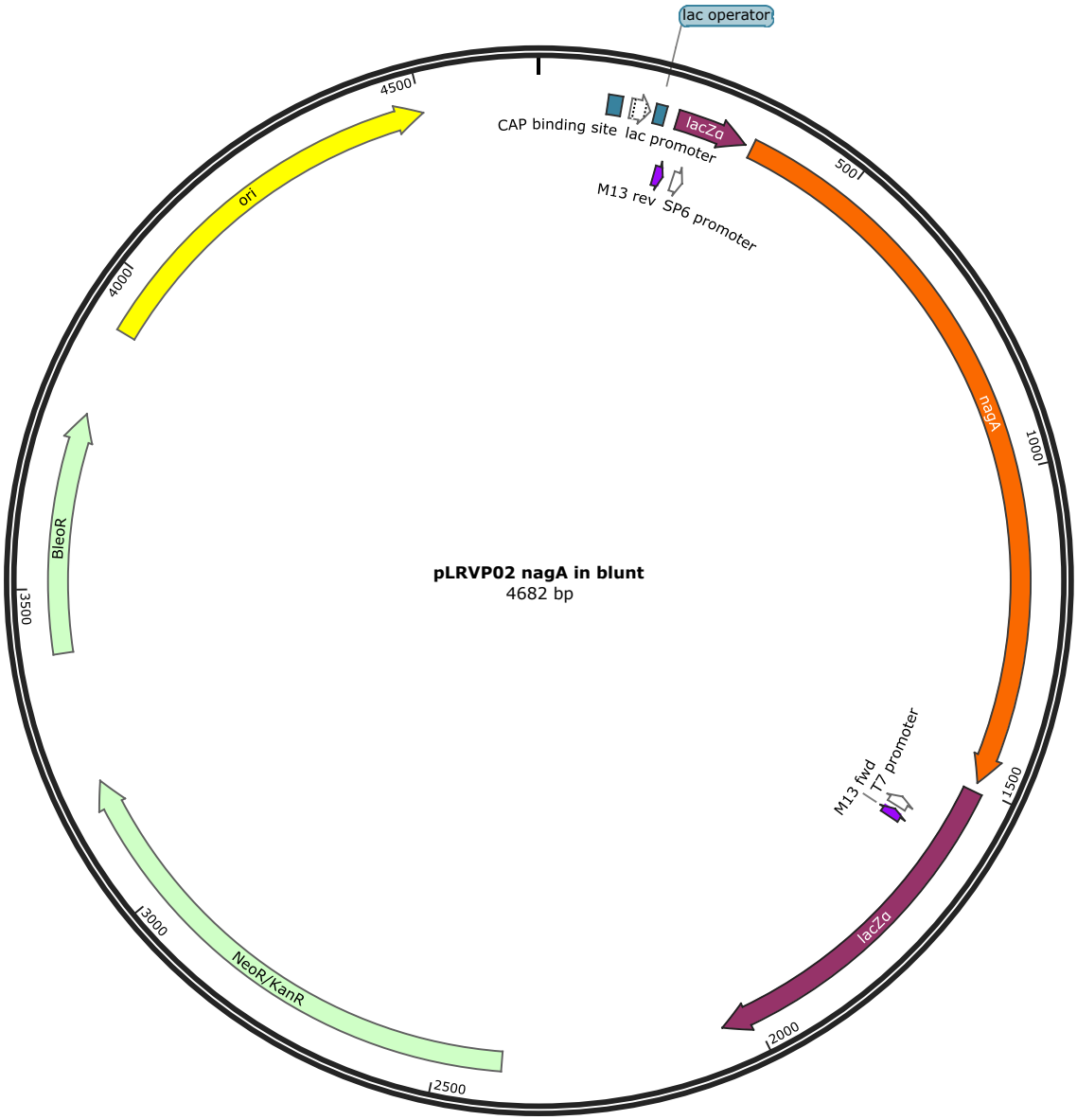
Sequence: pCR blunt II TOPO.dna (Circular / 3519 bp)
Features: 11 visible, 11 total



Printed from SnapGene® Viewer: Nov 14, 2016 9:56 PM Page 1

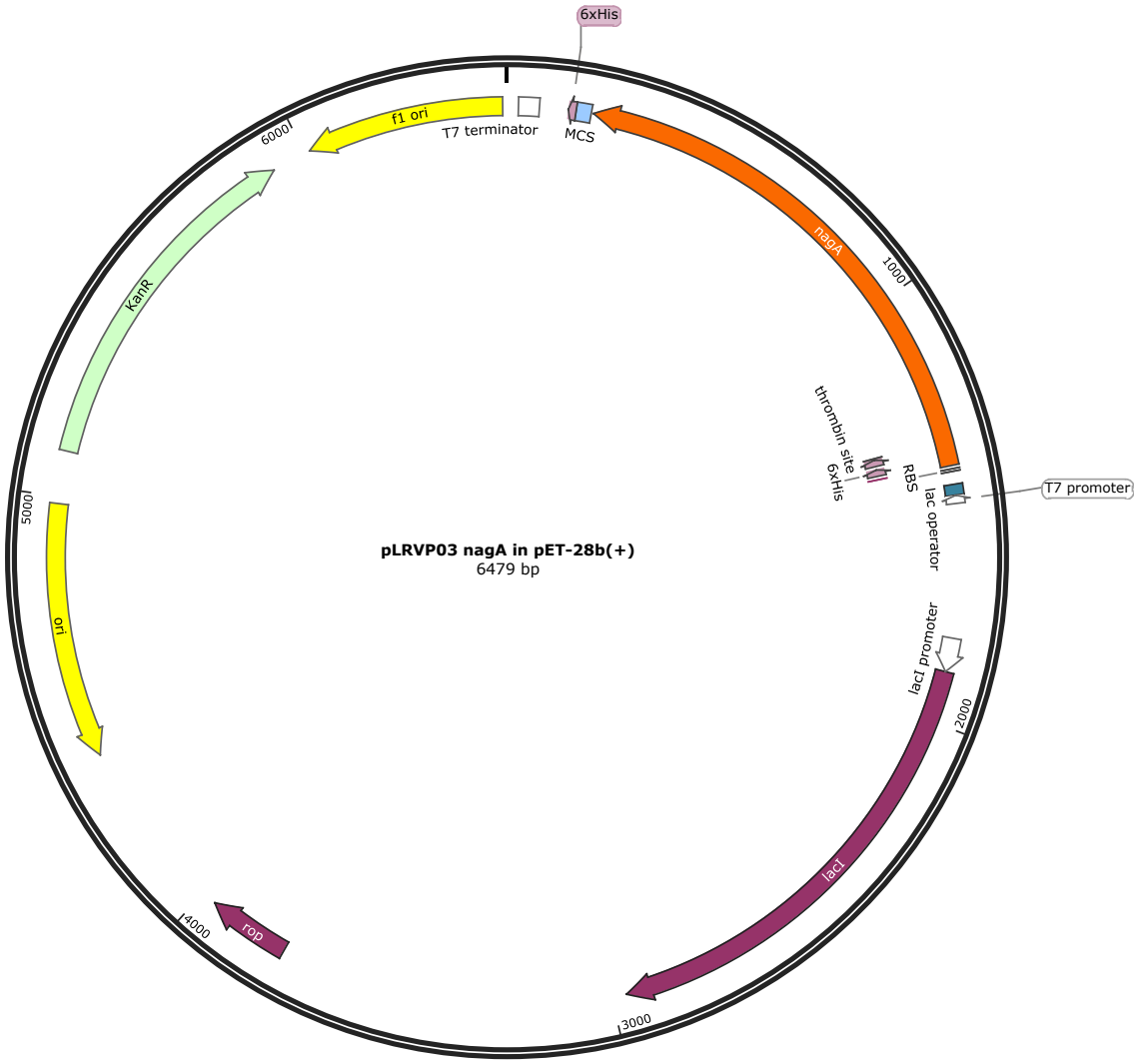
Plasmid map of blunt system

Sequence: pLRVP02 nagA in blunt.dna (Circular / 4682 bp)
Features: 13 visible, 13 total



Plasmid map of pLRVP02, *nagA* in blunt

Sequence: pLRVP03 nagA in pET-28b(+).dna (Circular / 6479 bp)
Features: 17 visible, 17 total



Plasmid map of pLRVP03, *nagA* in pET28b(+)

Sequence: pLRVP03.dna (Circular / 6479 bp)
Features: 16 visible, 16 total

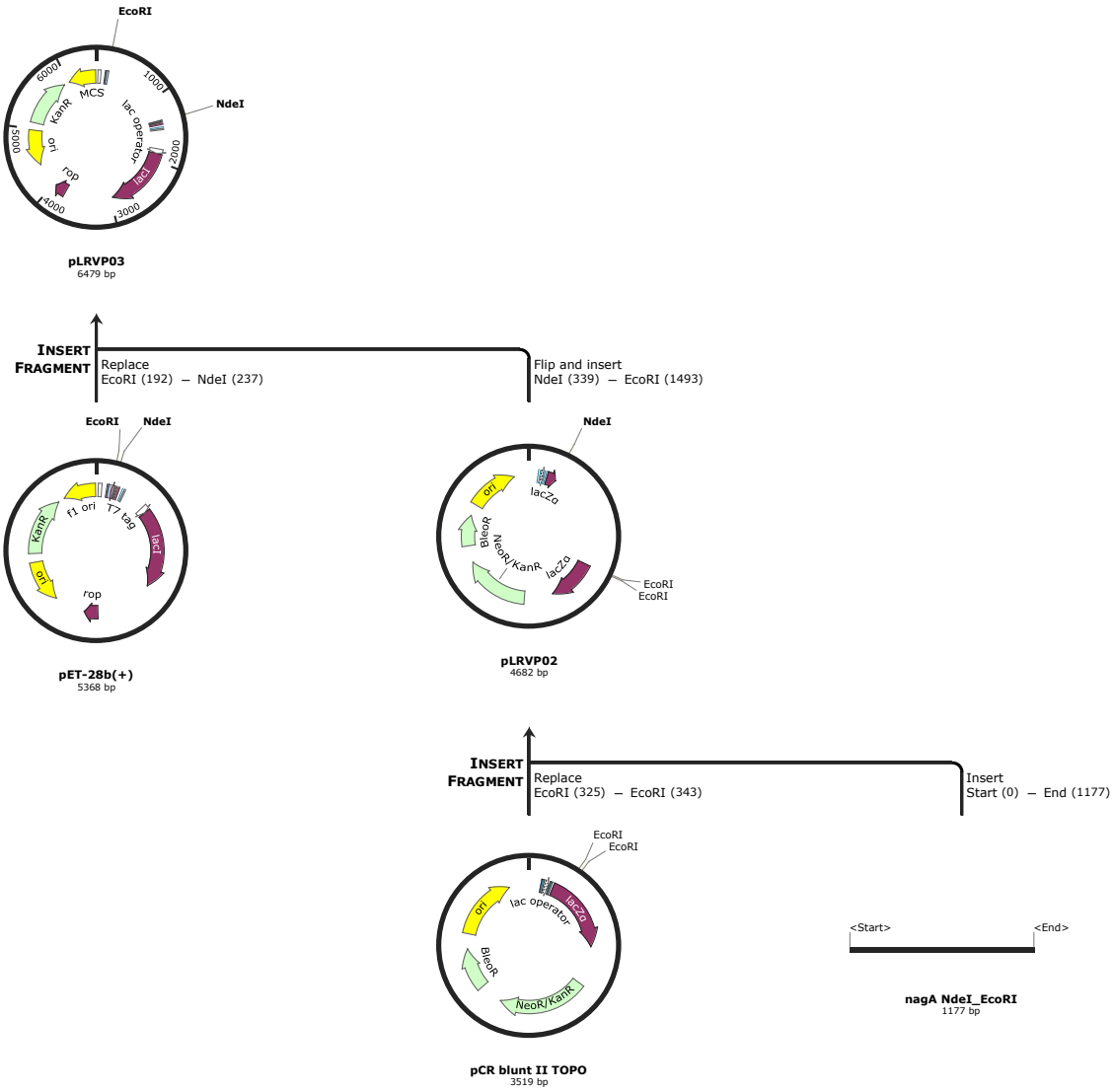


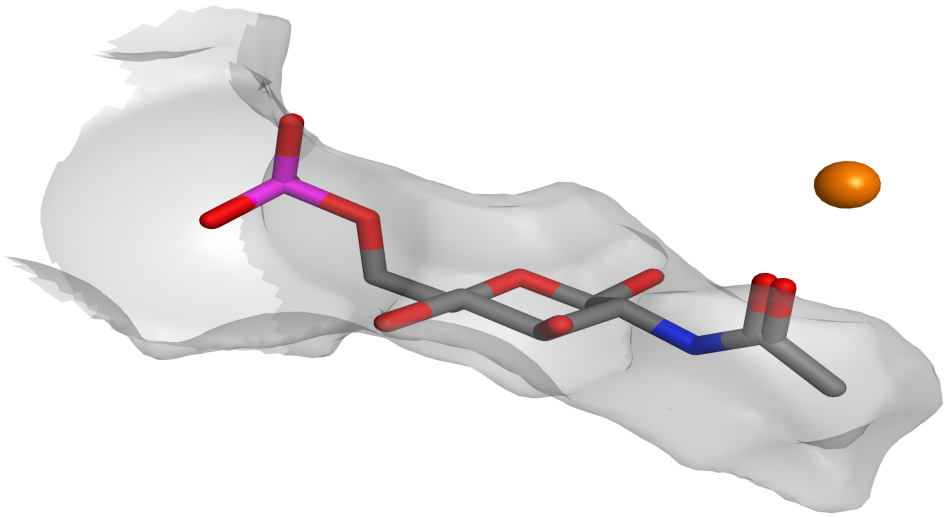
Diagram of the construction of plasmid pLRVP03

Ligand	Reactive Binding Mode			Flipped Binding Mode		
	Rank	Eint	Econf	Rank	Eint	Econf
inhibitor	1/179	-65.3	37.9	5/179	-57.7	35.0
acetyl	1/178	-51.9	42.7	2/178	-47.9	40.7
propyl	1/173	-53.6	43.7	2/173	-47.5	38.8
t-butyl	79/168	-41.9	60.5	1/168	-47.4	52.9
azide	73/188	-47.2	53.1	1/188	-51.2	45.5
n-butyl	48/179	-44.9	45.5	1/179	-47.8	37.5
i-butyl	2/180	-52.9	47.8	1/180	-48.3	42.8
benzyl	134/156	-32.2	67.9	1/156	-49.3	60.8

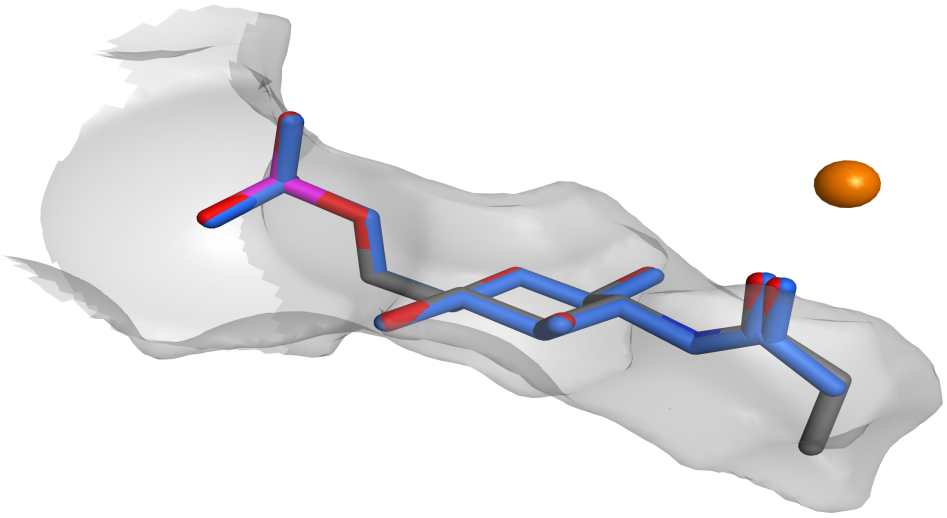
Rank indicates the occurrence of the specified ligand-binding mode in the set of all binding modes ranked by a score value calculated as the sum of Eint and Econf. Eint and Econf, calculated using the MMFF94s force field and a general born implicit solvation model, reported with units kcal/mol.

Docking experiment; reactive binding mode

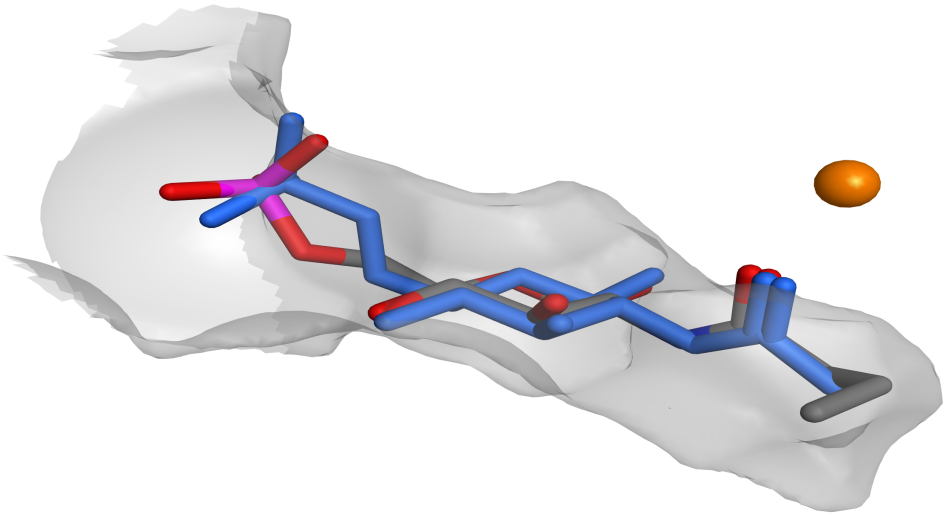
GlcNAc docking



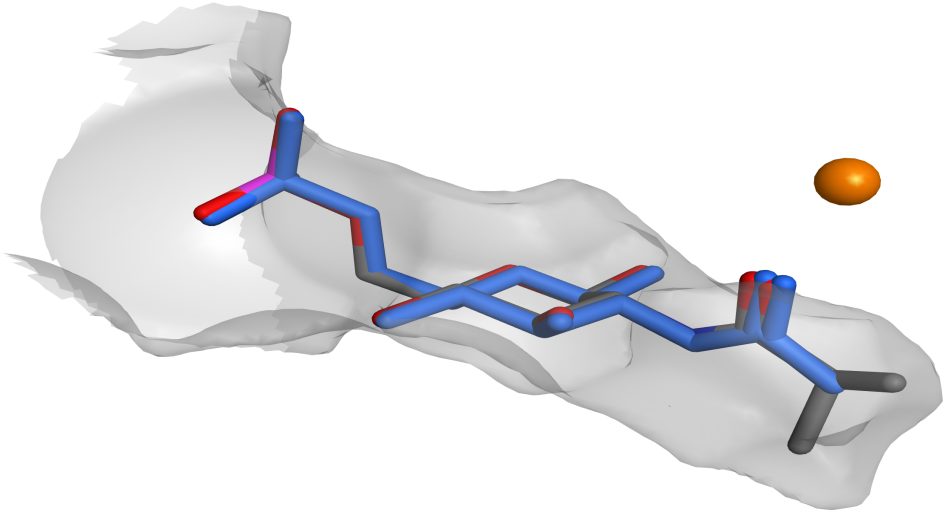
GlcNPr docking



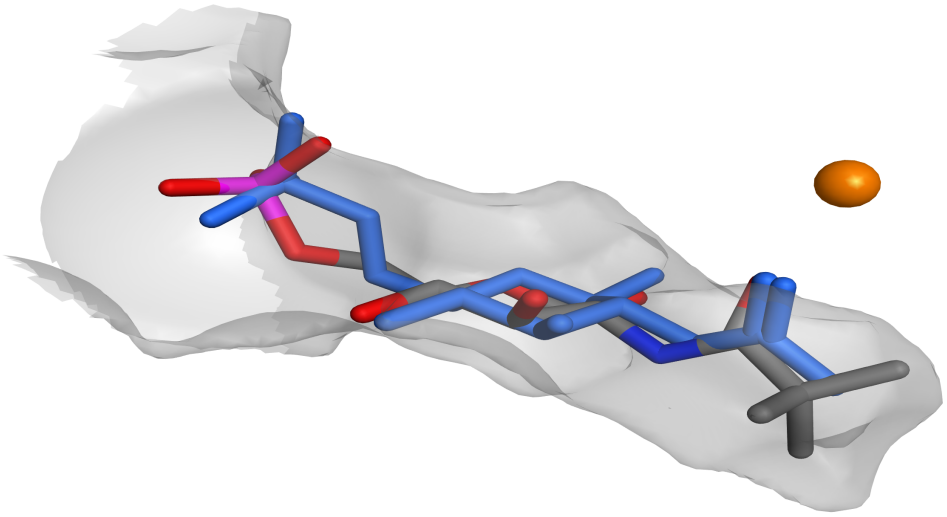
GlcNnBu docking



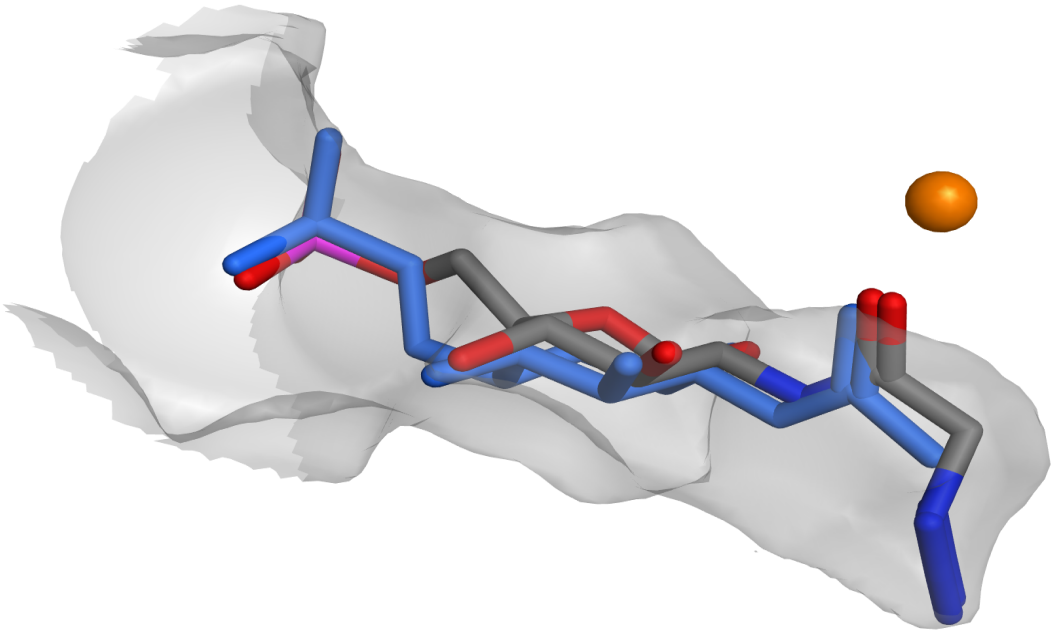
GlcNiBu docking



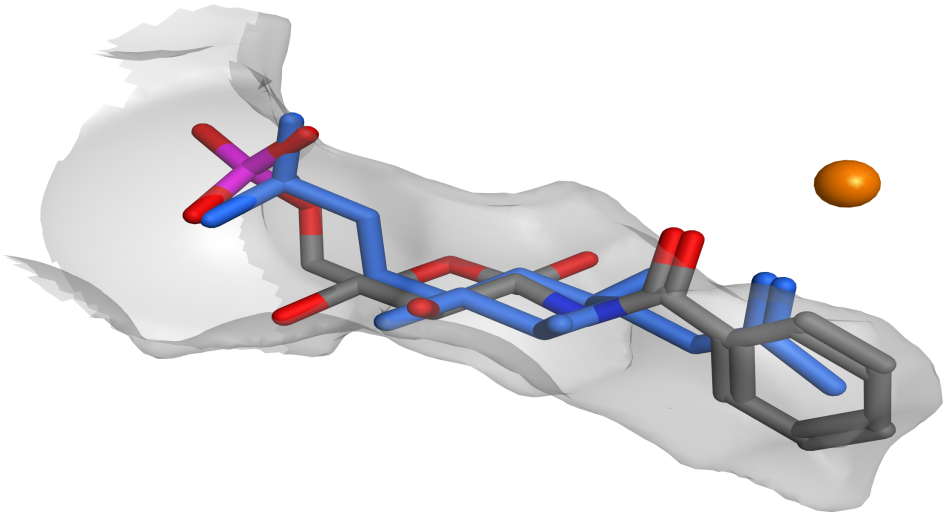
GlcNPiv docking



GlcNAz docking

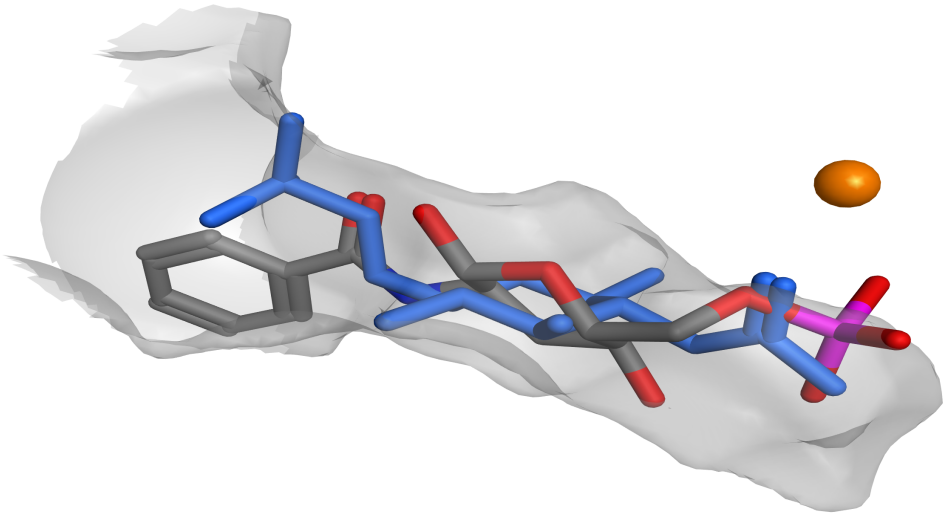


GlcNBz docking



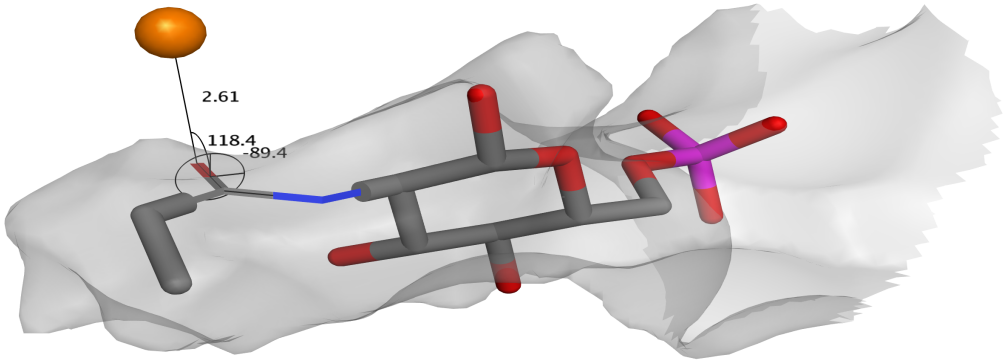
Docking experiment; flipped binding mode

GlcNBz docking

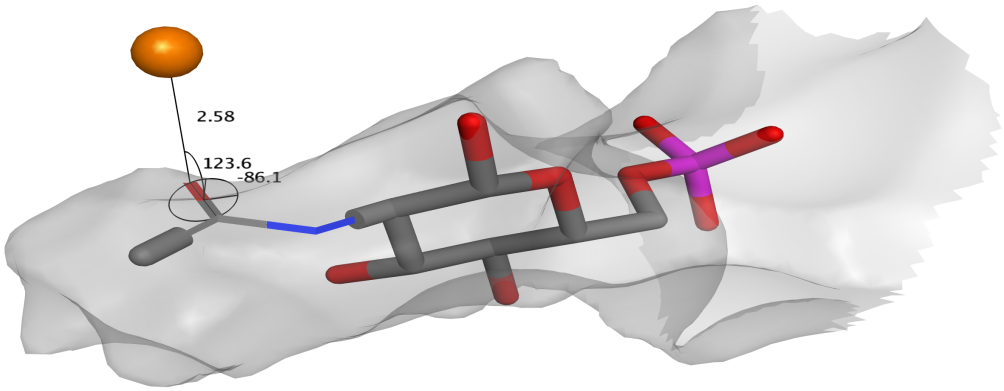


Geometrical analysis of the reactive binding mode

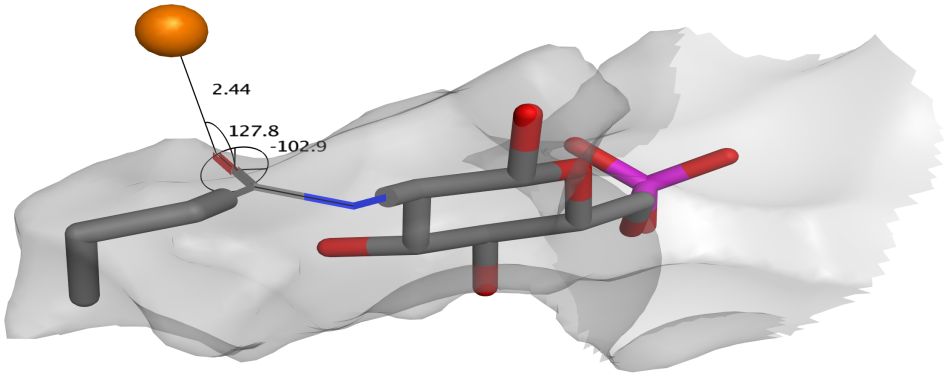
GlcNAc geometry



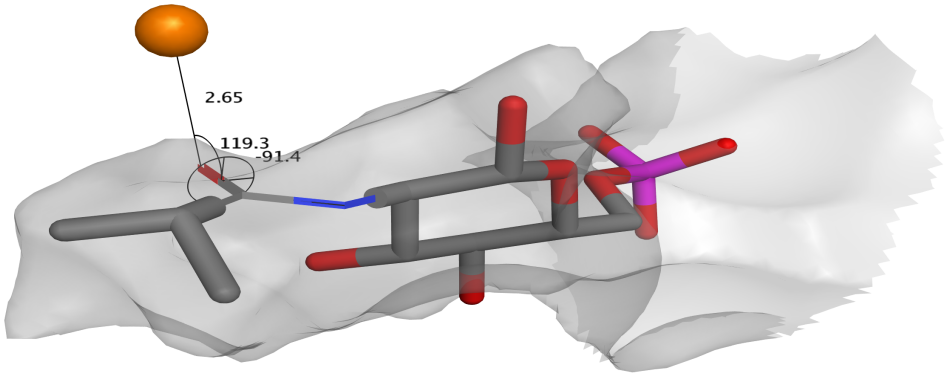
GlcNPr geometry



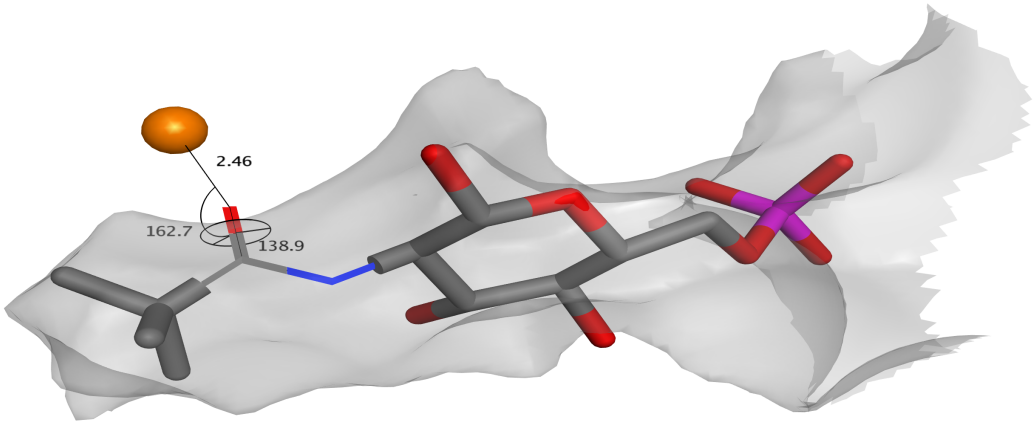
GlcNnBu geometry



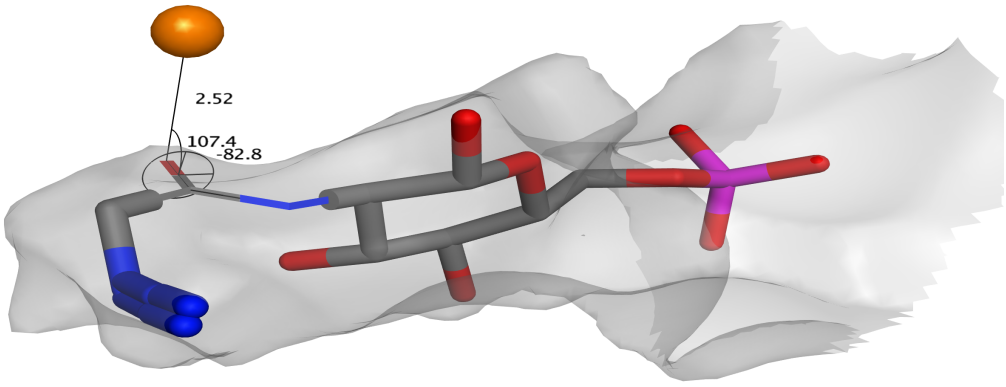
GlcNiBu geometry



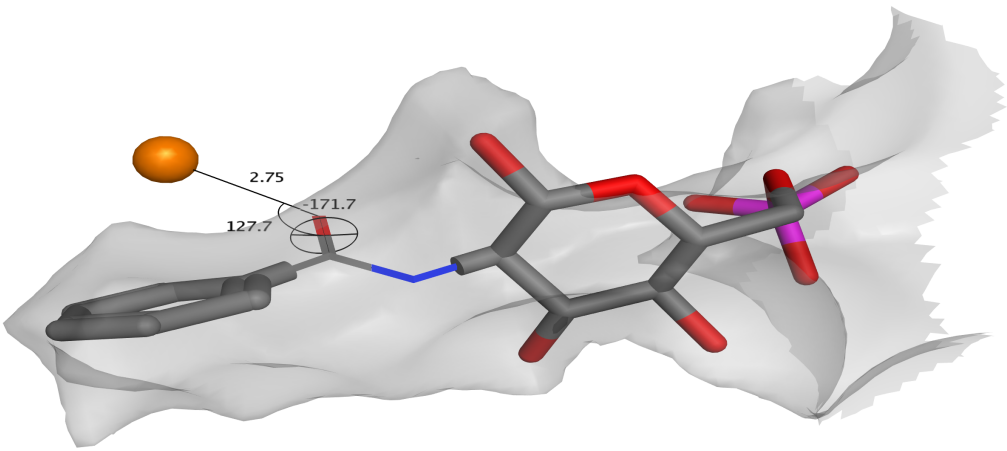
GlcNPiv geometry



GlcNAz geometry

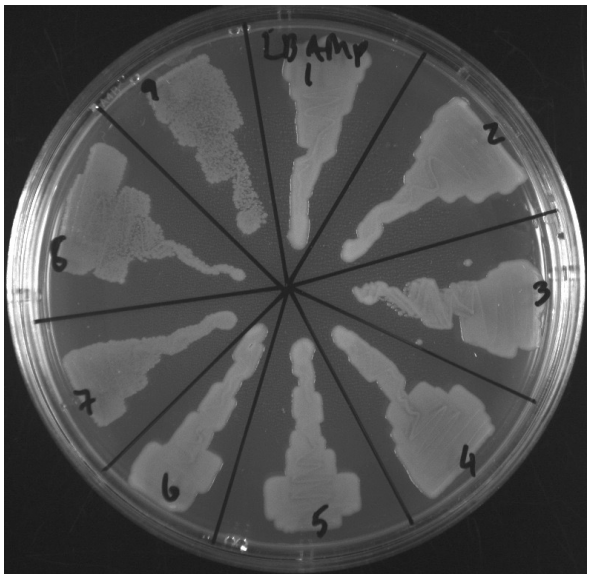


GlcNBz geometry

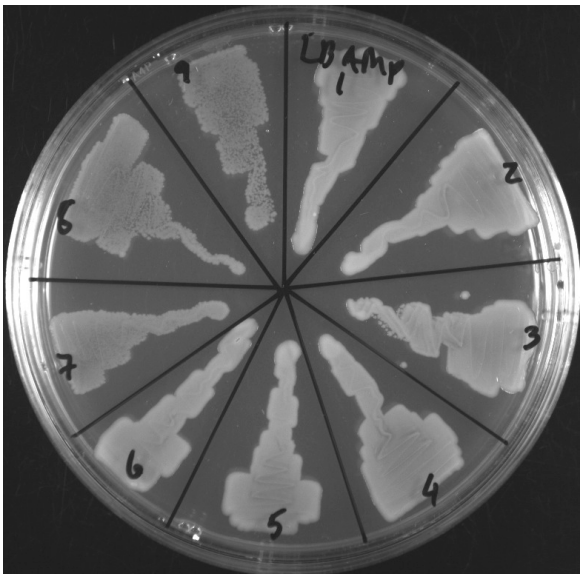


Results of the feeding experiment using LB media. Left, 24 h; right, 48 h; bottom, 72 h of incubation.

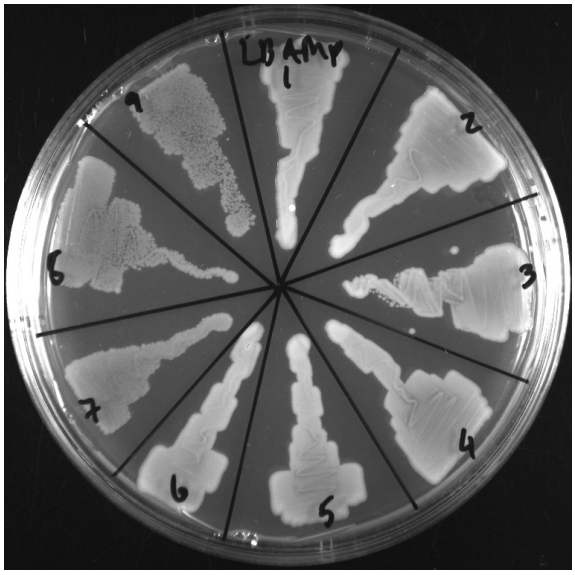
24



48

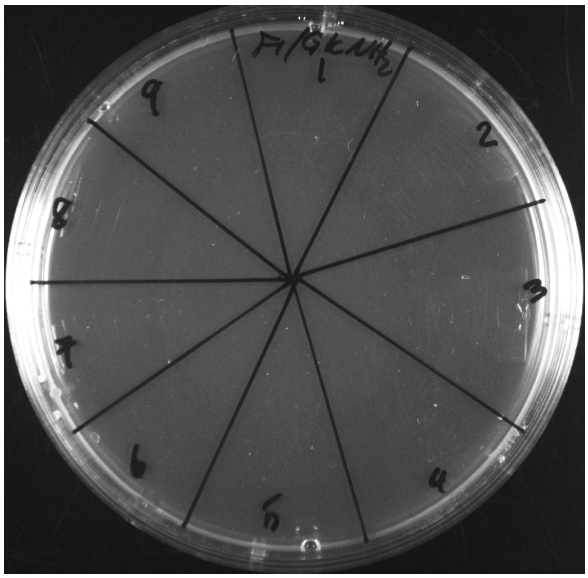


72

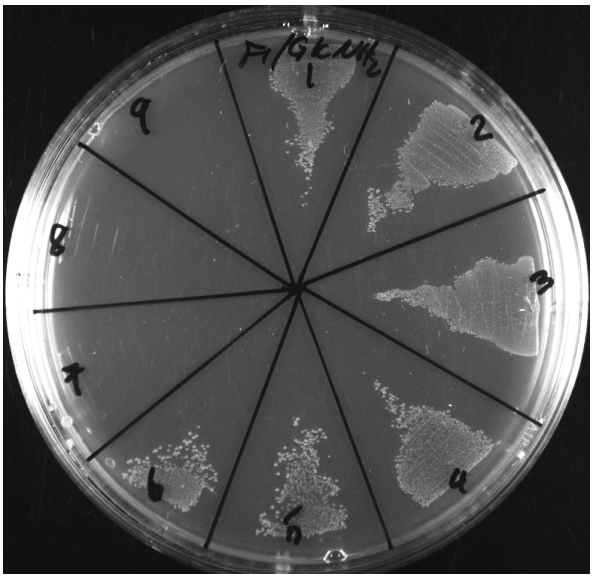


Results of the feeding experiment using $GlcNH_2$ as carbon source. Left, 24 h; right, 48 h; bottom, 72 h of incubation.

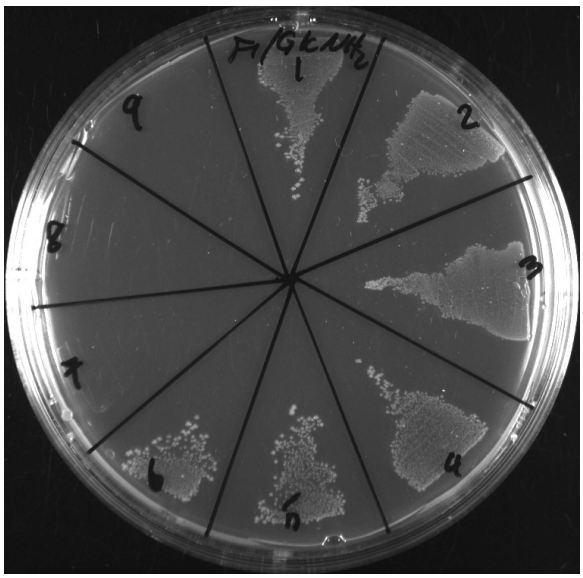
24



48

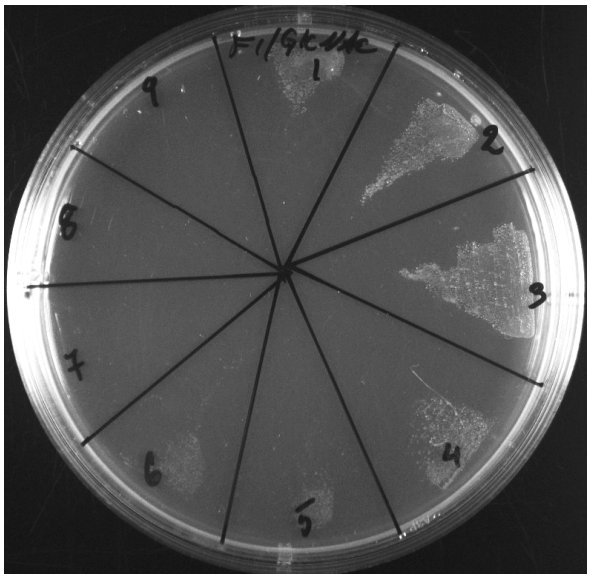


72

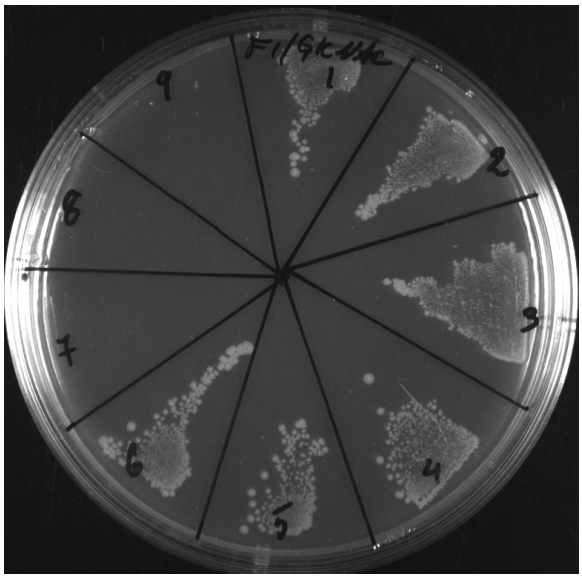


Results of the feeding experiment using GlcNAc as carbon source. Left, 24 h; right, 48 h; bottom, 72 h of incubation.

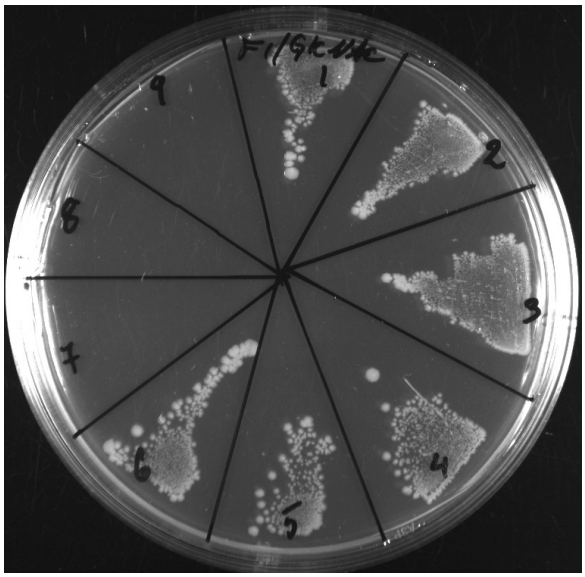
24



48

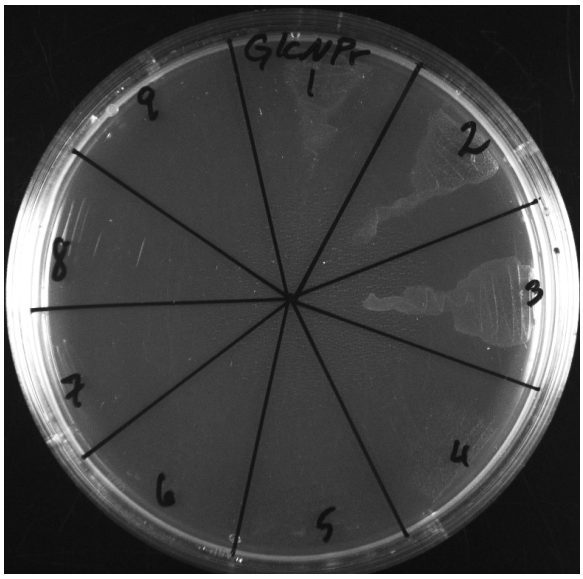


72

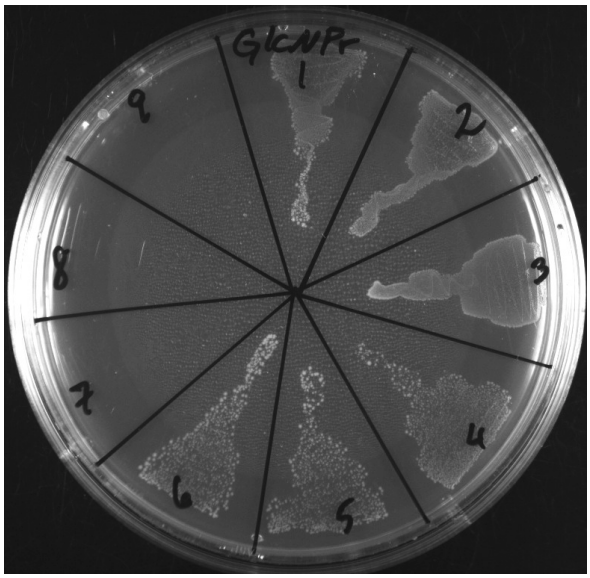


Results of the feeding experiment using GlcNPr as carbon source. Left, 24 h; right, 48 h; bottom, 72 h of incubation.

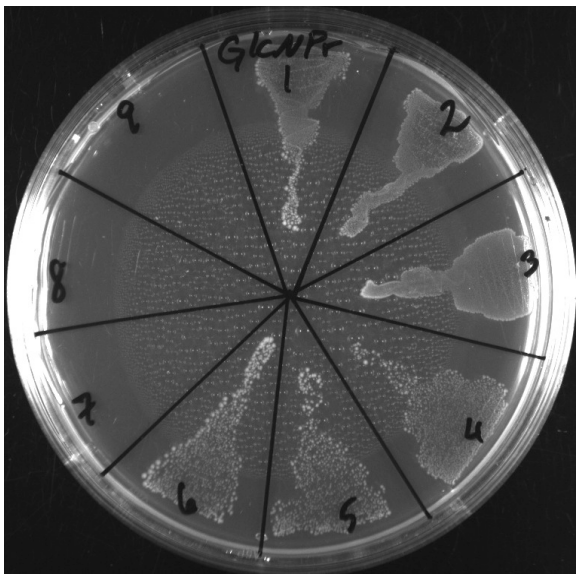
24



48

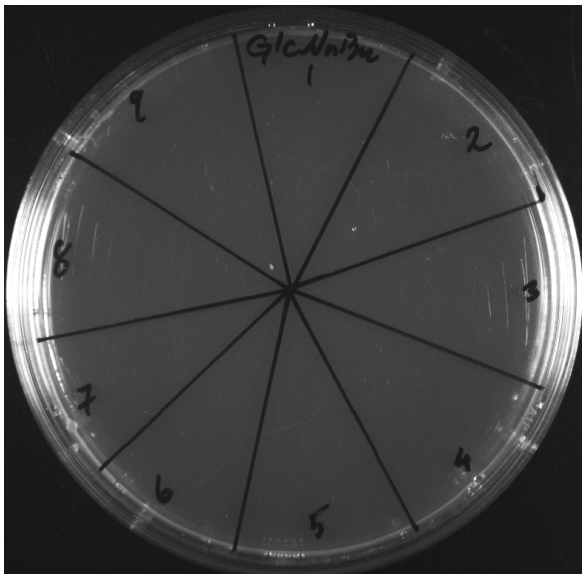


72

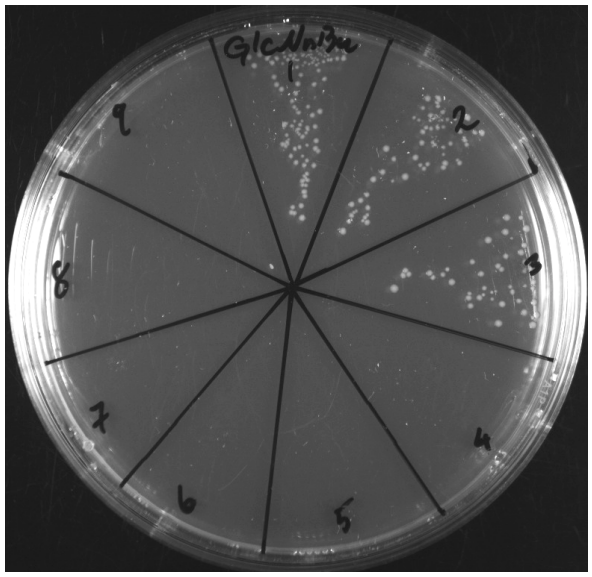


Results of the feeding experiment using GlcNBu as carbon source. Left, 24 h; right, 48 h; bottom, 72 h of incubation.

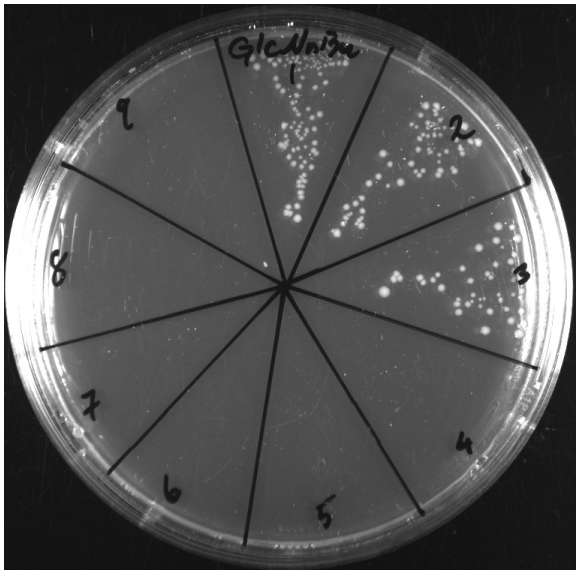
24



48

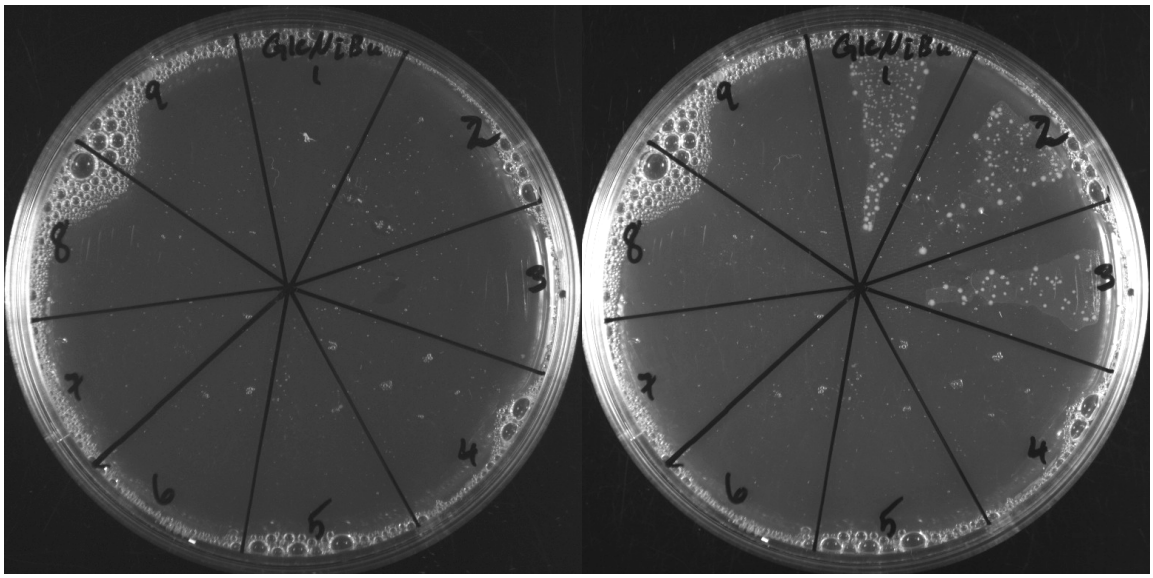


72

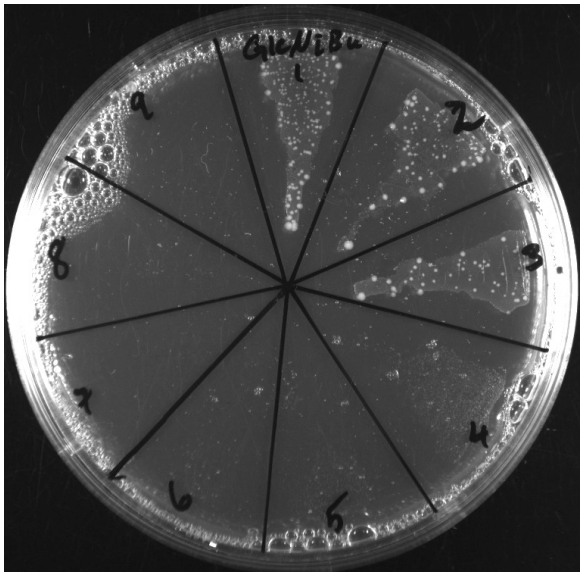


Results of the feeding experiment using GlcNiBu as carbon source. Left, 24 h; right, 48 h; bottom, 72 h of incubation.

24 48

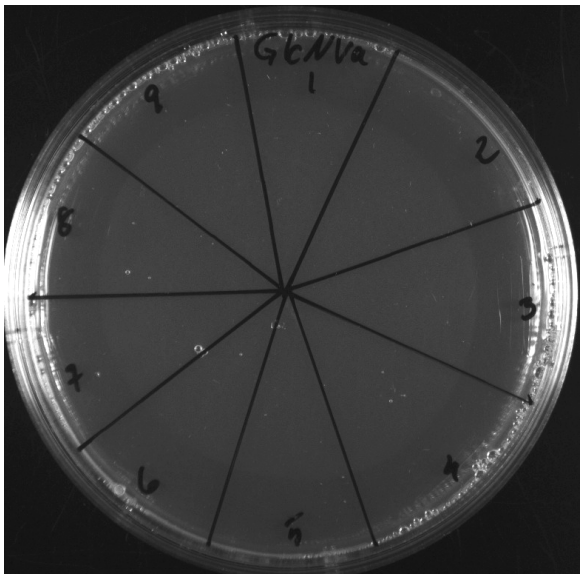


72

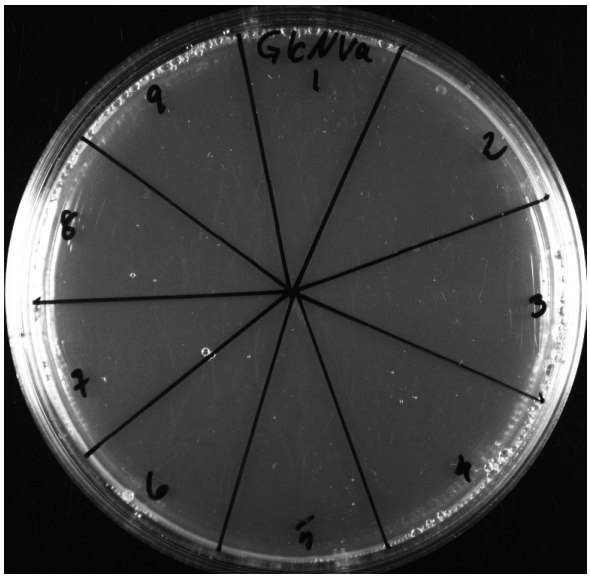


Results of the feeding experiment using GlcNVa as carbon source. Left, 24 h; right, 48 h; bottom, 72 h of incubation.

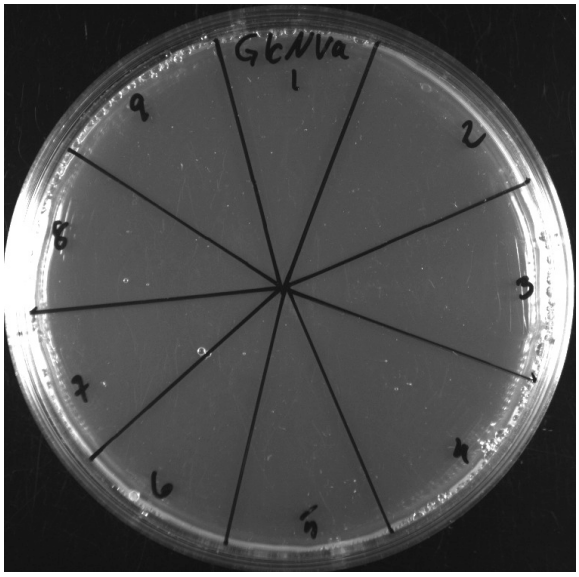
24



48

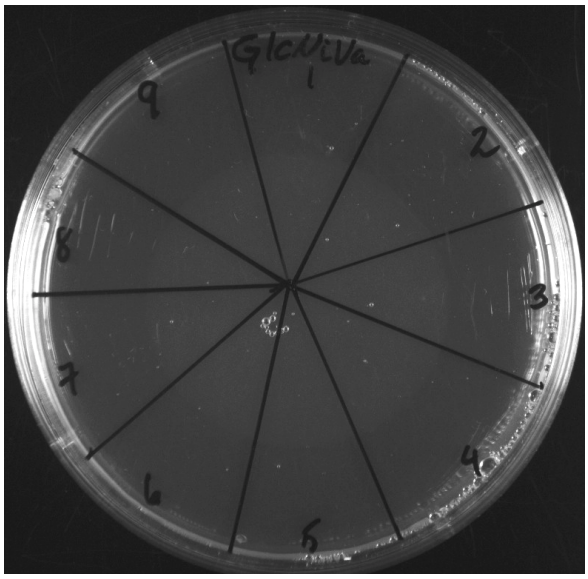


72



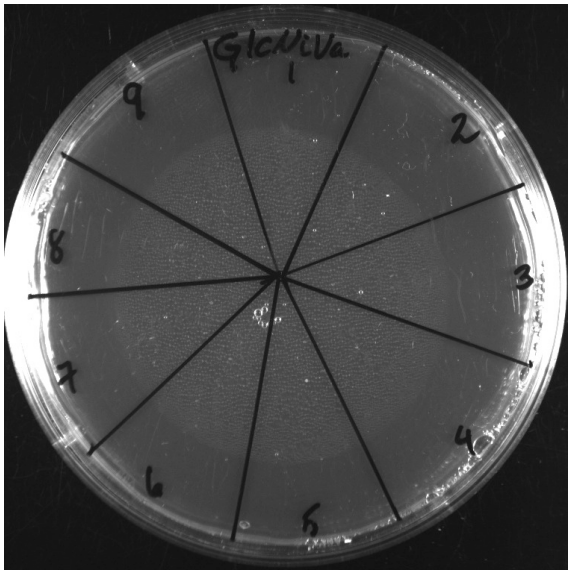
Results of the feeding experiment using GlcNiVa as carbon source. Left, 24 h; right, 48 h; bottom, 72 h of incubation.

24



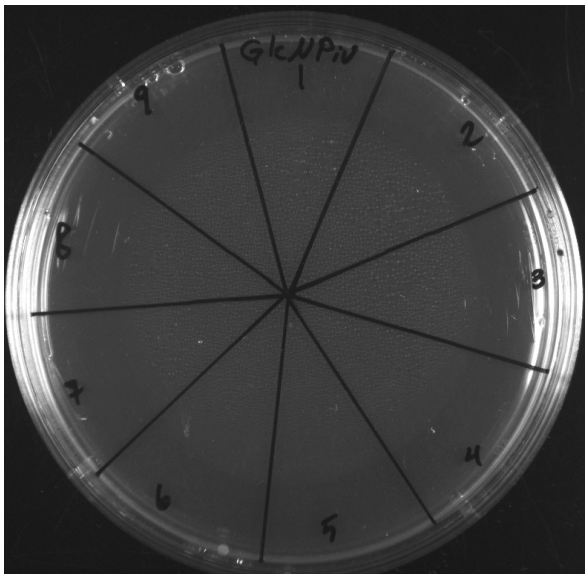
48

72

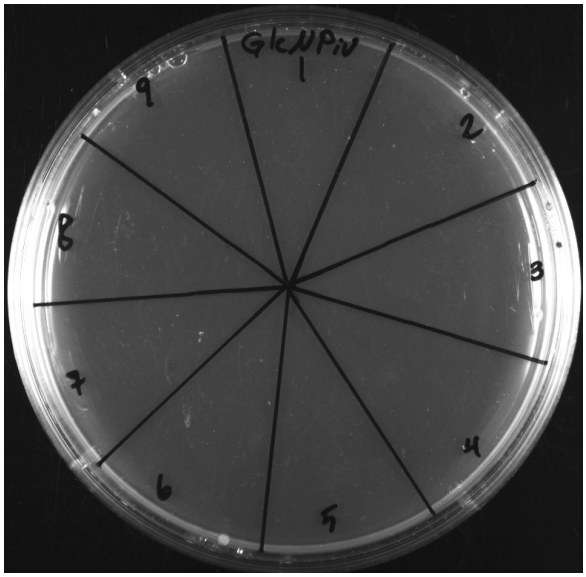


Results of the feeding experiment using GlcNPiv as carbon source. Left, 24 h; right, 48 h; bottom, 72 h of incubation.

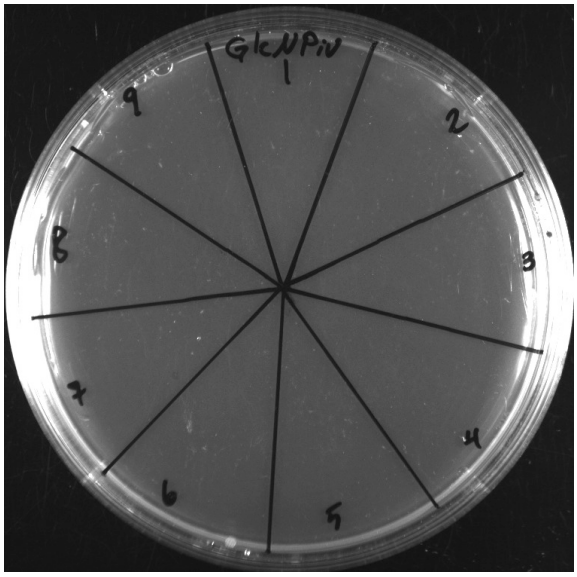
24



48

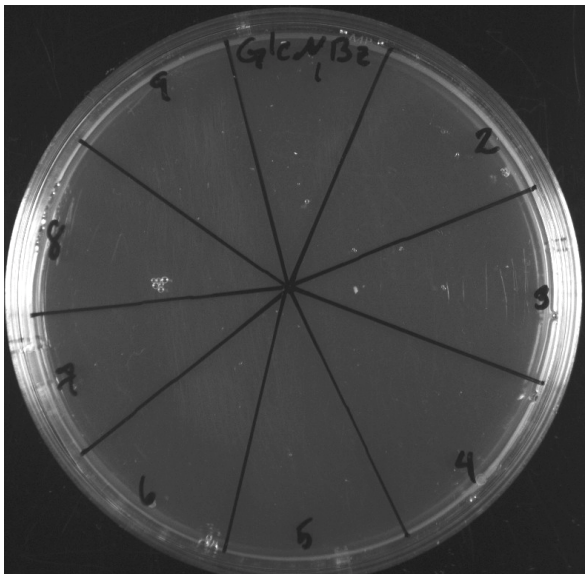


72

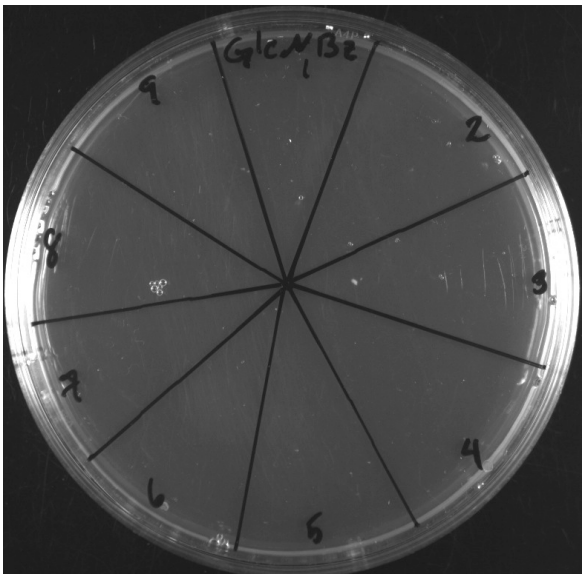


Results of the feeding experiment using GlcNBz as carbon source. Left, 24 h; right, 48 h; bottom, 72 h of incubation.

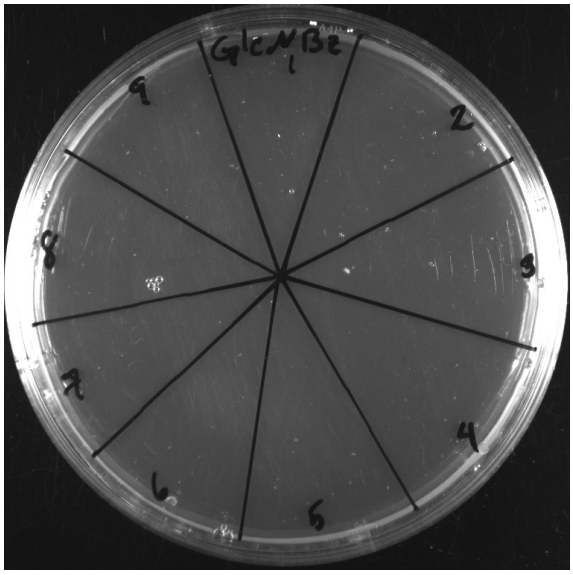
24



48

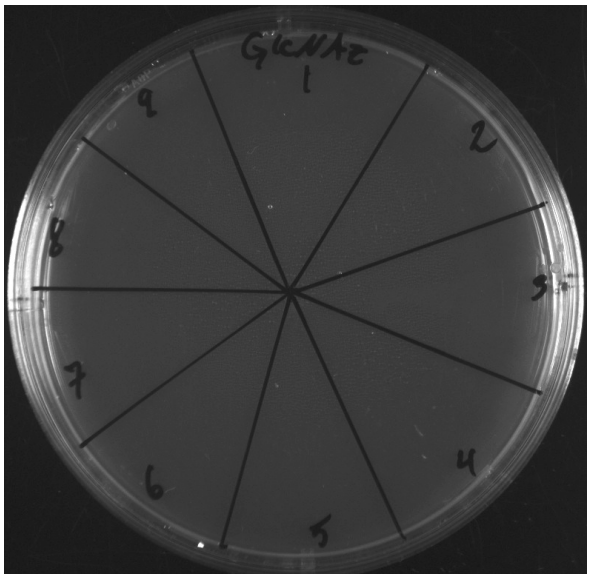


72

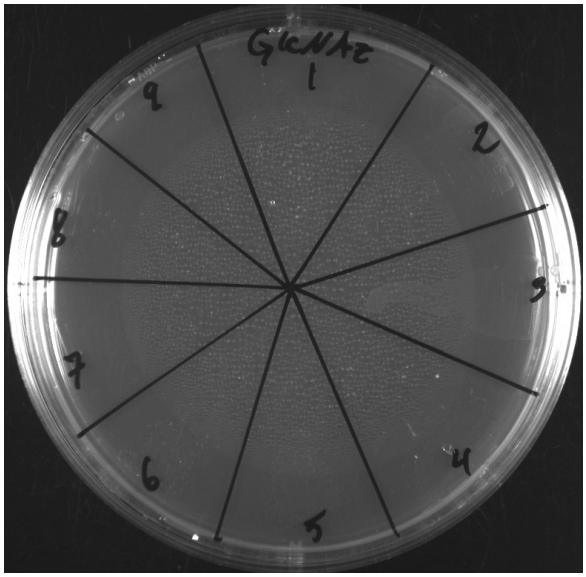


Results of the feeding experiment using GlcNAz as carbon source. Left, 24 h; right, 48 h; bottom, 72 h of incubation.

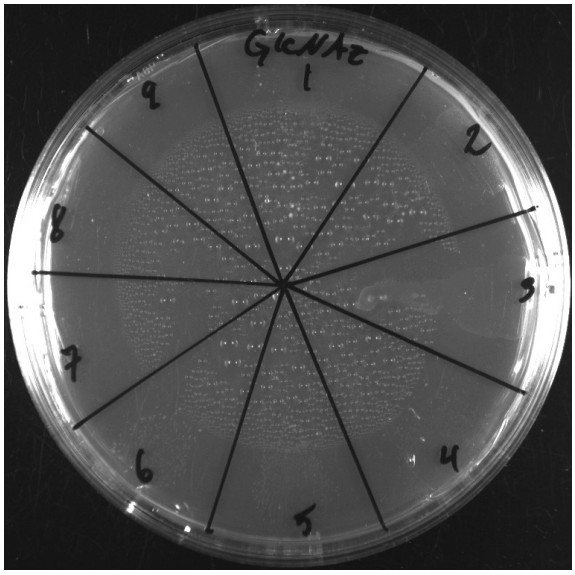
24



48



72



Chapter 3. Enzymatic and chemical phosphorylation of N-acylglucosamines to generate N-acylglucosamine-6-phosphates

3.1 Introduction

Kinases are important enzymes that catalyze the transfer of a phosphate group from a high-energy phosphate donor, *e.g*. ATP, to a specific substrate (fig 3.1). These substrates include different biological compounds, such as proteins, sugars or lipids.

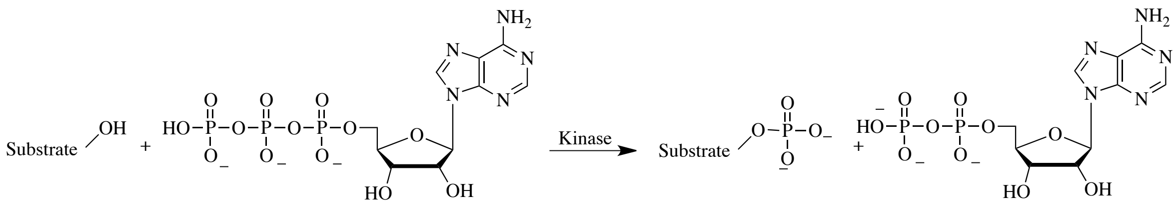


Figure 3. 1. General phosphorylation reaction catalyzed by kinases

Kinases are used to transmit signals and regulate complex biochemical processes in cells. The phosphorylation of substrates can activate, inactivate or modulate interactions with other molecules. Thus, mutations in kinases can produce different diseases in humans, including certain types of leukemia, neuroblastomas[1,2], glioblastoma[1,2], spinocerebellar ataxia[1], forms of agammaglobulinemia[1] among others[1].

The phosphorylation of sugars is a key step in cellular metabolism. During transport into the cell, monosaccharides are phosphorylated to trap them within the cell for further metabolic processing[3,4]. Transferring the γ-phosphate of ATP to different sugar substrates is catalyzed by three different non-homologous protein families (fig. 3.2): hexokinases, ribokinases and galactokinases[5].

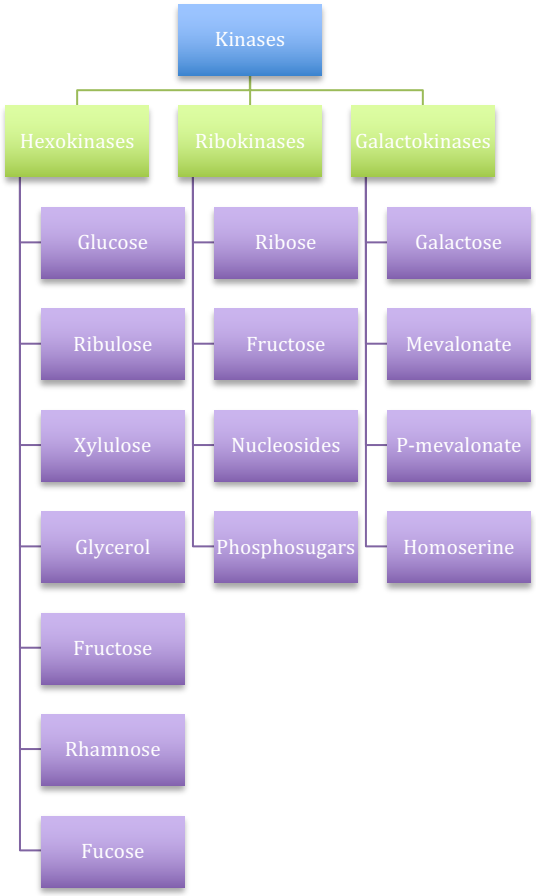


Figure 3. 2. Classification of the kinases super families and their substrates.

The hexokinase family contains glucokinases, ribulokinases, gluconokinases, xylulokinases, glycerokinases, fructokinases, rhamnokinases and fucokinases[5]. On the other hand, the galactokinase family contains enzymes capable of phosphorylating galactose, mevalonate, phosphomevalonate and homoserine[5]. The ribokinase family enzymes are able to transfer the ATP phosphate to sugars like ribose, fructose, sugar-containing molecules such as nucleosides, and sugar phosphate molecules like fructose-6-phosphate, fructose-1-phosphate and tagatose-6-phosphate[5]. In summary, members of the hexokinase family are capable of phosphorylating simple sugar substrates. Meanwhile, the

galactokinase family has a broader scope of substrates, while the ribokinase family has the broadest, being able to phosphorylate sugars, phosphosugars and nucleosides.

N-acetylglucosamine kinase belongs to the glucokinase group within the hexokinase family[6]. It catalyzes the phosphorylation of N-acetylglucosamine (GlcNAc) to *N*-acetylglucosamine-6-phosphate (GlcNAc-6-P); fig 3.3.

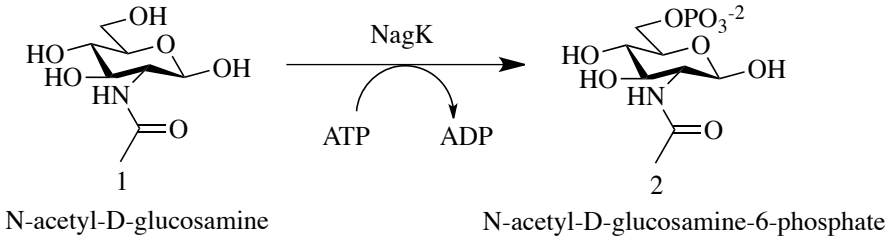


Figure 3. 3. GlcNAc phosphorylation reaction.

GlcNAc, from lysosomal degradation of oligosaccharides or nutritional sources, is an important substrate for the synthesis of UDP-GlcNAc, which is required in the biosynthesis of N- and O- glycans[7] and peptidoglycan[8].

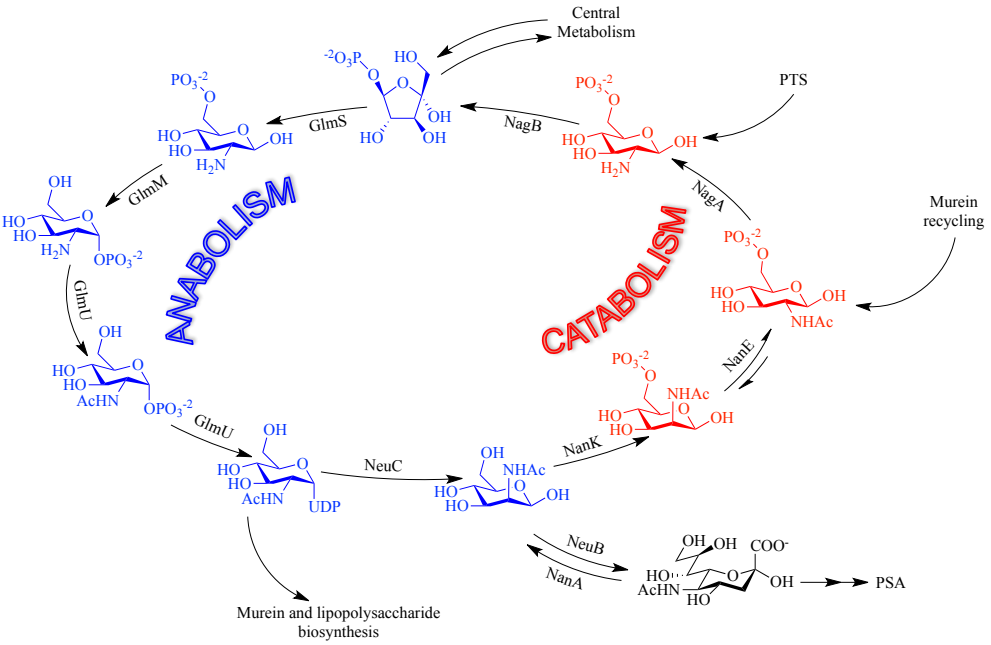


Figure 3. 4. Sialic acid metabolism.

N-acetyl-D-glucosamine kinase has an important role in *E. coli*. This enzyme is involved in the recycling of peptidoglycan from the bacterial cell wall[9] (fig. 3.5).

When *E. coli* grows in a medium rich in GlcNAc, the sugar is transported into the cell by a group of proteins and enzymes that form part of the phosphoenolpyruvate (PEP)-dependent phosphotransferase system superfamily. This transporter phosphorylates GlcNAc giving GlcNAc-6-P. GlcNAc-6-P is deacetyladed by NagA and the product, $GlcNH_2$-6-P, can enter glycolysis after been deaminated by NagB or go to biosynthesis pathways for murein and lipopolysaccharide by conversion to UDP-GlcNAc by GlmM and GlmU (fig. 3.4).

However, significant amounts of non-phosphorylated GlcNAc are released in the cytoplasm via the murein recycling pathway[10,11] (fig. 3.5). This recycling process involves the phosphorylation of GlcNAc by NagK.

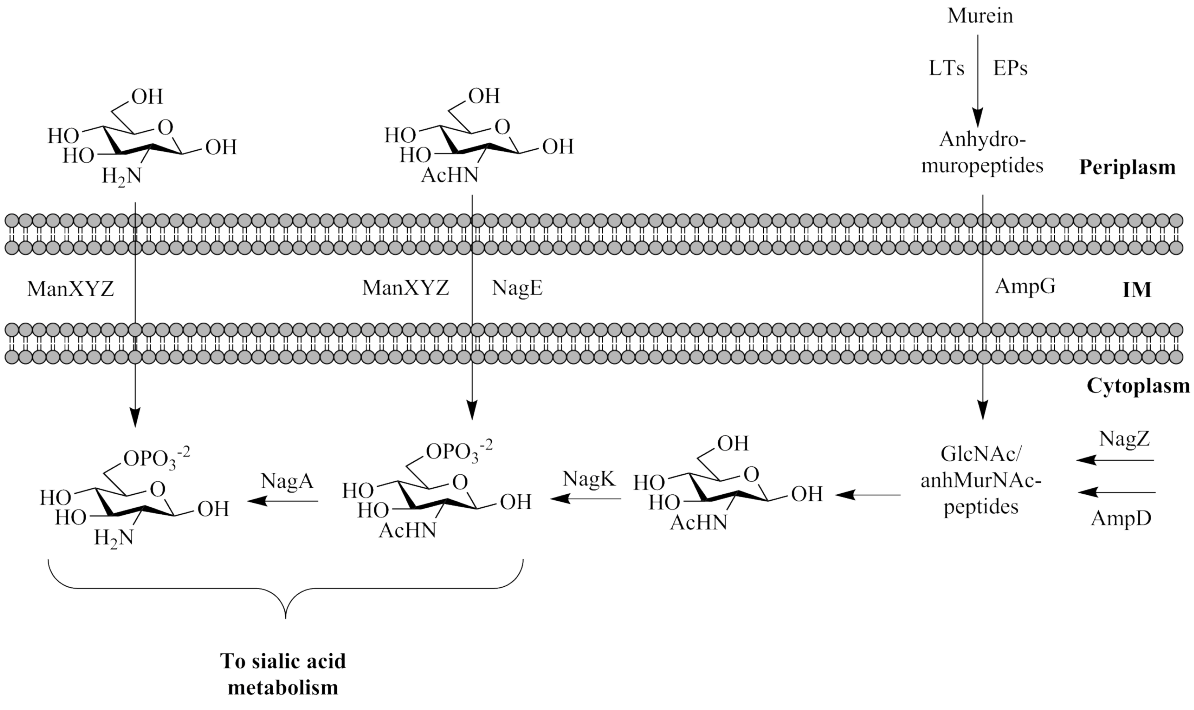


Figure 3. 5. Murein recycling and metabolism of GlcNAc and their connection with sialic acid metabolism. Adapted from Uehara, T. and Park, T.[9]

Hexokinase specificity has been studied intensively since 1960's. Researchers have shown that a change in carbon 1 or 2 has little or no effect in kinase activity. However, replacement of the hydroxyl group in carbon 3 or 4 decreased hexokinase activity[12]. Moreover, changes in carbons 2 and 4 decreased activity further (fig. 3.6). Hexokinases were completely inactive when configuration on carbons 2 and 3 or 3 and 4 differ from glucose[12]. More recent work examining the evolution of functional specificities of FGGY carbohydrate kinase super family members has shown similar trends[13]. Here, all the assays were done with several sugars that have hydroxyl groups in positions 2, 3 or 4. This is confirmed by Larion *et al*. where they show improvements in the kinase activity of kinases from ROK super family by mutagenesis[14]. This research group studied allokinase (AlsK) and *N*-acetyl-D-mannosamine kinase (NanK) from *E. coli*. By random mutagenesis they changed one amino acid from the enzymes and analyzed the enzymatic activity of the mutated enzymes with allose, altrose, 2-deoxyglucose, mannose and *N*-acetyl-D-mannosamine. Enzyme activity was measured by coupling the ADP production to the oxidation of NADH via the combined action of pyruvate kinase and lactate dehydrogenase. The results were fitted to Michaelis-Menten model and K_m and k_{cat} were determined and compared[14].

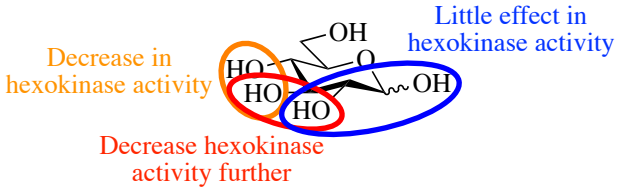


Figure 3. 6. Relationship between hydroxyl groups in glucose and the effect on the enzymatic activity of hexokinase.

In general terms, kinases can phosphorylate several sugars like compounds; however, certain configurations and substituents are not allowed. For this reason, a phosphorylation strategy becomes relevant in order to phosphorylate sugars that differ from their natural homologs.

Although, specific kinases for substituted sugars exist, e.g. *N*-acetyl-D-glucosamine kinase from *E. coli*[9,15–17], or glucosamine specific kinase from *V. cholerae*[18], a general kinase capable of phosphorylating sugars with acyl groups in position 2, 3 or 4 does not exist. For this reason, it is imperative to develop a synthetic strategy that allows the incorporation of a phosphate group in important positions such as carbon 6.

Classic sugar chemical modification involves the peracetylation[19,20] or benzylation[21] of the sugar to protect all the hydroxyl groups. After this first step, selective deprotection and protection of the groups is performed until the desired product is obtained. However, this classic approach involves numerous synthesis reactions and several purification steps.

Our goal was thus to develop a synthetic strategy that allowed researchers who were non-experts in carbohydrate synthesis to install a phosphate group in the C6 position in glucosamine and *N*-acylglucosamine, and more generally, in any primary hydroxyl group of a sugar. This will enable biochemists to better probe the substrate selectivities of carbohydrate processing enzymes with non-native substrates.

3.2 Results and discussion

3.2.1 Determination of kinase substrate specificity (ATP assay)

Phosphorylation of GlcNAc is an important step for the next catabolic reaction, deacetylation. It is possible that if the sugar is not phosphorylated, the deacetylation reaction may not proceed.

Kinases use ATP as a phosphate donor. This transfer produces ADP as a byproduct decreasing the concentration of ATP in solution. One method to measure the amount of ATP in vitro is the firefly luciferase assay. Luciferase is an enzyme that uses luciferin, ATP and oxygen to produce oxyluciferin, pyrophosphate, AMP and CO_2. During this process, luciferase is able to produce photons as part of the reaction. This metabolic process is used in certain organisms to produce bioluminescence (fig. 3.7) and the amount of light produced shows a linear correlation with the amount of ATP present in solution.

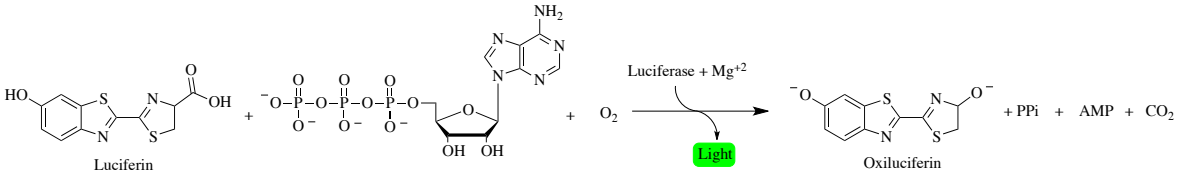


Figure 3. 7. Luciferin reaction with ATP in the presence of Luciferase.

NagK and NanK were overexpressed in *E. coli* BL21 to analyze their substrate tolerance. The first experiment compares the enzymatic activity of the two enzymes against three different sugars and a negative control. The phosphorylation activity was measured indirectly, quantifying the remaining ATP after ten minutes of reaction. Figure 3.8 shows the remaining ATP concentration in the enzymatic reactions with the three different sugar

substrates. The higher the concentration of ATP, the lower the level of phosphorylation of the sugar by the enzyme.

NagK shows a low concentration of ATP with GlcNAc, meanwhile the rest of the sugars have an ATP concentration similar to the negative control. The same behavior is observed with NanK and ManNAc.

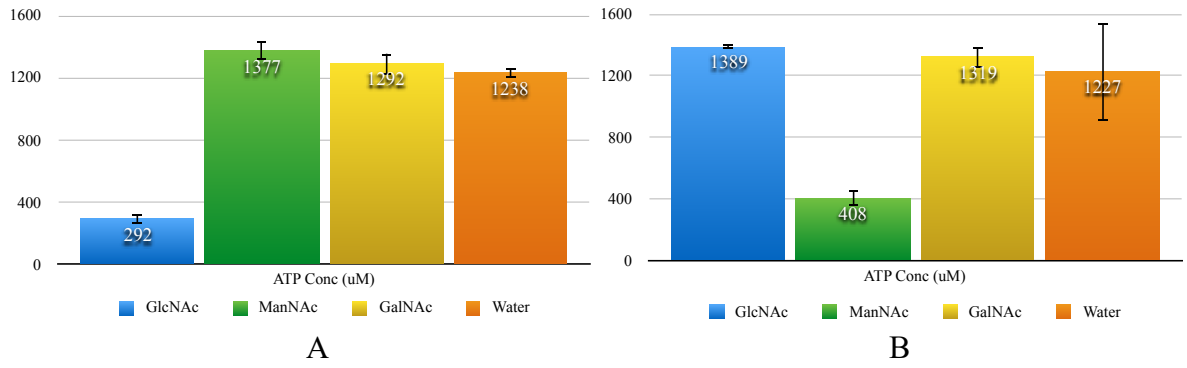


Figure 3. 8. Enzymatic activity of N-acetyl-D-glucosamine kinase –NagK (A)– and N-acetyl-D-mannosamine kinase –NanK (B)–.

This evidence shows that these kinases are substrate specific. ATP concentration decreases approximately three-fold for the native substrate. The ATP levels for the other sugars are consistent with the negative control, indicating no reaction has occurred.

After realizing the specificity of these kinases, it was important to corroborate that specificity on the glucosamine analogs we intended to study for sialic acid analog production.

Figure 3.9 shows the ATP consumption for all the *N*-acylglucosamine substrates used in chapter 2 as well as a few other aminosugars. Even though propyl, butyl and valeric glucosamines show certain degree of phosphorylation, the comparison with GlcNAc

proves the low efficiency of NagK to phosphorylate non-native substrates. Once again, the specificity of the NagK is evident.

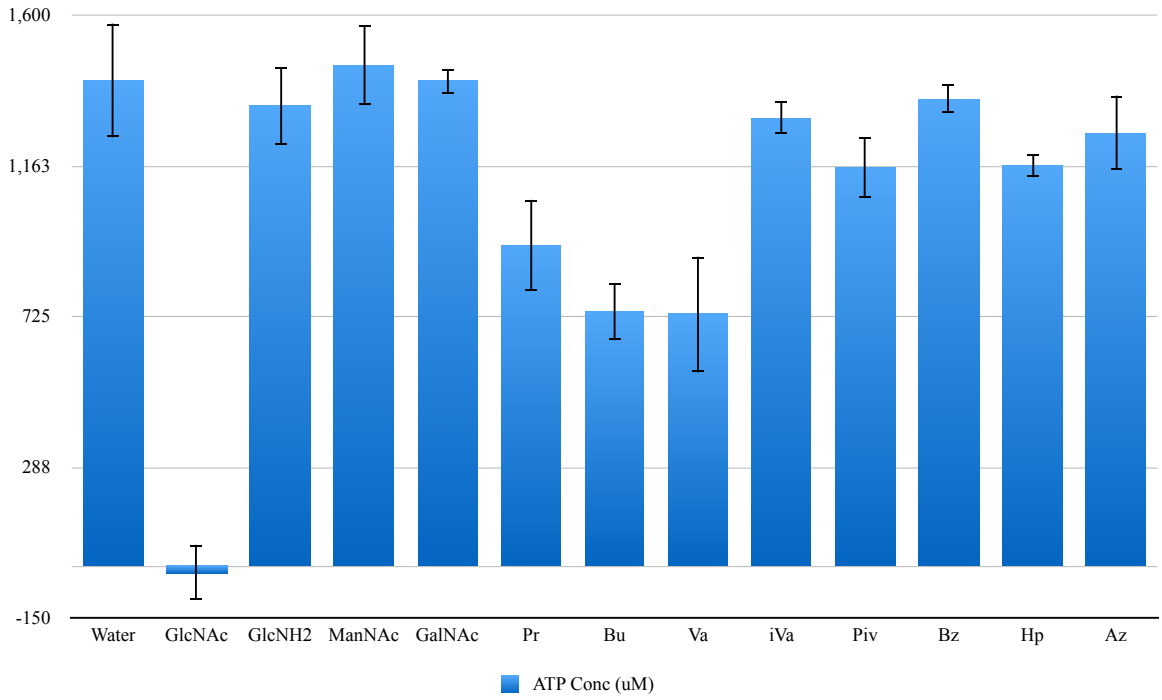


Figure 3. 9. Comparison of remaining ATP concentration in enzymatic phosphorylation reactions using NagK as catalyst.

3.2.2 Synthesis of N-acylglucosamine-6-phosphate

Enzymes are important and reliable biological catalysts and can be used in organic synthesis to achieve reactions in an easier and greener way[22]. However, some enzymatic systems are too complex, the reaction yields are too low, or the enzymatic specificity is too high for effective organic synthesis[22]. Moreover, although many enzymes have been thoroughly studied with regard to a substrate specificity, it is not certain or predictable how the same enzyme will interact with non-natural substrates[23]. In these cases, researchers

have to come with ideas to obtain the product of an enzymatic reaction with a substrate that is not compatible with the enzyme.

Evidence showed that kinases are substrate specific (fig. 3.7 and 3.8), for this reason, it was important to develop a synthetic route able to phosphorylate primary alcohols in sugars.

Sugar synthesis involves a numerous protection and deprotection reactions of selected functional groups[24]. There are several approaches to synthesis and modification of sugars, but these multistep strategies may cause environmental problems for large-scale production due to the use of toxic reagents and harmful organic solvents[25]. Other problems, such as low yield overall reactions or achieving stereoselective couplings for multifunctional carbohydrate donors and acceptors. Many of these problems can be traced back to the extensive protection of the functional groups required for typical carbohydrate chemistry[24].

One of the first protection methods developed was the peracetylation of sugars[26]. Other strategies have been developed, using orthogonal protecting groups that can be selectively removed in the presence of other protecting groups. Nevertheless, this synthetic strategy approach is long, and the final yields usually are low[24,27].

Complete protection of the OH groups is done to avoid secondary reactions were the alcohol can act as a nucleophile. However, these hydroxyl groups have different reactivity, according to the position and configuration. Sugars have broadly speaking three different kinds of hydroxyl groups. The primary alcohol is more reactive than the secondary because it is less hindered. The anomeric alcohol, on the other hand, has a very

special reactivity since it is a hemiacetal. It can undergo a number of unique reactions due to the presence of the second heteroatom attached to the anomeric carbon[28].

Since the hydroxyl group on carbon six is a primary alcohol, our first attempt towards an efficient synthesis was the direct phosphorylation of C6. This approach would give a P-protected phosphosugar that could be deprotected in the subsequent step. To attach the phosphate, several phosphoryl compounds were used. Due to the decreased steric hindrance of the C6 hydroxyl group, the reaction was expected to be straightforward. Many phosphoryl compounds and conditions were tested, and Table 3.1 shows attempts of direct phosphorylation. None of the reactions tested gave a desirable product.

Table 3. 1. Direct phosphorylation of GlcNAc with different phosphorylation agents and conditions.

Phosphorylating agent	Conditions	Result
Diphenyl phosphoryl chloride	DMF/K_2CO_3/RT	No desired product
Bis(2,2,2-trichloroethyl) phosphorochloridate	Dry pyridine/-35°C	No desired product
Bis(2,2,2-trichloroethyl) phosphorochloridate	THF/1-methylimidazole/0°C	No desired product
Dibenzyl phosphoryl chloride	Pyridine/RT	No desired product
Diethyl chlorophosphite	DMF/0°C	No desired product

Our next approach was to protect the anomeric alcohol due to its unique reactivity[28]. In this case, *N*-acetylglucosamine was mixed with benzyl alcohol[29–31] and acetyl chloride and heated overnight at 60 °C. Benzyl alcohol was chosen to protect the anomeric carbon for two reasons. First, it decreases the polarity of the sugar what make it easier to purify.

Second, it can be readily deprotected by hydrogenolysis with palladium on carbon catalyst[30,31] (fig 3.10).

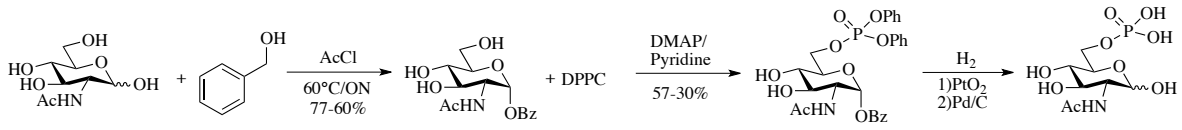


Figure 3. 10. Sugar phosphorylation synthetic strategy using *N*-acetyl-D-glucosamine as a model.

Once the benzylated sugar was obtained, the phosphorylation reaction was attempted. The anomeric protected sugar was dissolved in dry pyridine with DMAP and then, diphenyl phosphoryl chloride was added to the mixture[32–35]. The phosphorylation reaction was successful indicating that the anomeric carbon has to be protected in order to achieve further modifications.

The last reaction was the deprotection of the phosphosugar. In this case, hydrogenolysis with Pd/C to eliminate the benzyl group[21,31,34,36–38] and PtO_2 for the phenyl rings in the phosphate[35,39,40].

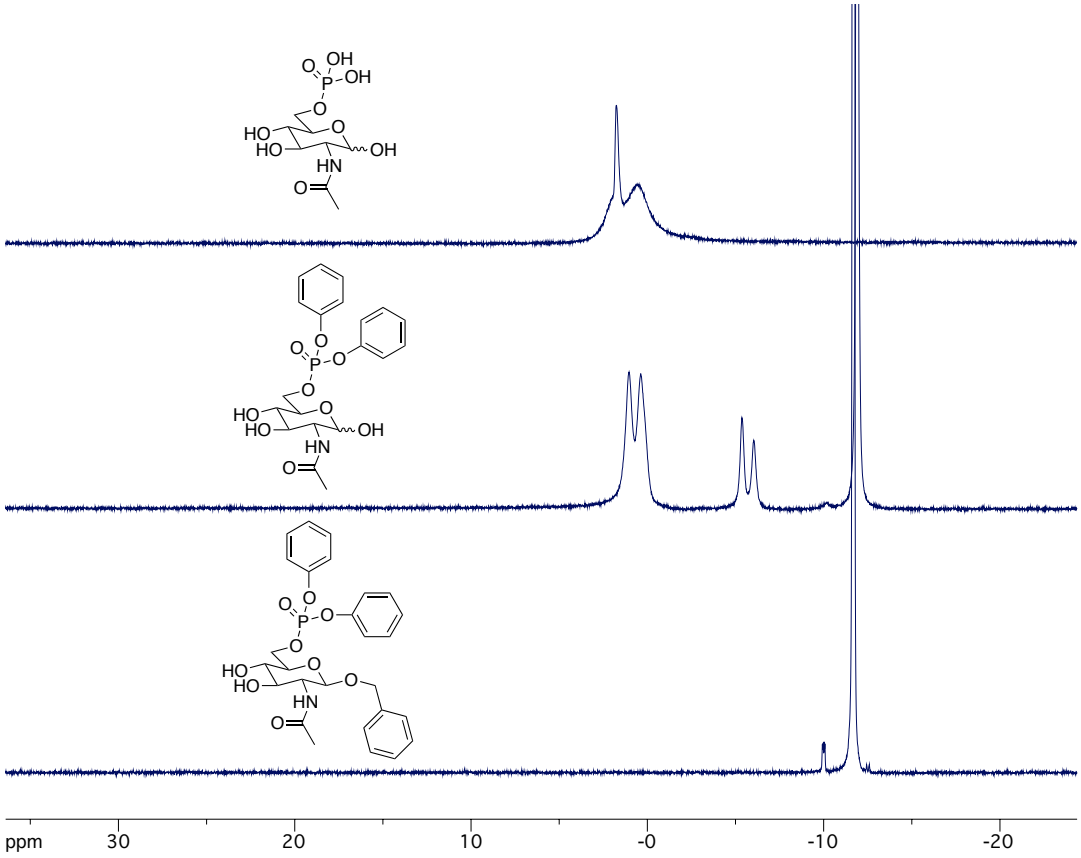


Figure 3. 11. Comparison of different deprotection stages by phosphorous NMR.

TLC and NMR were used to detect changes in the final product of the deprotection reactions. As is demonstrated in figure 3.11, the phosphorous signal in the NMR spectra for the fully protected phosphate has a shift between -11.7 ppm. After the first deprotection reaction with palladium on carbon, the phosphorus signal splits in three, the -10 ppm singlet being the strongest peak. The two doublets at lower field could correspond to partially deprotected sugars or could be due to the change in the environment of the molecule. However, the last NMR shows the fully deprotected GlcNAc, and the ^{31}P signal moved up

to 2 ppm and the signal at -10 ppm (completely protected phosphate) has disappeared. This experiment confirms that the hydrogenolysis of the benzyl and phenyl groups can be achieved with H_2 at low pressure. After the deprotection reaction, we faced a problem. Due to the lack of the aromatic groups, the polarity of the sugar increased substantially; this made the purification process a very complicated task. The reaction mixture was placed on a TLC and developed with a 20% methanol solution in chloroform, giving an Rf of 0. For this reason, obtain a product pure enough to be characterized was impossible.

The objective behind this synthetic strategy was to use different *N*-acylglucosamines to produce their correspondent phosphates. However, several of the acyl groups that we want to install in the glucosamine scaffold present multiple bonds, and hydrogenolysis reaction would reduce those multiple bonds[41–43].

For this reason, we came up with an alternate strategy. In this case, the synthesis begins with glucosamine hydrochloride instead. This approach allows the production of a higher quantity of glucosamine-6-phosphate than can be used later to install the different acyl substituents in the amino group using their correspondent NHS esters.

The first step is the protection of the nitrogen. The amino group was protected with benzyl chloroformate (CBZ)[30,31]. An advantage of this protective group is that it can be deprotected by hydrogenolysis like the others protective groups we used in the previous synthesis, which helps to keep the synthesis relatively short. After the protection of the amino group, we could continue with the synthesis we developed before (fig 3.12).

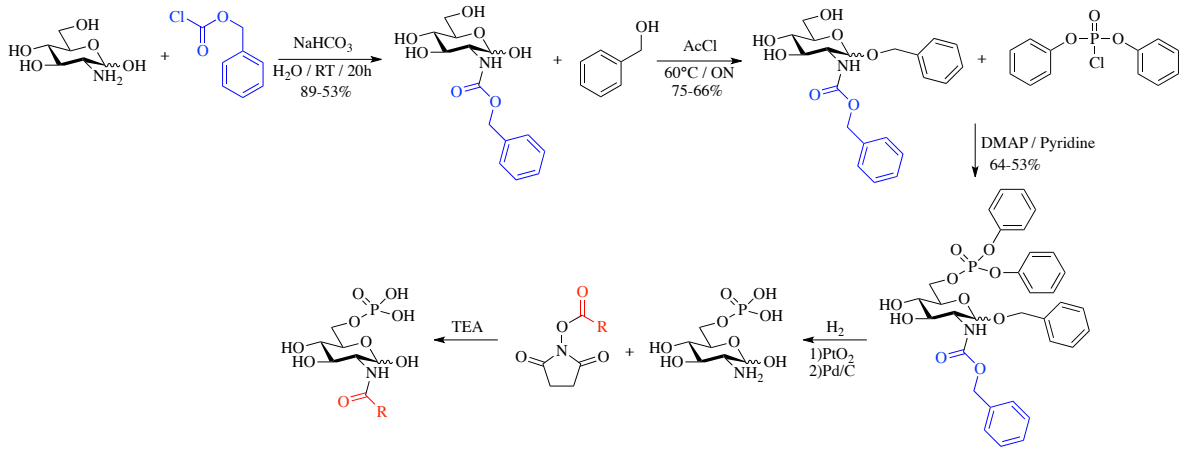


Figure 3. 12. Sugar phosphorylation strategy using glucosamine as starting material.

We were able to isolate a number of 6-phosphoacyl sugars; however, the purification of the final products was challenging. After hydrogenolysis removal of the protecting groups, the reaction mixture was filtered over a diatomite plug to remove the catalysts. The palladium on carbon and platinum oxide catalyst are fine powders that can be removed by filtration. However, filter paper, cotton, or glass wool plug are ineffective due to the small particle size. Therefore diatomite, a very effective filter aid, was used to ensure removal of the catalyst from the reaction mixture. The diatomite plug was washed with methanol to ensure complete recovery of the product. The methanolic solution was concentrated under vacuum until a syrup was obtained. The syrup was placed under an air stream to thoroughly dry the product and obtain a white powder. The NMR of this material (spectroscopic data figures SD 18, 25, 37, 42), shows that the aromatic signals are still present. While the ratio of integrations shows a low concentration of aromatic peaks, this suggests that deprotection may not have been complete.

One of the options to improve the purification of the final product is evaporate the methanol, then, dissolve the residue in a diluted solution of ammonium hydroxide and

extract the solution with an organic solvent such as DCM or ethyl acetate. After the extraction, the aqueous solution can be concentrated under vacuum to eliminate the remaining organic solvent and then lyophilized[40]. Additional options include use of a chromatographic column packed with a mixture 1:1 of activated carbon and diatomite, which has been reported to be useful fractionating complex sugar mixtures[44–46]. Thus, resynthesis of the phosophosugars followed by testing of these two options is the clear next stage for this study.

Phosphorylation is an important biological reaction. We have confirmed that the kinases involved in the sialic acid catabolic pathway have a high specificity towards their cognate substrates. This limits the scope of substrates to be used in the synthesis of the nonulosonic acids. For this reason, the development of a synthetic strategy to install a phosphate group at a specific position in the sugar became mandatory. We were able to design a synthesis that boasts a very low number of protecting group installation and deprotection steps. A key aspect of this methodology is the selective reactivity of thisphosphorylation reagent towards primary alcohols, such as C6 of hexopyranoses. The final challenge to address is the high polarity of the final product, which makes purification extremely difficult. Once overcome r, this synthetic approach is likely to provide rapid access to important phosphosugars for biochemical applications.

References

1. Lahiry, P., Torkamani, A., Schork, N. J. & Hegele, R. a. Kinase mutations in human disease: interpreting genotype-phenotype relationships. *Nat. Rev. Genet.* **11**, 60–74 (2010).
2. Bleeker, F. E. *et al.* Mutational profiling of kinases in glioblastoma. *BMC Cancer* **14**, 718 (2014).
3. McKEE, T. & McKEE, J. Carbohydrate metabolism. *Biochem. Mol. Basis Life* 43 (2011).
4. Merino, F. & Guixé, V. On the Specialization History of the ADP-Dependent Sugar Kinase Family. in *Gene Duplication* (ed. Friedberg, F.) 237–256 (2011).
5. Bork, P., Sander, C. & Valencia, A. *Convergent evolution of similar enzymatic function on different protein folds: the hexokinase, ribokinase, and galactokinase families of sugar kinases. Protein science : a publication of the Protein Society* **2**, (1993).
6. Boulanger, A. *et al.* Identification and regulation of the N-acetylglucosamine utilization pathway of the plant pathogenic bacterium Xanthomonas campestris pv. campestris. *J. Bacteriol.* **192**, 1487–1497 (2010).
7. Berger, M., Chen, H., Reutter, W. & Hinderlich, S. Structure and function of N-acetylglucosamine kinase. Identification of two active site cysteines. *Eur. J. Biochem.* **269**, 4212–4218 (2002).
8. Vollmer, W. & Bertsche, U. Murein (peptidoglycan) structure, architecture and biosynthesis in Escherichia coli. *Biochim. Biophys. Acta - Biomembr.* **1778**, 1714–1734 (2008).

9. Uehara, T. & Park, J. T. The N -Acetyl-d-Glucosamine Kinase of Escherichia coli and Its Role in Murein Recycling The N -Acetyl- D -Glucosamine Kinase of Escherichia coli and Its Role in Murein Recycling. **186**, 7273–7279 (2004).

10. Cheng, Q., Li, H., Merdek, K. & Park, J. T. Molecular characterization of the beta-N-acetylglucosaminidase of Escherichia coli and its role in cell wall recycling. *J. Bacteriol.* **182**, 4836–4840 (2000).

11. Vötsch, W. & Templin, M. F. Characterization of a β-N-acetylglucosaminidase of Escherichia coli and elucidation of its role in muropeptide recycling and β-lactamase induction. *J. Biol. Chem.* **275**, 39032–39038 (2000).

12. Lange, C. F. & Kohn, P. Substrate specificity of hexokinases. *J. Biol. Chem.* **236**, 1–5 (1961).

13. Zhang, Y., Zagnitko, O., Rodionova, I., Osterman, A. & Godzik, A. The FGGY carbohydrate kinase family: Insights into the evolution of functional specificities. *PLoS Comput. Biol.* **7**, (2011).

14. Larion, M., Moore, L. B., Thompson, S. M. & Miller, B. G. Divergent evolution of function in the ROK sugar kinase superfamily: Role of enzyme loops in substrate specificity. *Biochemistry* **46**, 13564–13572 (2007).

15. Konopka, J. B. N-acetylglucosamine (GlcNAc) functions in cell signaling. *Scientifica (Cairo).* **2012**, 631–632 (2012).

16. Uehara, T. *et al.* Recycling of the Anhydro- N -Acetylmuramic Acid Derived from Cell Wall Murein Involves a Two-Step Conversion to N Recycling of the Anhydro-N -Acetylmuramic Acid Derived from Cell Wall Murein Involves a Two-Step Conversion to N -Acetylglucosamine-Phosphat. *J. Bacteriol.* **187**, 3643–3649

(2005).

17. Plumbridge, J. An alternative route for recycling of N-acetylglucosamine from peptidoglycan involves the N-acetylglucosamine phosphotransferase system in Escherichia coli. *J. Bacteriol.* **191**, 5641–5647 (2009).

18. Park, J. K., Wang, L. X. & Roseman, S. Isolation of a glucosamine-specific kinase, a unique enzyme of Vibrio cholerae. *J. Biol. Chem.* **277**, 15573–15578 (2002).

19. Ichikawa, Y. & Lee, Y. C. Synthesis of a branched glycopeptide derivative containing terminal D-mannose 6-phosphate residues. *Carbohydr. Res.* **198**, 235–246 (1990).

20. Wagner, G. K., Pesnot, T. & Field, R. a. A survey of chemical methods for sugar-nucleotide synthesis. *Nat. Prod. Rep.* **26**, 1172–1194 (2009).

21. Durka, M. *et al.* Systematic synthesis of inhibitors of the two first enzymes of the bacterial heptose biosynthetic pathway: Towards antivirulence molecules targeting lipopolysaccharide biosynthesis. *Chem. - A Eur. J.* **17**, 11305–11313 (2011).

22. Davis, B. G. & Boyer, V. Biocatalysis and enzymes in organic synthesis. *Nat. Prod. Rep.* **18**, 618–640 (2001).

23. Koeller, K. M. & Wong, C. H. Enzymes for chemical synthesis. *Nature* **409**, 232–240 (2001).

24. Wang, P. G. Sugars synthesized in a snap. *Nat. Chem. Biol.* **3**, 309–310 (2007).

25. Fernandez-Lorente, G. *et al.* Regio-selective deprotection of peracetylated sugars via lipase hydrolysis. *Tetrahedron* **59**, 5705–5711 (2003).

26. Nicholas, S. D. & Smith, F. Acetylation of Sugars : Abstract : Nature.

27. Wang, C.-C., Kulkarni, S. S., Lee, J.-C., Luo, S.-Y. & Hung, S.-C. Regioselective

one-pot protection of glucose. *Nat. Protoc.* **3**, 97–113 (2008).

28. Miljkovic, M. *Carbohydrates*. (Springer New York, 2009). doi:10.1007/978-0-387-92265-2

29. Guo, J. & Ye, X. S. Protecting groups in carbohydrate chemistry: Influence on stereoselectivity of glycosylations. *Molecules* **15**, 7235–7265 (2010).

30. Jones, G. B., Lin, Y., Xiao, Z., Kappen, L. & Goldberg, I. H. Molecular probes of DNA bulges: Functional assay and spectroscopic analysis. *Bioorganic Med. Chem.* **15**, 784–790 (2007).

31. Hasuoka, A. *et al.* Total synthesis of Novel antibiotics pyloricidin A, B and C and their application in the study of Pyloricidin derivatives. *J. Antibiot. (Tokyo).* **55**, 191–203 (2002).

32. Sabesan, S. & Neira, S. Synthesis of glycosyl phosphates and azides. *Carbohydr. Res.* **223**, 169–185 (1992).

33. Chełmecka, E., Pasterny, K., Gawlik-Jedrysiak, M., Szeja, W. & Wrzalik, R. Theoretical and experimental studies on methyl α-d-glucopyranoside derivatives. *J. Mol. Struct.* **834**–**836**, 498–507 (2007).

34. Szabo, P. Phosphorylated sugars. *J. Chem. Soc. Perkin Trans. 1* **4**, 2–7 (1989).

35. Summeren, R. P. Van, Moody, D. B., Feringa, B. L. & Minnaard, A. J. Total Synthesis of Enantiopure - D -Mannosyl Phosphomycoketides from Mycobacterium tuberculosis. 4546–4547 (2006).

36. Fairweather, J. K., Karoli, T. & Ferro, V. The synthesis of phosphorylated disaccharide components of the extracellular phosphomannan of Pichia (Hansenula) holstii NRRL Y-2448. *Bioorganic Med. Chem.* **12**, 6063–6075 (2004).

37. Kim, B. S., Kim, B. T. & Hwang, K. J. A practical method to cleave diphenyl phosphonate esters to their corresponding phosphonic acids in one step. *Bull. Korean Chem. Soc.* **30**, 1391–1393 (2009).

38. Liu, J., Shen, G. & Ichikawa, Y. Overproduction of CMP-sialic acid synthetase for organic synthesis. *J. Am. Chem. Soc.* **114**, 3901–3910 (1992).

39. Reck, F., Marmor, S., Fisher, S. & Wuonola, M. A. Inhibitors of the bacterial cell wall biosynthesis enzyme MurC. *Bioorg. Med. Chem. Lett.* **11**, 1451–1454 (2001).

40. Huestis, M. P., Aish, G. a, Hui, J. P. M., Soo, E. C. & Jakeman, D. L. Lipophilic sugar nucleotide synthesis by structure-based design of nucleotidylyltransferase substrates. *Org. Biomol. Chem.* **6**, 477–484 (2008).

41. Maegawa, T., Akashi, A. & Sajiki, H. A mild and facile method for complete hydrogenation of aromatic nuclei in water. *Synlett* 1440–1442 (2006). doi:10.1055/s-2006-939719

42. Corey, E. & Link, J. A General, Catalytic, and Enantioselective Synthesis of a-Amino Acids. *J. Am. Chem. Soc.* 1906–1908 (1992).

43. Chandrasekhar, S., Prakash, S. J. & Lohitha, C. Poly (ethylene Glycol) (400) as Superior Solvent Medium against Ionic Liquids for Catalytic Hydrogenations with PtO 2 been found to be a superior solvent over the ionic liquids by severalfold in promoting the hydrogenation of various functional groups. *J. Org. Chem.* **71**, 2196–2199 (2006).

44. Sandoval, M. *et al.* Screening of strains and recombinant enzymes from Thermus thermophilus for their use in disaccharide synthesis. *J. Mol. Catal. B Enzym.* **74**, 162–169 (2012).

45. Nobre, C., Teixeira, J. A. & Rodrigues, L. R. Fructo-oligosaccharides purification from a fermentative broth using an activated charcoal column. *N. Biotechnol.* **29**, 395–401 (2012).

46. Morales, V., Sanz, M. L., Olano, a. & Corzo, N. Rapid Separation on Activated Charcoal of High Oligosaccharides in Honey. *Chromatographia* **64**, 1–6 (2006).

47. Zhang, J., Eisink, N. N. H. M., Witte, M. D. & Minnaard, A. J. Regioselective Manipulation of GlcNAc Provides Allosamine, Lividosamine, and Related Compounds. *J. Org. Chem.* **84**, 516–525 (2019).

48. Janiak, A. M., Hoffmann, M., Milewska, M. J. & Milewski, S. Hydrophobic derivatives of 2-amino-2-deoxy-D-glucitol-6-phosphate: A new type of D-glucosamine-6-phosphate synthase inhibitors with antifungal action. *Bioorganic Med. Chem.* **11**, 1653–1662 (2003).

49. Zhang, Z., Ren, S., Wan, S., Li, W. & Jiang, T. Synthesis of D-Glucosamine modified Benzo d 1,2 selenazol-3-(2H)-one derivatives. *Synth. Commun.* **40**, 3438–3446 (2010).

3.3 Experimental Section

3.3.1 Plasmid Production

Two strains of chemical competent *E. coli* Top 10 were transformed with plasmids pBRL07 and pSRH04 that carries nanK and nagK respectively. The strains were placed on a glass test tube and incubated on ice for 5 min. 5 μL of plasmid solution were placed in the bacterial suspension and incubated for 30 min on ice.

After the incubation period the test tubes were placed in a 42 °C water bath for 45 s and returned to the ice bath for 2 min. 1 mL of LB broth was added to each tube and they were incubated at 37 °C for half an hour; then, the bacterial suspension was inoculated on LB plates supplemented with kanamycin and incubated overnight at 37 °C.

Single colonies were taken and inoculated in 5 mL of LB broth supplemented with kanamycin and incubated for 18 h. After the incubation time, the cells were spun down and the plasmids were extracted using the P1, P2, P3 protocol.

3.3.2 Gene expression and protein extraction from *E. coli*

To produce the seed culture one single colony of the transformants was used to inoculate 10 mL of LB broth supplemented with kanamycin. The culture was incubated at 37 °C and 200 rpm overnight. Next day, 400 mL of fresh LB were supplemented with 400 μL of 1000 × kanamycin solution and inoculated with the seed culture. The bacterial culture was incubated at 37 °C with shaking until OD_{600} reaches 0.3-0.6.

When the proper optical density was reached, IPTG 1M was added to a final concentration of 0.5 mM and the culture was placed in the incubator at 30 °C with shaking.

After 12 to 18 hours of incubation, the culture was centrifuged at 5000 × g for 20 min. at 4 °C. Cells were resuspended in 50 mL of lysis buffer[9] and placed in an ice bath. The cells were lysed by sonication using 5 pulses of 20 s with 30 s intervals of cooling.

The cell suspension was centrifuged at 9000×g for 1 h and the supernatant was transferred to a clean falcon tube.

The recombinant proteins were purified by Ni affinity chromatography.

3.3.3 Nickel purification for 6X-His tagged proteins

For our lab purposes, proteins are generally expressed using pET-vectors (Novagen). These proteins are equipped with a 6X-His tag, which adopts a conformation that is optimal for coordination to divalent nickel ions, Ni^{2+}. We use Ni-NTA superflow (Qiagen) for the resin, and the amount of this resin to the lysate containing the tagged protein should be determined empirically.

The lysate from the protein production was incubated with nickel resin, Ni-NTA superflow, for 1-4 hours at 4 °C with gentle rocking/shaking. After the incubation, the mixture was poured into a column and the flow-through was collected and identified. The column was washed with 10 × resin volume using elution buffer[10] 0 mM imidazole and the fraction was collected and labeled 0 mM. Then the column was wash/eluted with increasing concentration of imidazole (10× 20 mM, 5× 100 mM twice and 5× 250 mM twice). The

---

[9] Lysis buffer = 100 mM sodium phosphate, 300 mM NaCl, 10% (v/v) glycerol, 1 mg/mL lysozyme, 1 µg/mL pepstatin A, 1—2 µg/mL leupeptin, pH 8.0

[10] Solutions: 100 mM Tris, 300 mM NaCl, X mM imidazole (X = 0, 20, 100, 250), pH 7.4

Adjust to pH 7.4 after the addition of imidazole.

fractions were kept in ice to prevent protein denaturation and the fractions were analyzed by SDS-PAGE.

3.3.4 Determination of kinase activity through quantification of unreacted ATP

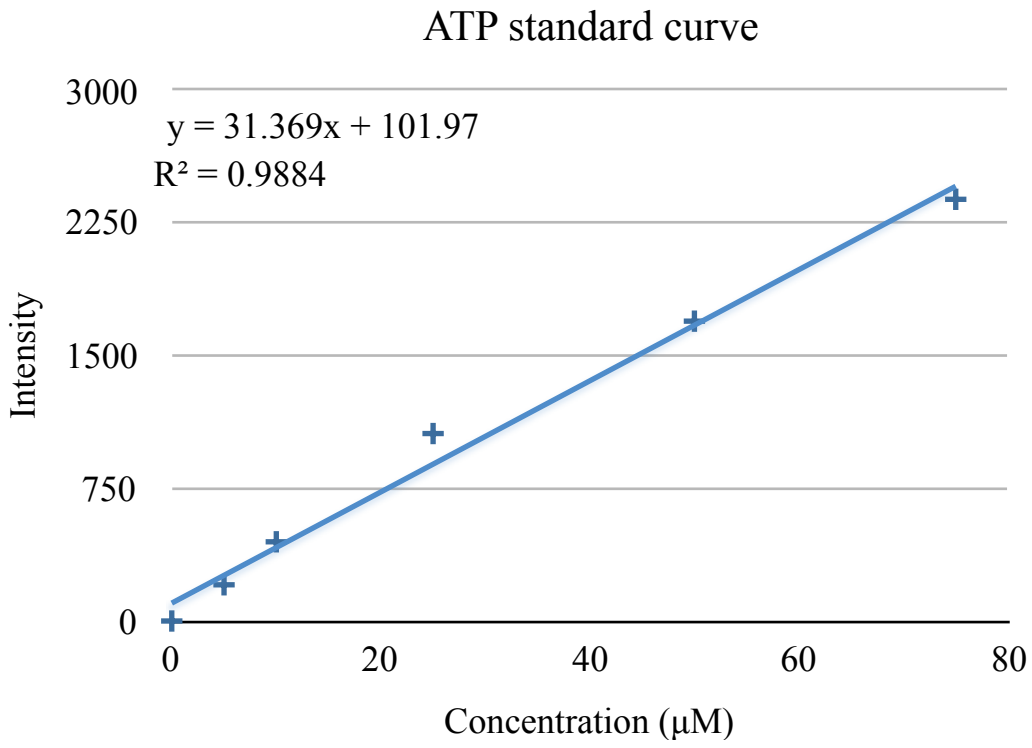
This test allows determining indirectly how active a kinase is through the quantification of the ATP that remains in solution.

200 µL enzymatic reactions were set up in triplicate with the following conditions: phosphate buffer pH 7.4 50 mM, NaCl 50 µM, ATP 1.2 mM, sugar 1 mM, kinase 3µM. The reactions were incubated for 10 minutes at room temperature.

ATP quantification was achieved using the Adenosine 5′-triphosphate (ATP) Bioluminescent Assay Kit from Sigma-Aldrich Co.

In a 96 well plate, 100 µL of ATP assay mix solution were placed and allow to stand at room temperature for approximately 3 minutes, then, 100 µL of the sample or blank were added to the well and mixed thoroughly. The plate was placed on a plate reader under luminometer configuration to measure the amount of light produced.

To be able to quantify the concentration of remnant ATP, an ATP standard curve was made.

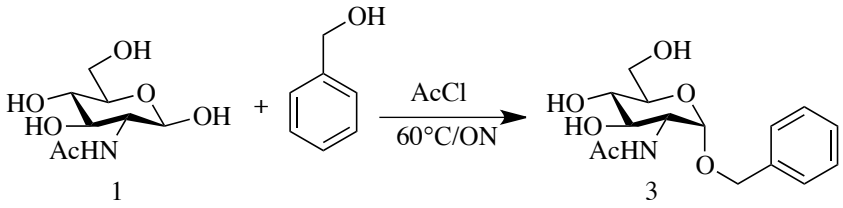


ATP standard curve made with luciferase assay

3.3.5 Phosphosugar synthesis General procedures

Anomeric protection with benzyl alcohol

O-benzyl-2-acetylamino-2-deoxy-D-glucopyranose



N-acetylglucosamine (3 g, 13.6 mmol, 1 eq.) were placed in a 100 mL round bottom flask containing 25 mL of benzyl alcohol. 0.5 eq. of acetyl chloride were added, and a condenser was placed in vertical position. The system was heated at 60 °C overnight.

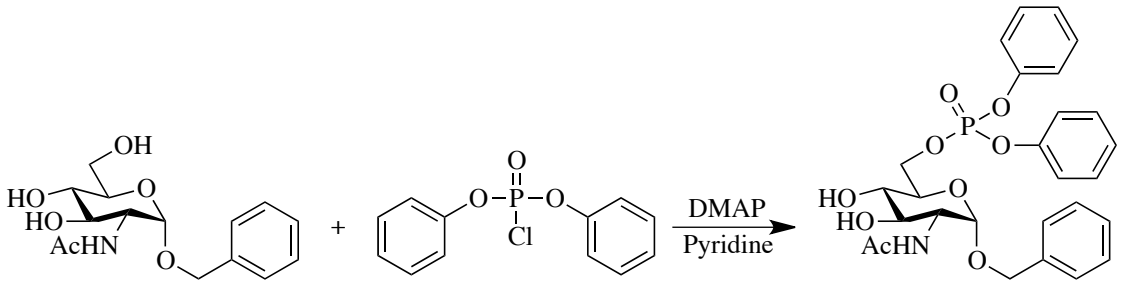
Next day, the reaction mixture was poured over 250 mL of ether and placed at -40 °C for two hours. The precipitate was filtered and washed with ether. The solid was dried and

purified by column chromatography using DCM:MeOH 5% to yield 2.55 g (60%) as a mixture of anomers (R_f: 0.22 in DCM:MeOH 10%).

^{1}H-NMR (400 MHz; MeOD): δ 7.42-7.30 (m, 5H), 4.87 (s, 1H), 4.76 (d, *J* = 12.0 Hz, 1H), 4.52 (d, *J* = 12.0 Hz, 1H), 3.92-3.83 (m, 2H), 3.75-3.66 (m, 3H), 3.38 (t, *J* = 9.2 Hz, 1H), 1.95 (s, 3H). This data is consistent with the literature values[47].

Sugar phosphorylation

O-benzyl-6-diphenoxyphosphoryl-2-acetylamino-2-deoxy-D-glucopyranose[34]



Benzyl protected GlcNAc (486 mg, 1.56 mmol, 1 eq) was placed in a 25 mL oven dry round bottom flask; 19 mg (0.16 mmol, 0.1 eq) of DMAP and 7 mL of dry pyridine were added to the flask and mixed until dissolved. Then, 462 mg of diphenyl phosphoryl chloride (1.7 mmol, 1.1 eq) were added and the reaction mixture was stirred overnight under argon atmosphere.

Next day, the reaction mixture was diluted with 50 mL of EtOAc and washed 6 x 50 mL with brine. The organic phase was dried with $MgSO_4$ and concentrated under vacuum. The crude was purified by column chromatography using DCM:MeOH 2% to yield 267 mg (30%) as a mixture of anomers (R_f: 0.2 in DCM:MeOH 5%).

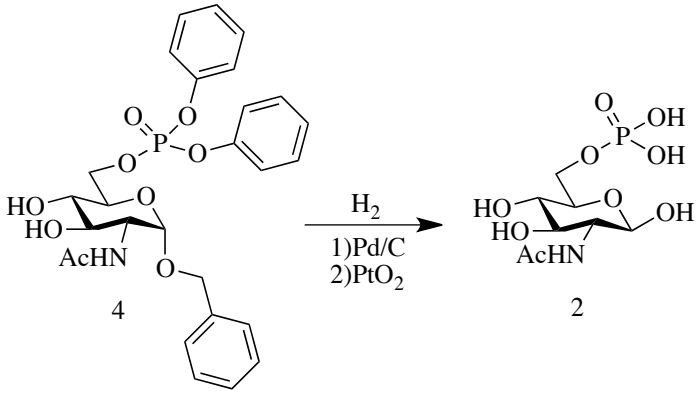
^{1}H-NMR (300 MHz; MeOD; Me_4Si) 7.37-7.44 (m, 1 H), 7.22-7.33 (m, 3 H), 4.82 (d, *J* = 3.6 Hz), 4.65 (d, *J* 12.0), 4.57 (ddd, *J* = 11.0, 7.1 and 2.0 Hz), 4.42-4.50 (m), 4.43 (d, *J* =

12.0 Hz), 3.91 (dd, *J* = 10.8 and 3.6 Hz), 3.85 (d, *J* = 5.4 Hz), 3.72 (dd, *J* = 10.8 and 8.7 Hz), 3.40 (dd, *J* = 10.0 and 8.7 Hz), 1.96 (1 H, s).

^{13}C-NMR (101 MHz; MeOD; Me_4Si) 172.16, 150.45 (d, *J* = 7.3 Hz), 137.32, 129.66, 128.00, 127.92, 127.50, 125.43 (d, *J* = 0.9 Hz), 119.78 (d, *J* = 4.9 Hz), 96.18, 71.15, 70.86 (d, *J* = 6.7 Hz), 70.42, 69.02, 68.50 (d, *J* = 6.3 Hz), 53.79, 21.11.

Hydrogenolysis

N-Acetyl-D-Glucosamine-6-phosphate

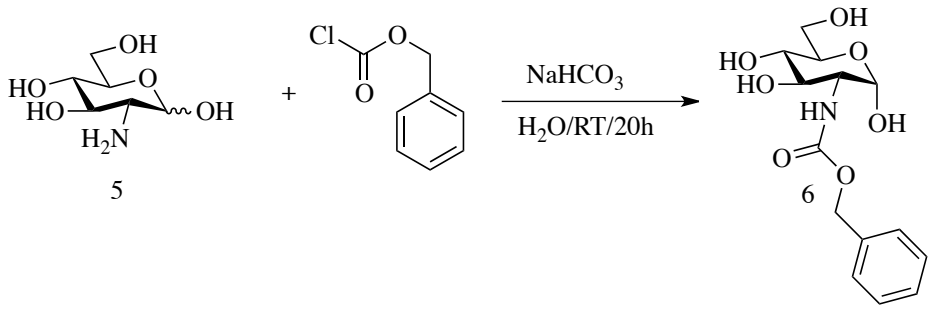


The protected sugar (131 mg) were placed in a 100 mL round bottom flask containing 10 mL of MeOH. 50 mg of Pd/C 10% were added to the solution. The solution was degassed under vacuum and saturated with hydrogen under 1 atm. pressure. The reaction was carried out for 12 to 18 h. After the 18 h, the mixture was filter through a pad of diatomite and 50 mg of PtO_2 were added.

^{1}H-NMR (400 MHz; D_2O): δ 5.06 (s,), 4.60 (d, *J* = 15.9 Hz, 1H), 4.04-3.99 (m, 2H), 3.74-3.39 (m, 6H), 1.90 (s, 3H).

Amine group protection with CBZ

2-carbobenzyloxyamino-2-deoxy-D-glucopyranose[30,48,49]

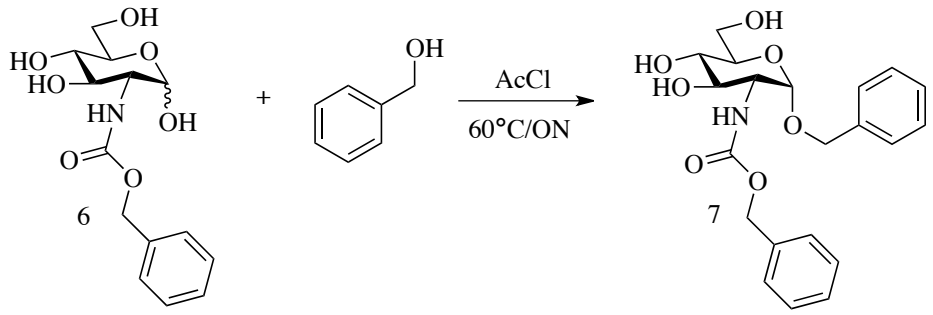


Glucosamine hydrochloride (1 g, 4.6 mmol, 1 eq) and benzyl chloroformate (722µL, 5,06 mmol, 1.1 eq) were added to a round bottom flask containing 30 mL of distilled water with sodium bicarbonate (1.2 g, 14 mmol, 3 eq). The mixture was stirred 20 hours at room temperature until completion by TLC (silica gel, DCM:MeOH 15%). The reaction mixture was extracted with 3x50 mL of THF/EtOAc 2:1. The organic phase was dried with $MgSO_4$ and then the solvent was evaporated *in vacuo* to yield 998 mg of a white solid (89-53%, Rf: 0.4). HRMS calcd for $C_{14}H_{19}NO_7Na$ 336.1059; found 336.1010 Da

^{1}H-NMR (300 MHz; MeOD): δ 7.41-7.29 (m, 6H), 5.13 (d, J = 3.3 Hz, 1H), 5.10 (s, 2H), 4.58 (d, J = 8.1 Hz,), 3.79-3.59 (m, 7H).

Anomeric protection with benzyl alcohol

O-benzyl-2-carbobenzyloxyamino-2-deoxy-D-glucopyranose[30,31]



CBZGlcNH (2.23 g , 7.1 mmol, 1 eq) was placed in a 250 mL round bottom flask containing 30 mL of benzyl alcohol and 0.5 eq of acetyl chloride were added and a condenser was placed in reflux position. The system was heated at 60 °C overnight.

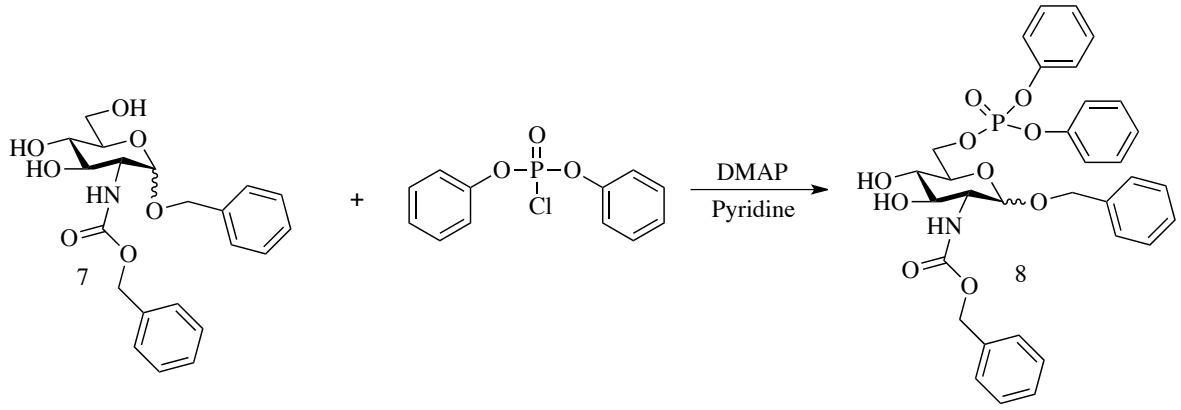
Next day, the reaction mixture was poured over 300 mL of ether and placed at -40 °C for two hours. The precipitate was filtered and washed with ether. The solid was dried and purified by column chromatography using DCM:MeOH 10% to yield 1.9 g (66%) as a mixture of anomers (Rf: alpha anomer 0.4, beta anomer 0.3). HRMS calcd for $C_{21}H_{25}NO_7Na$ 426.1529; found 426.1514 Da.

^{1}H-NMR beta anomer (400 MHz; MeOD): δ 7.38-7.26 (m, 11H), 5.51 (s,), 5.18-5.04 (m, 2H), 4.90 (d, *J* = 12.2 Hz, 2H), 4.62 (d, *J* = 12.1 Hz, 1H), 4.47 (d, *J* = 7.0 Hz, 1H), 3.92 (dd, *J* = 11.9, 2.2 Hz, 1H), 3.72 (dd, *J* = 11.9, 5.8 Hz, 1H), 3.50-3.42 (m, 2H), 3.28 (ddd, *J* = 9.2, 6.3, 2.6 Hz, 1H).

^{1}H-NMR alpha anomer (400 MHz; MeOD): δ 7.39-7.27 (m, 10H), 5.09 (q, *J* = 9.6 Hz, 2H), 4.90 (d, *J* = 3.2 Hz, 1H), 4.75 (d, *J* = 12.1 Hz, 1H), 4.52 (d, *J* = 12.1 Hz, 1H), 3.84 (dd, *J* = 11.4, 1.8 Hz, 1H), 3.72 (t, *J* = 5.6 Hz, 1H), 3.69-3.62 (m, 3H), 3.39 (d, *J* = 9.2 Hz, 1H).

Sugar phosphorylation

O-benzyl-6-diphenoxyphosphoryl-2-carbobenzyloxyamino-2-deoxy-D-glucopyranose



Benzyl anomeric protected CBZGlcNH (500 mg, 1.24 mmol, 1 eq) was placed in a 25 mL oven dry round bottom flask. 30.3 mg (0.25 mmol, 0.2 eq) of DMAP and 6 mL of dry pyridine were added to the flask and mixed until dissolved. Then, 367 mg of diphenyl phosphoryl chloride (1.36 mmol, 1.1 eq) were added and the reaction mixture was stirred overnight under argon atmosphere.

Next day, the reaction mixture was diluted with 50 mL of EtOAc and washed 6 x 50 mL with brine. The organic phase was dry with $MgSO_4$ and concentrated under vacuum. The crude was purified by column chromatography using DCM:MeOH 2% to yield 502 mg (64%) as a mixture of anomers (R_f alpha anomer 0.3, beta anomer 0.2 in DCM:MeOH 5%).

HRMS calcd for $C_{28}H_{32}NO_9PNa$ 658.1818; found 658.1825 Da

^{1}H-NMR alpha anomer (400 MHz; MeOD): δ 7.38-7.17 (m, 20H), 5.45 (s,), 5.16-5.03 (m, 2H), 4.79 (d, *J* = 12.1 Hz, 1H), 4.67 (t, *J* = 8.5 Hz, 1H), 4.50 (dd, *J* = 15.7, 8.3 Hz, 3H), 3.56 (t, *J* = 8.6 Hz, 3H), 3.39 (dd, *J* = 37.7, 11.9 Hz, 1H).

^{31}-P NMR (122 MHz; MeOD): δ -11.9

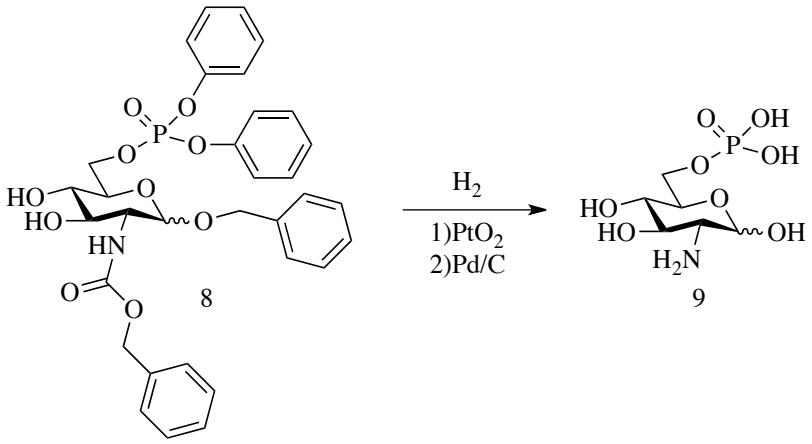
^{1}H-NMR beta anomer (400 MHz; MeOD): δ 7.43-7.22 (m, 19H), 7.17 (dd, *J* = 8.6, 7.4 Hz, 1H), 6.81-6.77 (m,), 5.50 (s, 1H), 5.09 (q, *J* = 9.4 Hz, 2H), 4.84 (d, *J* = 3.2 Hz, 1H), 4.63 (d, *J* = 12.0 Hz, 1H), 4.56 (ddd, *J* = 10.8, 7.0, 1.8 Hz, 1H), 4.49-4.42 (m, 2H), 3.87-3.83 (m, 1H), 3.70-3.61 (m, 2H), 3.42-3.37 (m, 1H).

^{31}-P NMR (122 MHz; MeOD): δ -11.9

^{13}C NMR (101 MHz; MeOD) *δ* 175.92 (s,), 150.45 (d, *J* = 7.3 Hz, 1C), 137.29 (s,), 129.67 (s, 3C), 127.98 (d, *J* = 6.3 Hz, 4C), 127.51 (s, 1C), 125.44 (d, *J* = 1.2 Hz, 2C), 119.79 (d, *J* = 4.7 Hz, 3C), 96.14 (s, 1C), 71.15 (s, 1C), 70.86 (d, *J* = 6.8 Hz, 1C), 70.44 (s, 1C), 68.95 (s, 1C), 68.53 (d, *J* = 6.3 Hz, 1C), 53.67 (s, 1C), 28.59 (s, 1C), 9.03 (s, 1C).

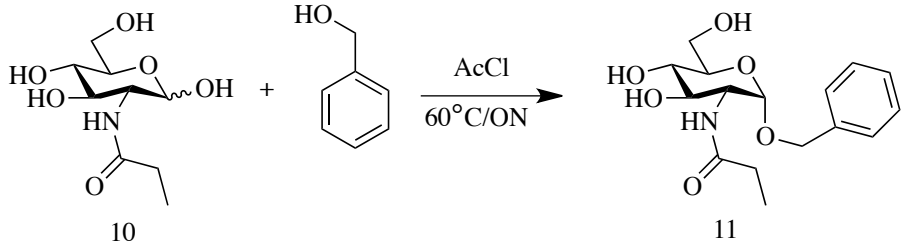
Hydrogenolysis

Glucosamine-6-phosphate



The previous procedures were followed to synthesize the following compounds

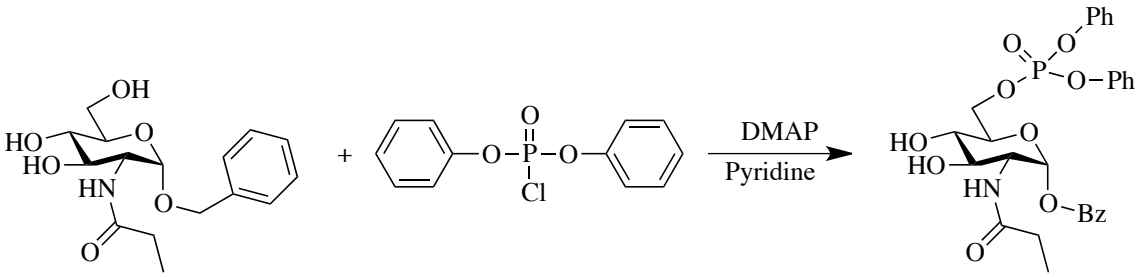
O-benzyl-2-propylamino-2-deoxy-glucopyranose



For the reaction 100 mg of 10, 120 µL of AcCl were dissolved in 10 mL of benzyl alcohol. The procedure for the anomeric protection reaction was followed to yield 47 mg of 11 (34%).

^{1}H-NMR (400 MHz; MeOD): δ 7.41-7.28 (m, 5H), 4.87 (s, 1H), 4.76 (d, *J* = 12.0 Hz, 1H), 4.52 (d, *J* = 12.0 Hz, 1H), 3.89 (td, *J* = 13.0, 5.1 Hz, 2H), 3.76-3.66 (m, 3H), 3.38 (t, *J* = 9.2 Hz, 1H), 2.24 (q, *J* = 7.4 Hz, 2H), 1.11 (t, *J* = 7.6 Hz, 3H).

O-benzyl-6-diphenoxyphosphoryl-2-propylamino-2-deoxy-glucopyranose



11 12

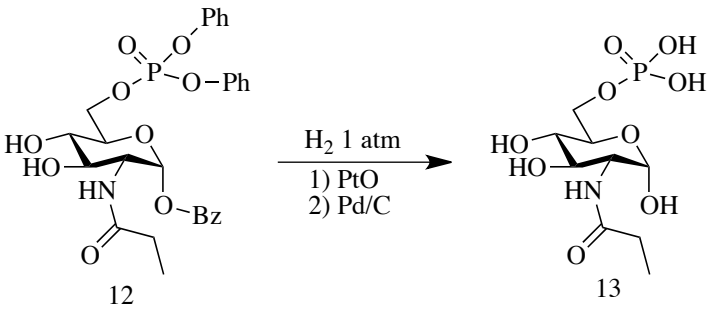
For this reaction, the phosphorylation reaction procedure was followed; 55 mg of 11, 50.3 mg (1.1 eq) of DPCP, 2 mg (0.1 eq) of DMAP and 1 mL of pyridine to yield 23 mg of 12 (39%).

^{1}H-NMR (400 MHz; MeOD): δ 7.41 (td, *J* = 7.9, 3.0 Hz, 4H), 7.32-7.24 (m, 11H), 4.83 (d, *J* = 3.6 Hz, 1H), 4.64 (d, *J* = 12.0 Hz, 1H), 4.58 (ddd, *J* = 10.9, 7.0, 1.8 Hz, 1H), 4.48 (dd, *J* = 7.9, 5.6 Hz, 1H), 4.43 (d, *J* = 12.0 Hz, 1H), 3.90 (td, *J* = 12.4, 4.0 Hz, 2H), 3.73 (t, *J* = 9.7 Hz, 1H), 3.40 (t, *J* = 9.4 Hz, 1H), 2.24 (qd, *J* = 7.6, 1.5 Hz, 2H), 1.11 (t, *J* = 7.6 Hz, 3H).

^{13}C-NMR (101 MHz; MeOD): δ 175.92 (s), 150.45 (d, *J* = 7.3 Hz), 137.29 (s), 129.67 (s), 128.01 (s), 127.95 (s), 127.51 (s), 125.45 (s), 125.44 (s), 119.82 (s), 119.77 (s), 96.14 (s), 71.15 (s), 70.86 (d, *J* = 6.8 Hz), 70.44 (s), 68.95 (s), 68.53 (d, *J* = 6.3 Hz), 53.67 (s), 28.59 (s), 9.03 (s).

Hydrogenolysis

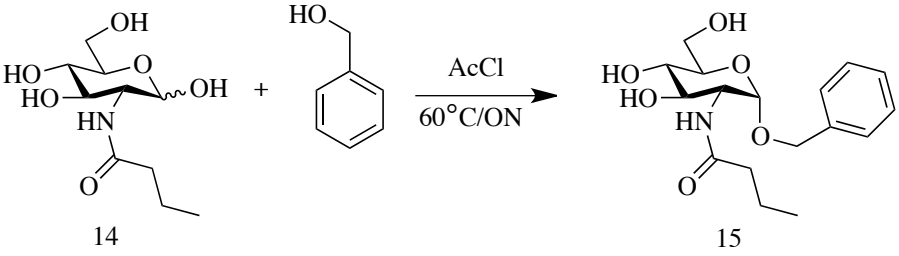
N-propyl-D-glucosamine-6-phosphate



For this reaction, the hydrogenolysis reaction procedure was followed; 23 mg of 12, 15 mg of Pd/C, 15 mg of PtO_2 and 5 mL of MeOH to yield 20 mg of crude product.

^{1}H-NMR (300 MHz; D_2O): δ 5.32 (s, 1H), 3.92 (d, *J* = 17.2 Hz, 1H), 3.80 (d, *J* = 5.3 Hz, 2H), 3.62 (d, *J* = 9.2 Hz, 2H), 3.39 (s, 1H), 3.19 (d, *J* = 21.3 Hz, 1H), 2.17 (s, 1H), 0.99 (s, 3H).

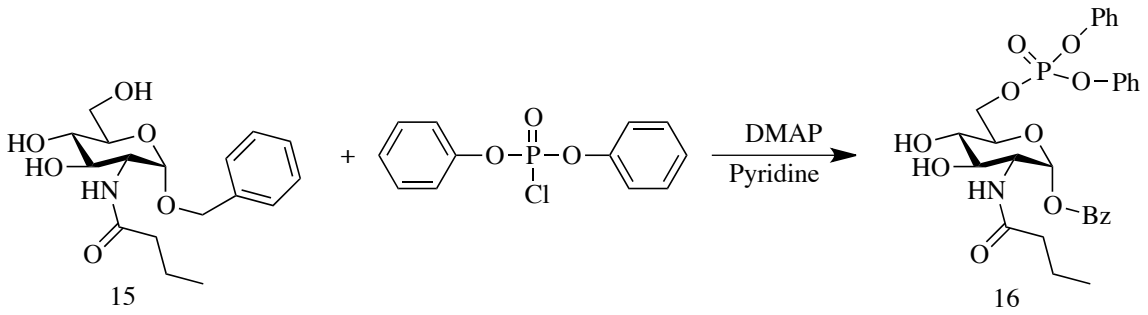
O-benzyl-2-*n*-butylamino-2-deoxy-glucopyranose



For the reaction 200 mg of 14, 120 µL of AcCl were dissolved in 10 mL of benzyl alcohol. The procedure for the anomeric protection reaction was followed to yield 113 mg of 15 (41%).

^{1}H-NMR (300 MHz; MeOD): δ 7.41-7.29 (m, 5H), 4.87 (d, *J* = 3.6 Hz, 1H), 4.75 (d, *J* = 12.0 Hz, 1H), 4.50 (d, *J* = 12.0 Hz, 1H), 3.94-3.69 (m, 5H), 3.39 (d, *J* = 9.0 Hz, 1H), 2.20 (t, *J* = 7.5 Hz, 2H), 1.61 (sextet, *J* = 7.4 Hz, 2H), 0.94 (t, *J* = 7.4 Hz, 3H).

O-benzyl-6-diphenoxyphosphoryl-2-*n*-butylamino-2-deoxy-glucopyranose



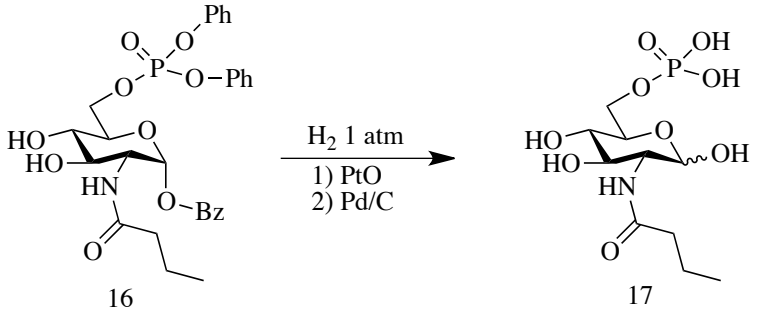
15 16

For this reaction, the phosphorylation reaction procedure was followed; 97 mg of 15, 86 mg (1.1 eq) of DPCP, 3.5 mg (0.1 eq) of DMAP and 1.5 mL of pyridine to yield 99 mg of 16 (61%).

^{1}H-NMR (400 MHz; MeOD): δ 7.40 (td, *J* = 7.8, 3.2 Hz, 4H), 7.30-7.22 (m, 11H), 5.49 (s, 1H), 4.84 (d, *J* = 3.5 Hz, 1H), 4.64 (d, *J* = 11.9 Hz, 1H), 4.61-4.44 (m, 2H), 4.41 (d, *J* = 12.0 Hz, 1H), 3.91 (ddd, *J* = 21.7, 10.3, 4.4 Hz, 2H), 3.74 (t, *J* = 9.7 Hz, 1H), 3.43-3.32 (m, 2H), 2.20 (t, *J* = 7.4 Hz, 2H), 1.62 (sextet, *J* = 7.4 Hz, 2H), 0.94 (t, *J* = 7.4 Hz, 3H).

Hydrogenolysis

N-n-butyl-D-glucosamine-6-phosphate

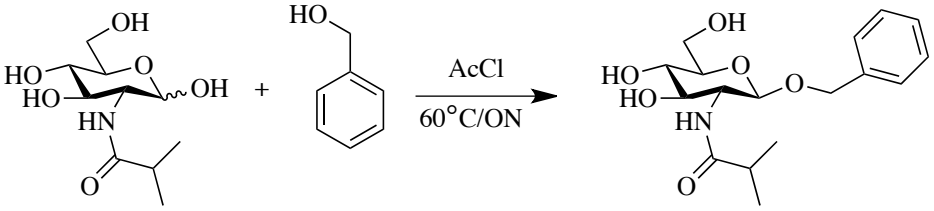


16 17

For this reaction, the hydrogenolysis reaction procedure was followed; 99 mg of 16, 15 mg of Pd/C, 15 mg of PtO_2 and 5 mL of MeOH to yield 61 mg of crude product.

^{1}H-NMR (400 MHz; D_2O): δ 7.31 (t, *J* = 7.9 Hz, 1H), 7.13 (t, *J* = 4.7 Hz, 2H), 5.08 (dd, *J* = 15.1, 3.4 Hz,), 4.75 (d, *J* = 3.5 Hz,), 4.60 (dd, *J* = 12.1, 8.5 Hz,), 4.21-3.76 (m, 3H), 3.70-3.39 (m, 3H), 2.18 (q, *J* = 5.9 Hz, 2H), 1.52 (q, *J* = 7.3 Hz, 2H), 0.82 (td, *J* = 7.4, 3.4 Hz, 3H).

O-benzyl-2-*i*-butylamino-2-deoxy-D-glucopyranose

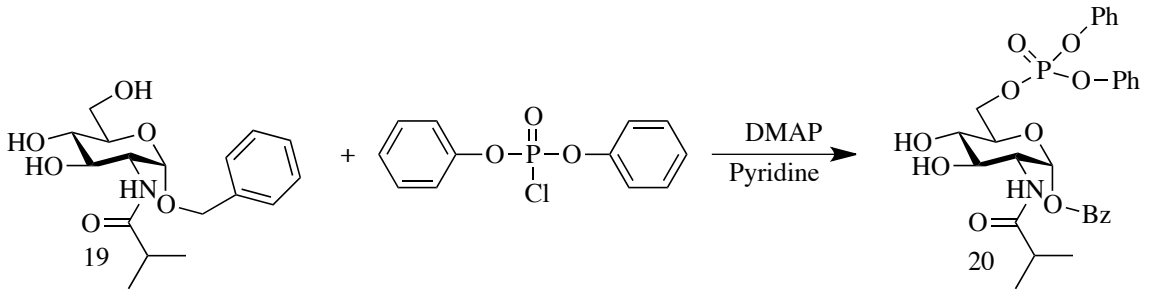


18 19

For the reaction 250 mg of 18 and 120 µL of AcCl were dissolved in 10 mL of benzyl alcohol. The procedure for the anomeric protection reaction was followed to yield 115 mg of 19 (33%).

^{1}H-NMR (400 MHz; MeOD): δ 7.41-7.29 (m, 5H), 4.88 (s, 6H), 4.75 (d, *J* = 12.0 Hz, 1H), 4.51 (d, *J* = 12.0 Hz, 1H), 3.91-3.85 (m, 2H), 3.76-3.67 (m, 3H), 3.38 (t, *J* = 9.2 Hz, 1H), 2.51 (dt, *J* = 13.7, 6.8 Hz, 1H), 1.09 (dd, *J* = 11.8, 6.9 Hz, 6H).

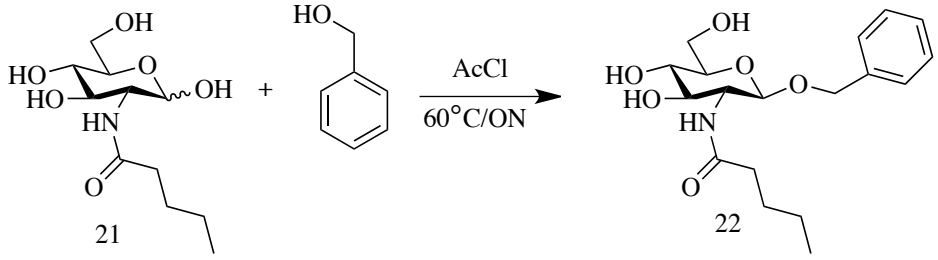
O-benzyl-6-diphenoxyphosphoryl-2-*i*-butylamino-2-deoxy-D-glucopyranose



For this reaction, the phosphorylation reaction procedure was followed; 100 mg of 19, 86 mg (1.1 eq) of DPCP, 3.5 mg (0.1 eq) of DMAP and 1.5 mL of pyridine to yield 38 mg of 20 (22%).

^{1}H-NMR (300 MHz; MeOD): δ 7.44-7.38 (m, 4H), 7.32-7.25 (m, 11H), 4.82 (d, *J* = 3.5 Hz, 1H), 4.64 (d, *J* = 12.0 Hz, 1H), 4.59-4.45 (m, 2H), 4.44-4.40 (m, 1H), 3.88 (dt, *J* = 10.3, 4.9 Hz, 2H), 3.73 (dd, *J* = 10.8, 8.7 Hz, 1H), 3.39 (dd, *J* = 10.0, 8.7 Hz, 1H), 2.51 (7, *J* = 6.9 Hz, 1H), 1.09 (dd, *J* = 9.7, 6.9 Hz, 6H).

O-benzyl-2-*n*-valeroylamino-2-deoxy-glucopyranose



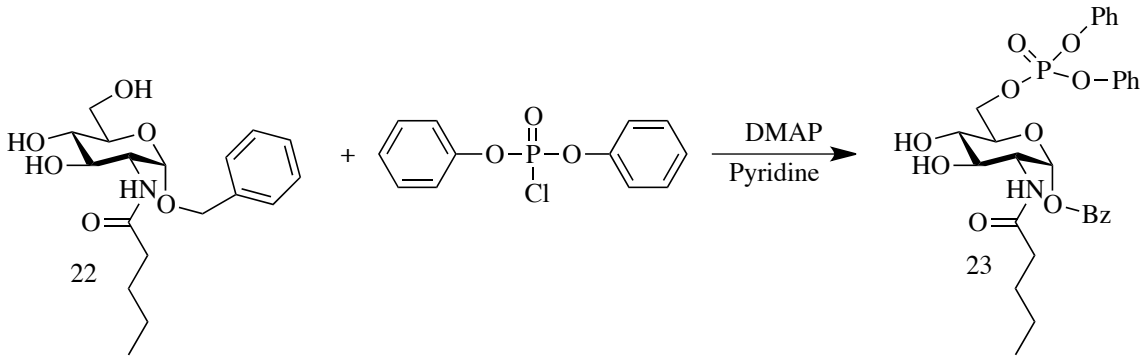
For the reaction 251 mg of 21 and 120 µL of AcCl were dissolved in 10 mL of benzyl alcohol. The procedure for the anomeric protection reaction was followed to yield 139 mg of 22 (41%).

^{1}H-NMR (400 MHz; MeOD): δ 7.34 (ddt, *J* = 22.1, 15.1, 7.5 Hz, 5H), 4.76 (d, *J* = 12.0 Hz, 1H), 4.50 (d, *J* = 11.9 Hz, 1H), 3.93-3.84 (m, 2H), 3.75-3.66 (m, 3H), 3.38 (t, *J* = 9.1

Hz, 1H), 2.23 (t, *J* = 7.5 Hz, 2H), 1.57 (quintet, *J* = 7.6 Hz, 2H), 1.40-1.36 (m, 1H), 0.94 (t, *J* = 7.3 Hz, 3H).

^{13}C-NMR (101 MHz; MeOD): δ 175.22 (s, 1C), 137.54 (s, 1C), 127.97 (s, 1C), 127.91 (s, 1C), 127.42 (s, 1C), 96.11 (s, 1C), 72.69 (s, 1C), 71.20 (s, 1C), 71.06 (s, 1C), 68.70 (s, 1C), 61.34 (s, 1C), 53.90 (s, 1C), 46.52 (s, 1C), 35.32 (s, 1C), 27.78 (s, 1C), 21.98 (s, 1C), 12.77 (s, 1C), 7.84 (s, 1C).

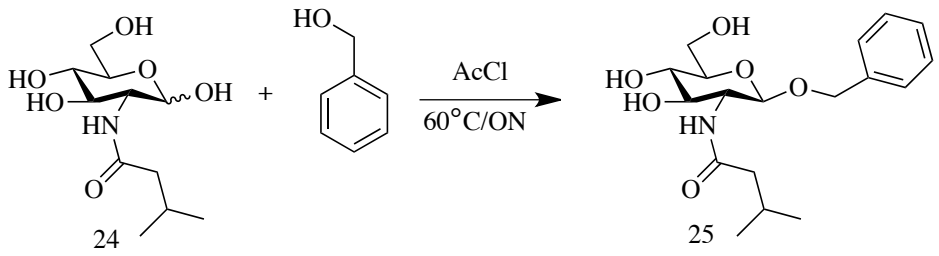
O-benzyl-6-diphenoxyphosphoryl-2-*n*-valeroylamino-2-deoxy-glucopyranose



For this reaction, the phosphorylation reaction procedure was followed; 85 mg of 15, 71 mg (1.1 eq) of DPCP, 3 mg (0.1 eq) of DMAP and 2 mL of pyridine to yield 46 mg of 23 (32%).

^{1}H-NMR (300 MHz; MeOD): δ 7.43-7.37 (m, 4H), 7.32-7.24 (m, 11H), 5.50 (s, 2H), 4.83 (d, *J* = 3.6 Hz, 1H), 4.64 (d, *J* = 11.9 Hz, 1H), 4.61-4.55 (m, 1H), 4.51-4.44 (m, 1H), 4.43-4.39 (m, 1H), 3.89 (ddd, *J* = 13.3, 9.9, 4.0 Hz, 2H), 3.72 (dd, *J* = 10.8, 8.7 Hz, 1H), 3.43-3.39 (m, 1H), 2.22 (t, *J* = 7.5 Hz, 2H), 1.62-1.52 (m, 2H), 1.34 (dq, *J* = 15.1, 7.5 Hz, 3H), 0.93 (t, *J* = 7.3 Hz, 3H).

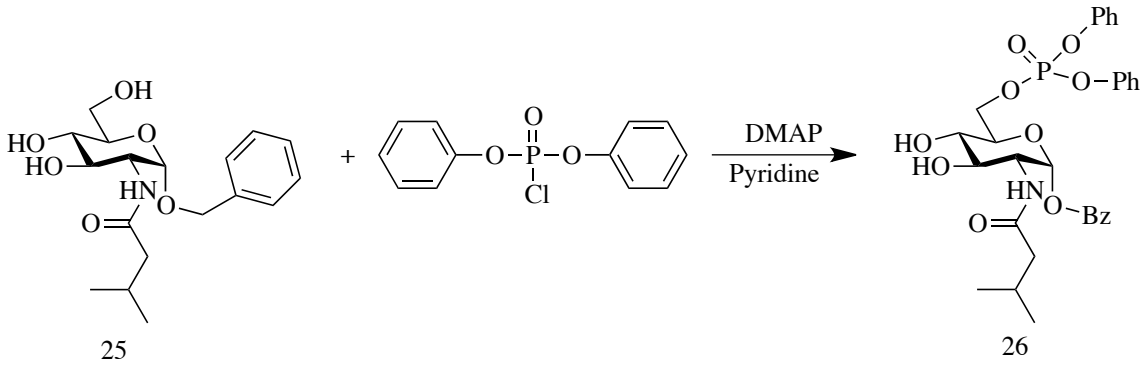
O-benzyl-2-*i*-valeroylamino-2-deoxy-glucopyranose



For the reaction 250 mg of 24 and 120 µL of AcCl were dissolved in 10 mL of benzyl alcohol. The procedure for the anomeric protection reaction was followed to yield 167 mg of 25 (50%).

^{1}H-NMR (400 MHz; MeOD): δ 7.39-7.36 (m, 2H), 7.34-7.27 (m, 4H), 4.87 (s, 2H), 4.74 (d, *J* = 11.8 Hz, 1H), 4.47 (d, *J* = 11.8 Hz, 1H), 3.93-3.89 (m, 1H), 3.83 (d, *J* = 9.8 Hz, 1H), 3.72-3.64 (m, 4H), 3.38-3.34 (m, 2H), 2.09-2.06 (m, 2H), 2.05-1.98 (m, 1H), 0.92 (d, *J* = 13.4 Hz, 7H).

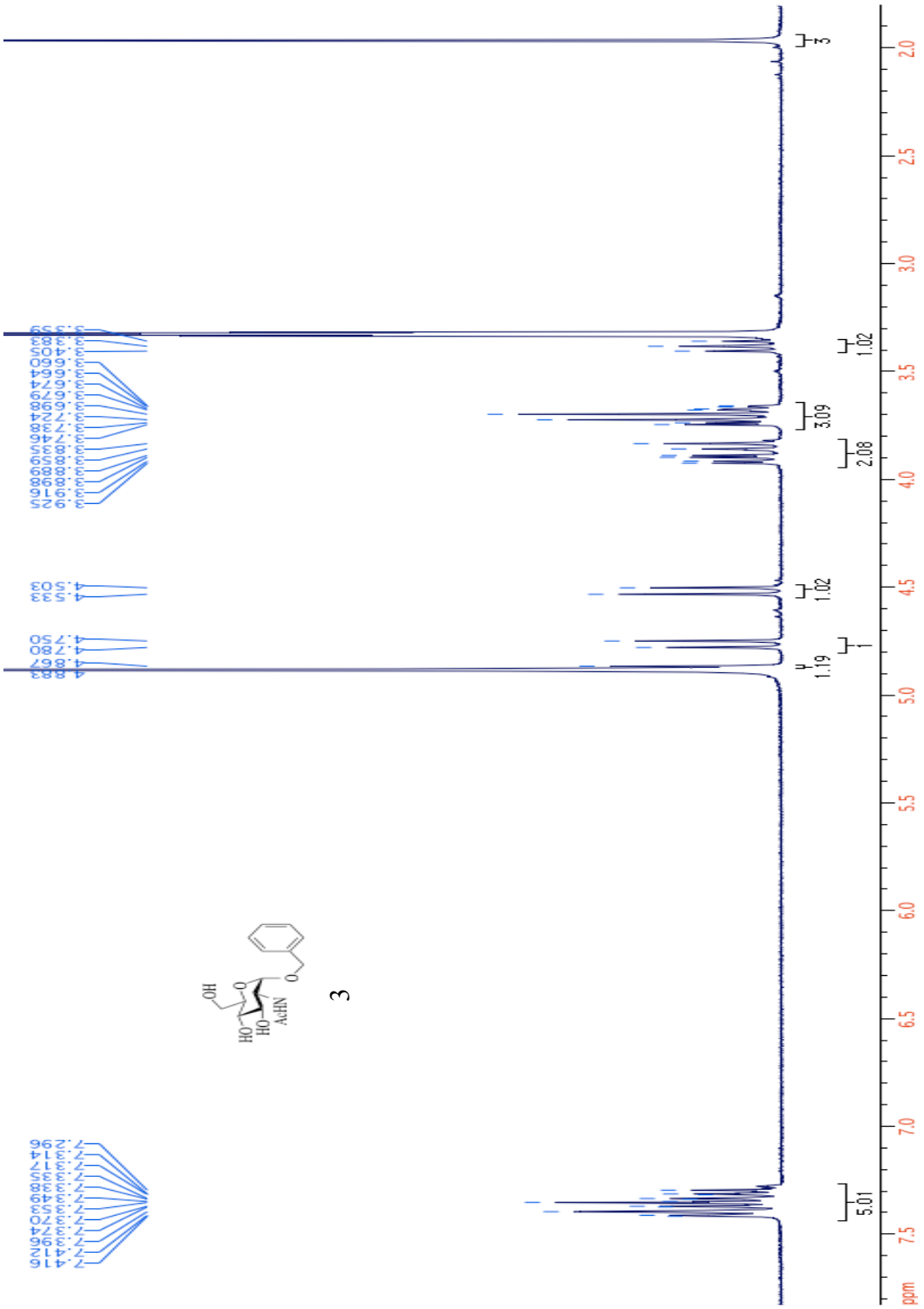
O-benzyl-6-diphenoxyphosphoryl-2-*i*-valeroylamino-2-deoxy-glucopyranose



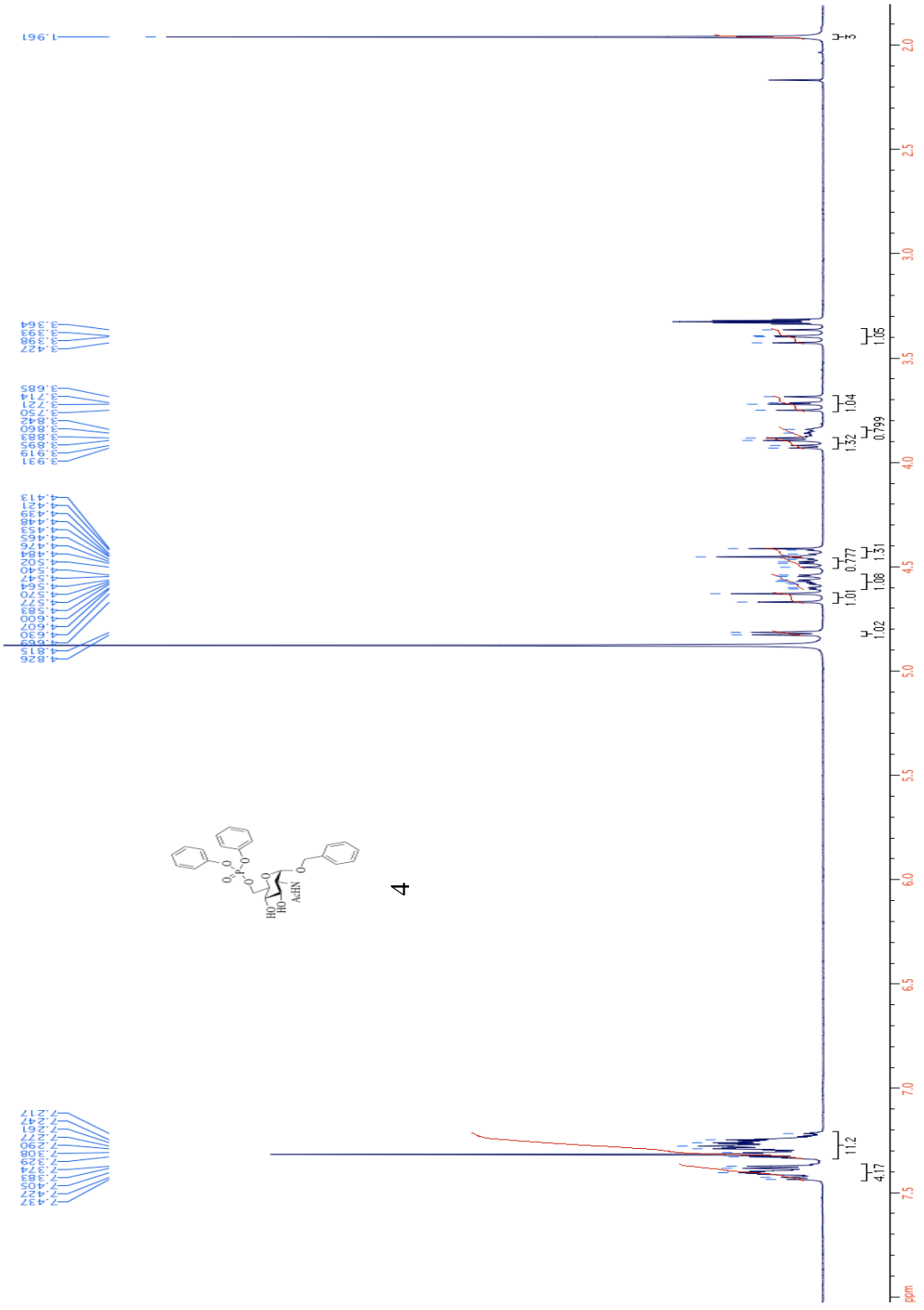
For this reaction, the phosphorylation reaction procedure was followed; 101 mg of 25, 84.5 mg (1.1 eq) of DPCP, 3.5 mg (0.1 eq) of DMAP and 2 mL of pyridine to yield 76 mg of 26 (46%).

^{1}H-NMR (300 MHz; MeOD): δ 7.42-7.36 (m, 4H), 7.29-7.22 (m, 10H), 4.81 (d, *J* = 3.6 Hz, 1H), 4.62 (d, *J* = 11.8 Hz, 1H), 4.55 (td, *J* = 5.5, 2.0 Hz, 1H), 4.49-4.42 (m, 1H), 4.39 (d, *J* = 11.8 Hz, 1H), 3.91 (dd, *J* = 10.8, 3.6 Hz, 1H), 3.85 (dd, *J* = 10.0, 5.5 Hz, 1H), 3.70 (dd, *J* = 10.8, 8.7 Hz, 1H), 3.38 (dd, *J* = 10.2, 8.9 Hz, 1H), 2.11-2.07 (m, 2H), 2.04-1.97 (m, 1H), 0.92 (t, *J* = 6.1 Hz, 5H).

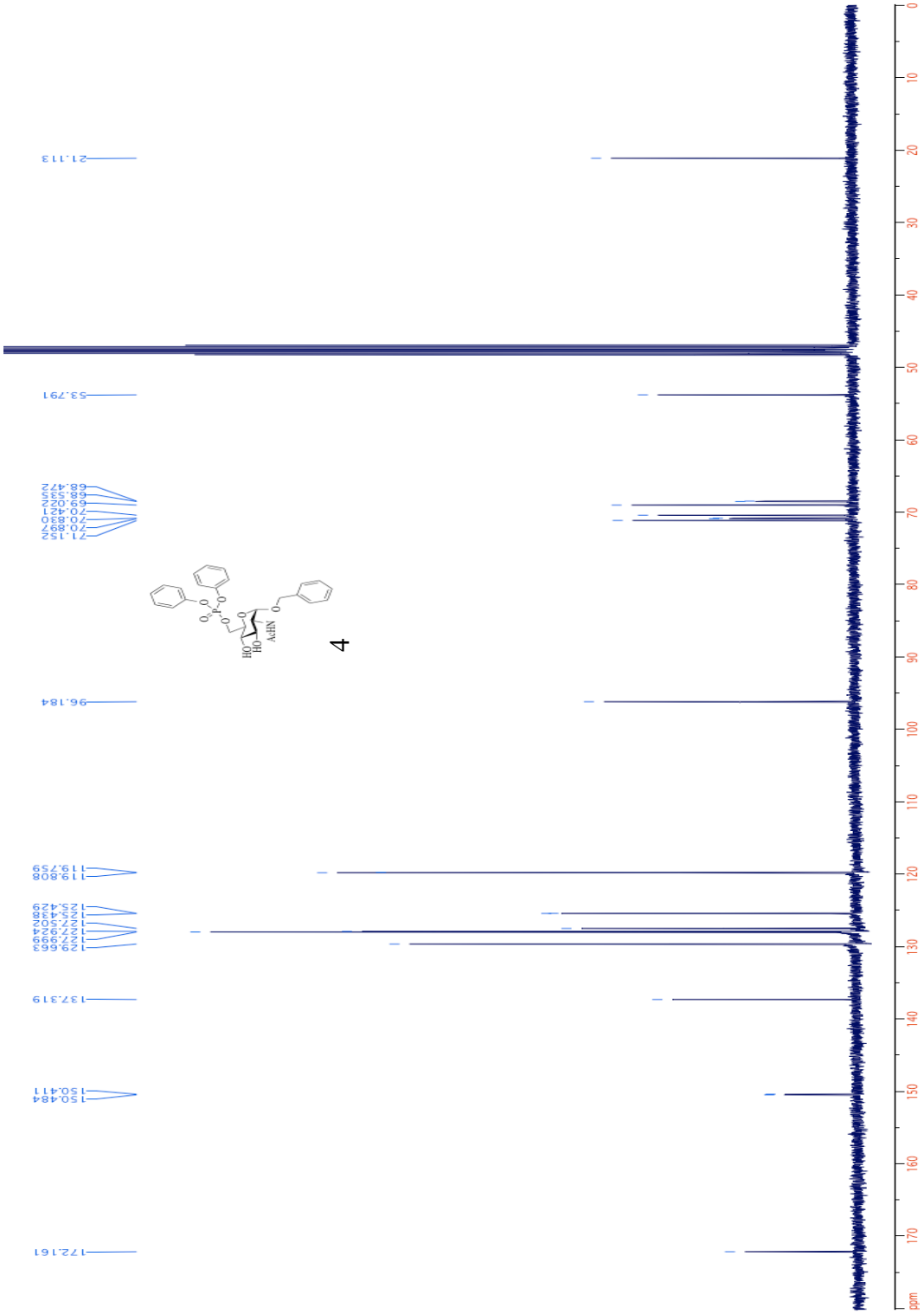
3.3.5 Spectroscopic Data



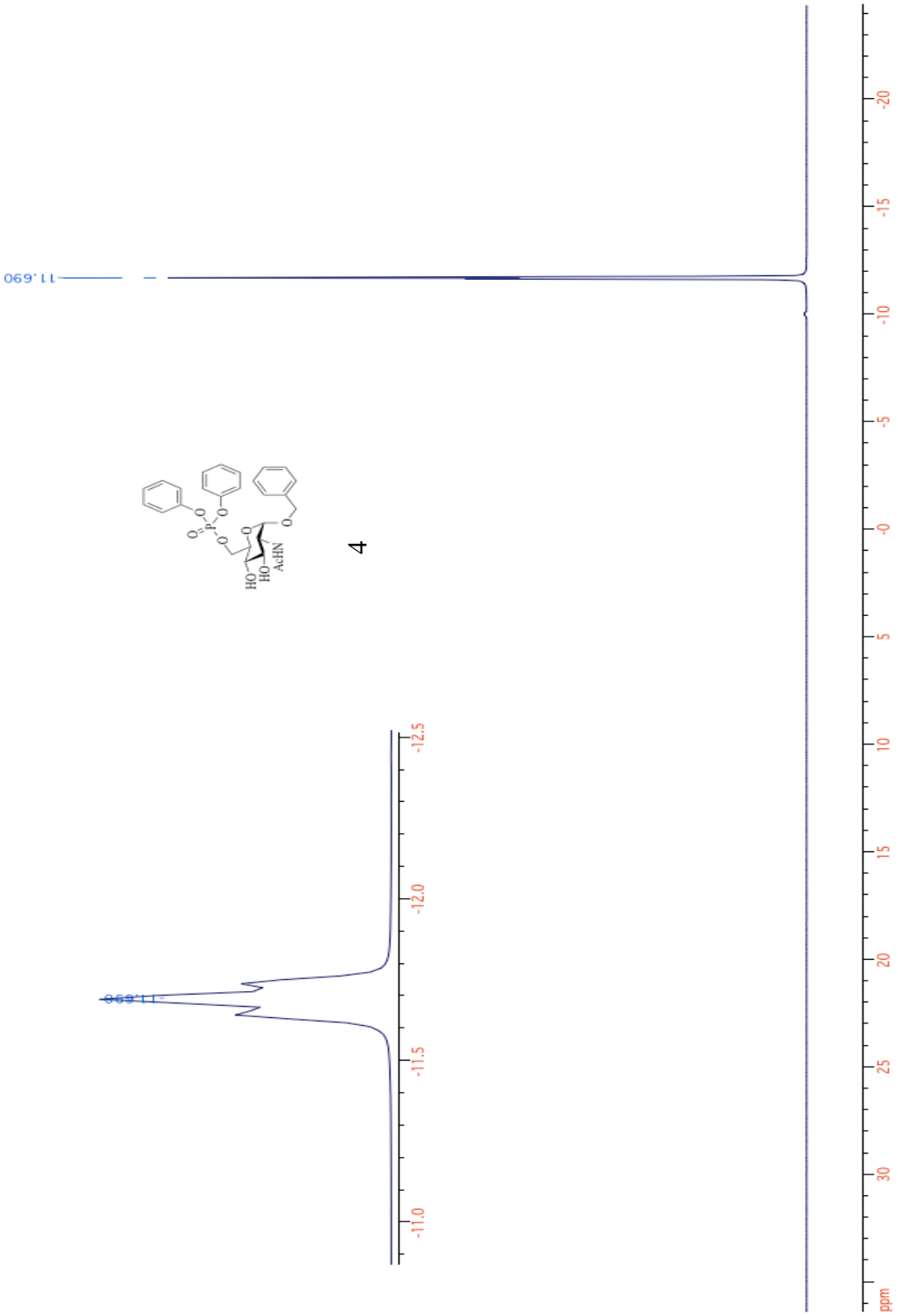
SD 1. ^{1}H-NMR of O-benzyl-2-acetylamino-2-deoxy-D-glucopyranose



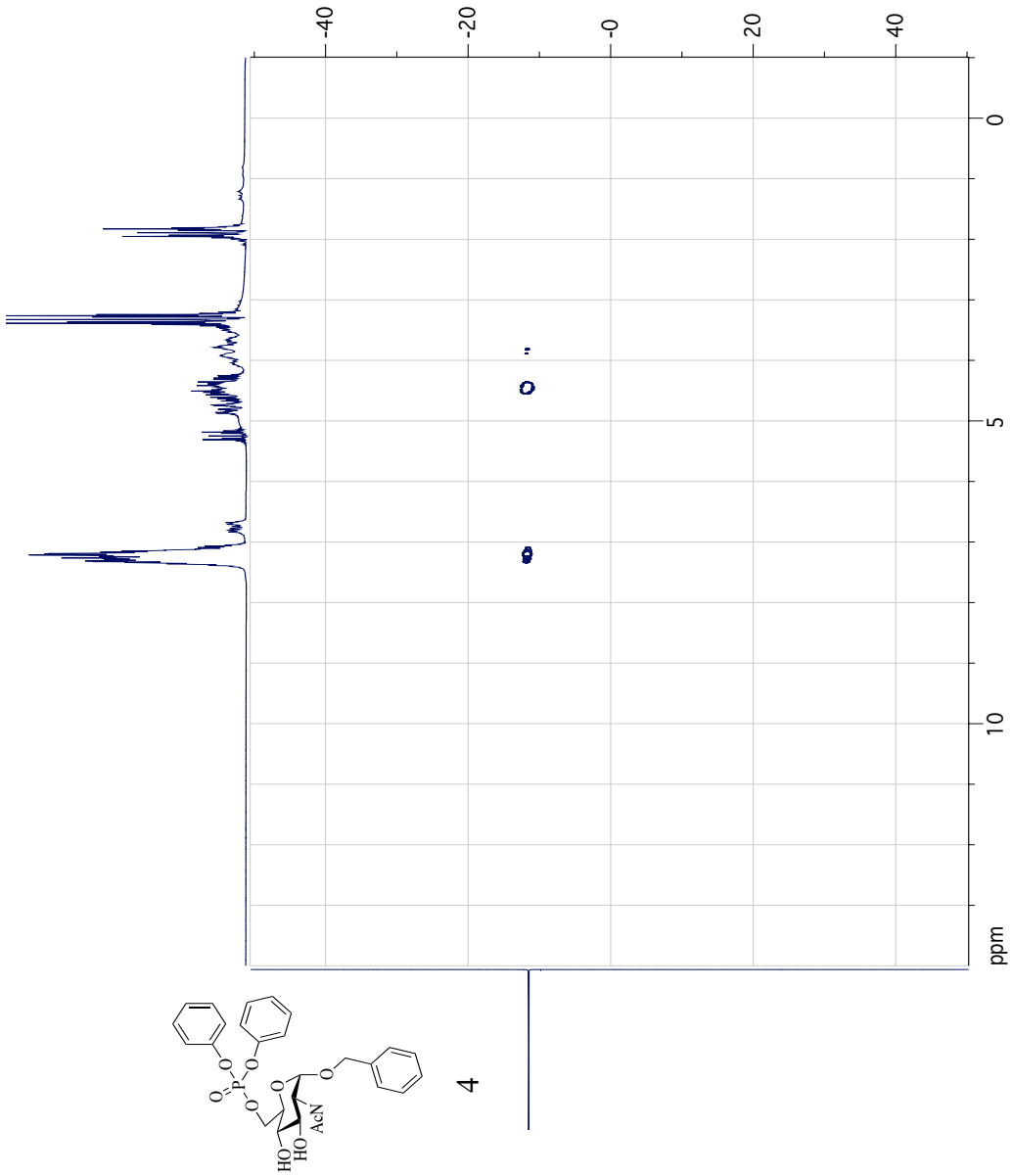
SD 2. ^{1}H-NMR of O-benzyl-6-diphenoxyphosphoryl-2-acetylamino-2-deoxy-D-glucopyranose



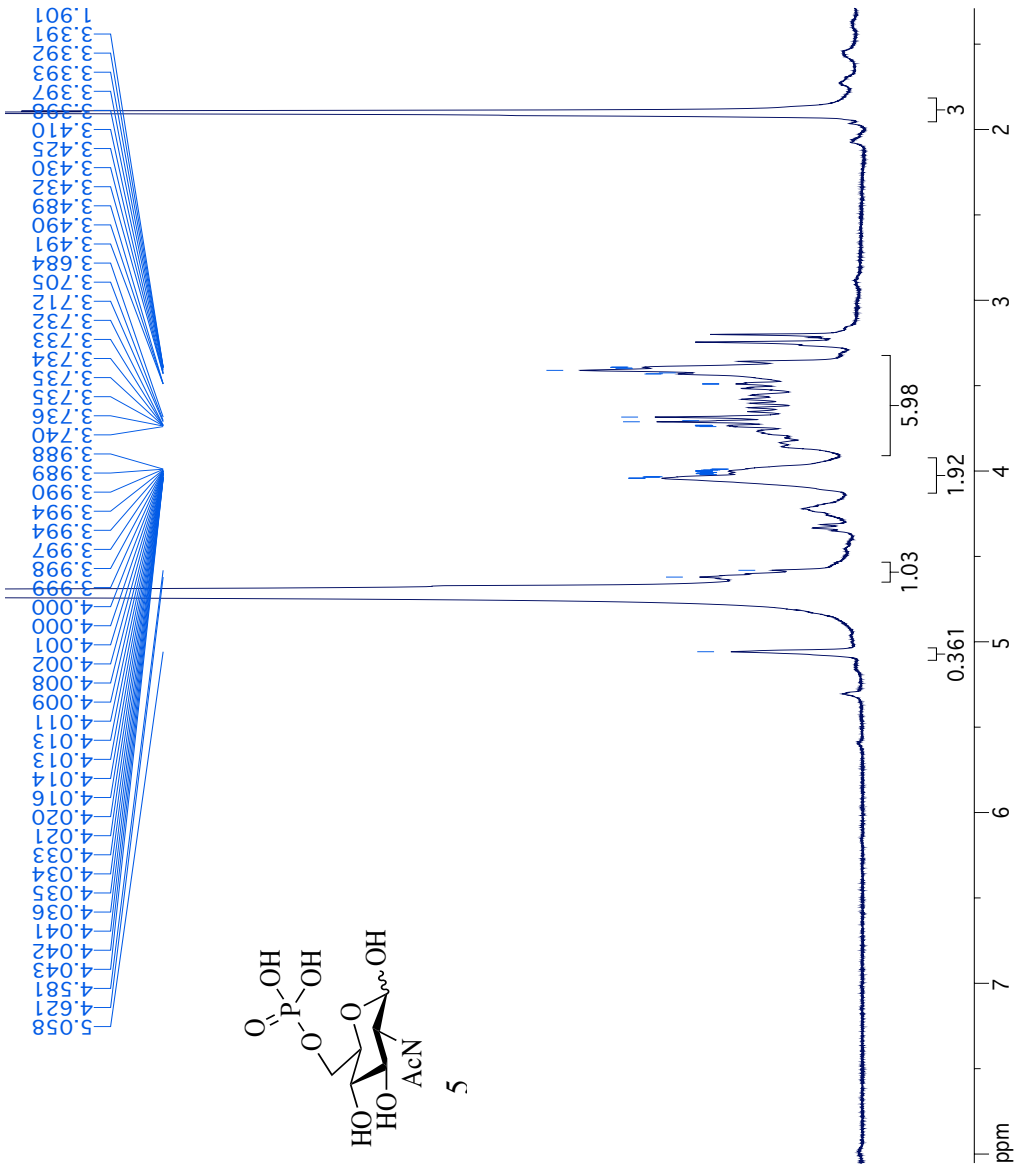
SD 3. ^{13}C-NMR of O-benzyl-6-diphenoxyphosphoryl-2-acetylamino-2-deoxy-D-glucopyranose



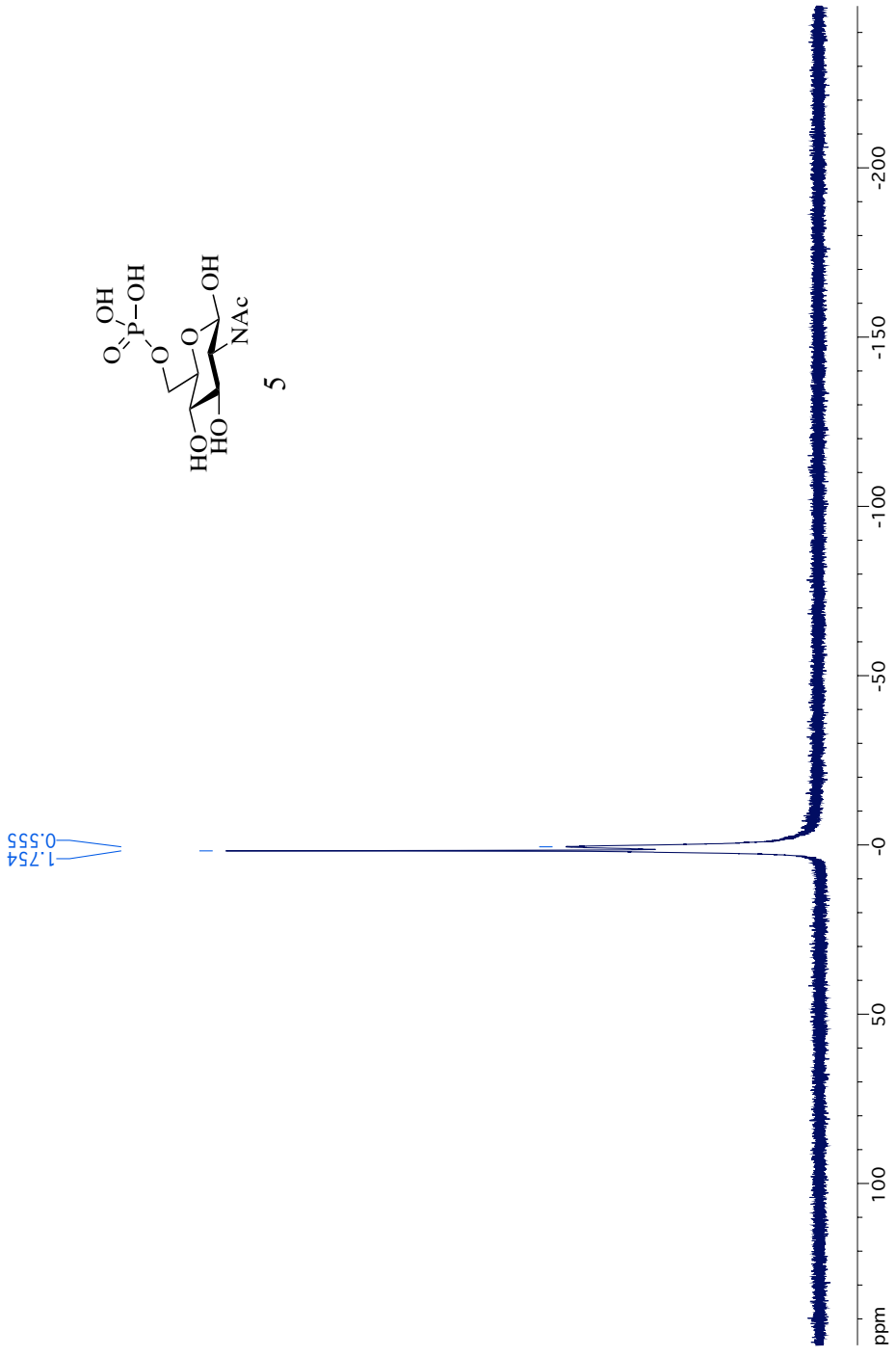
SD 4. ^{31}P-NMR of O-benzyl-6-diphenoxyphosphoryl-2-acetylamino-2-deoxy-D-glucopyranose



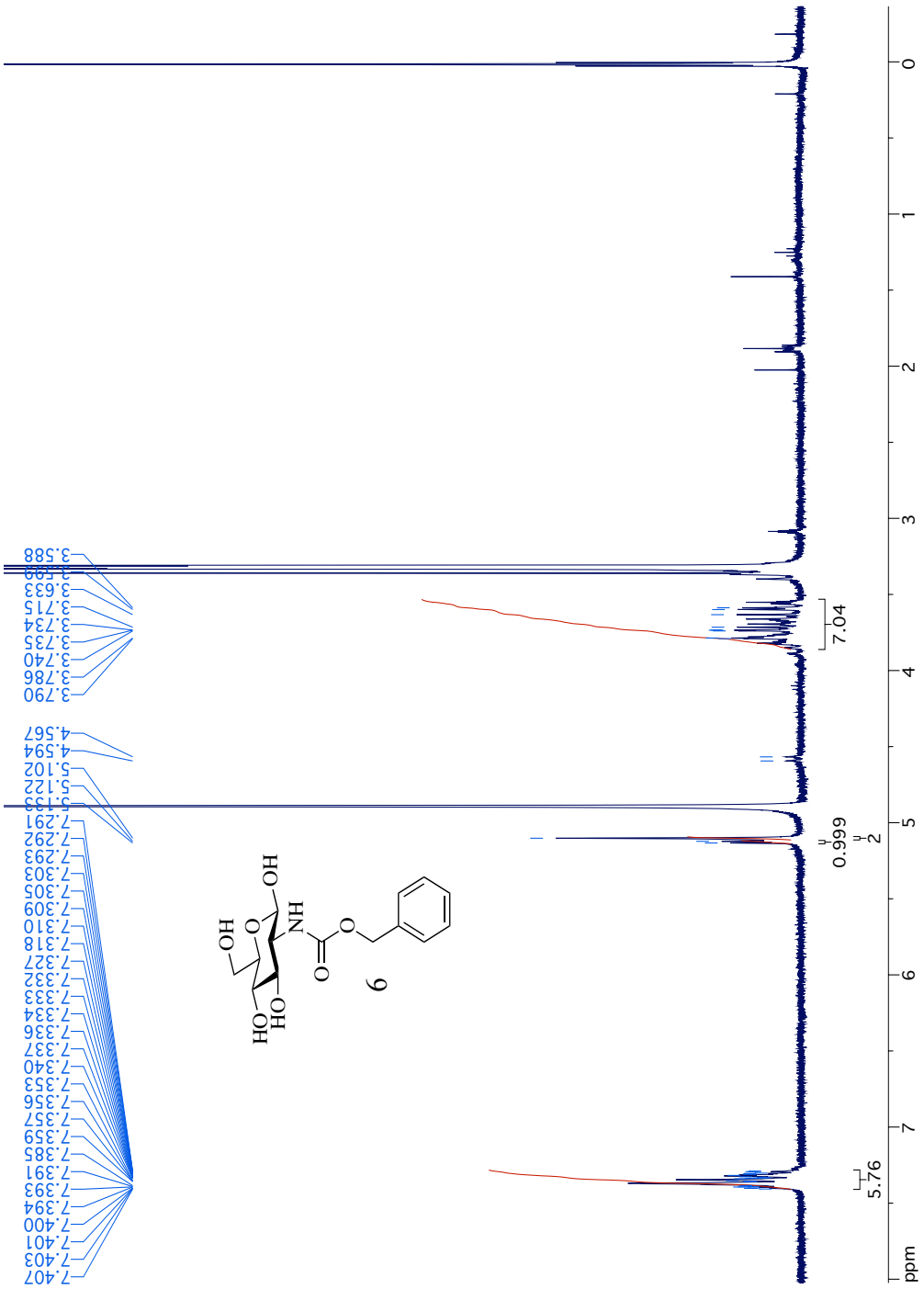
SD 5. ^{1}H-^{31}P HMBC of O-benzyl-6-diphenoxyphosphoryl-2-acetylamino-2-deoxy-D-glucopyranose



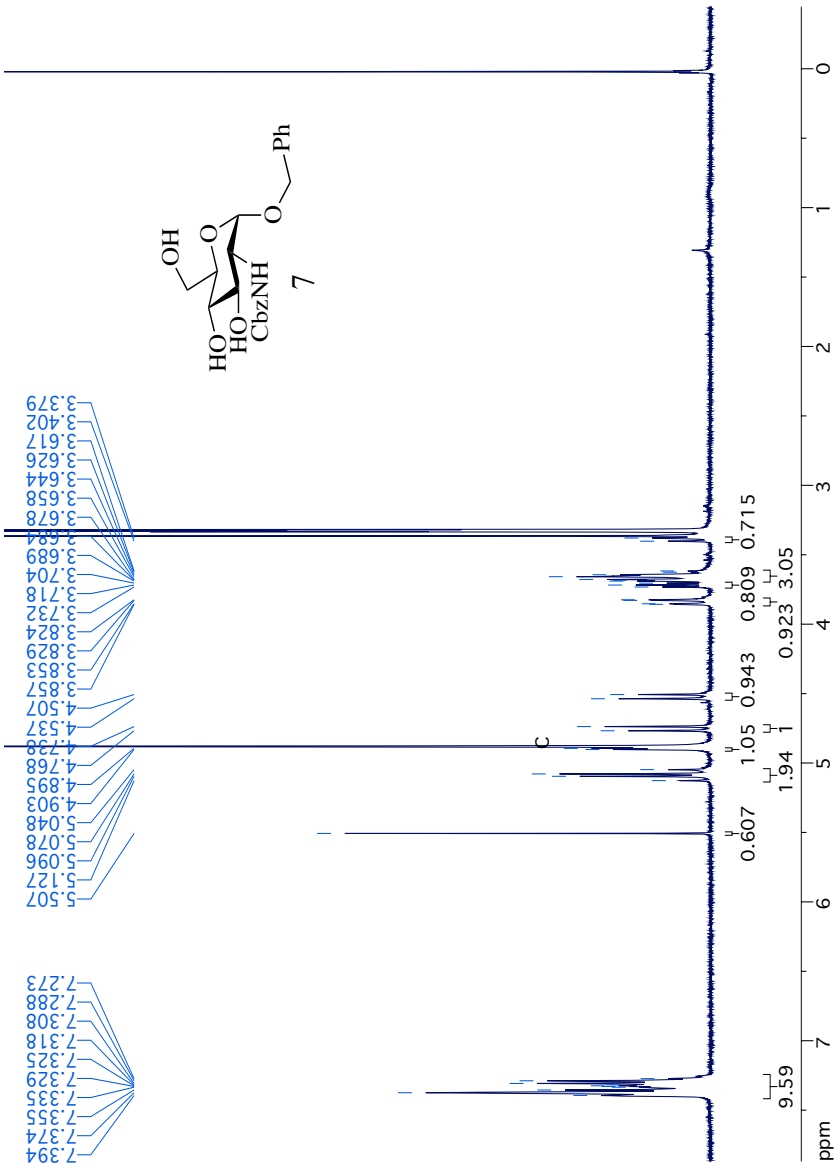
SD 6. ^{1}H-NMR of N-acetylglucosamine-6-phosphate



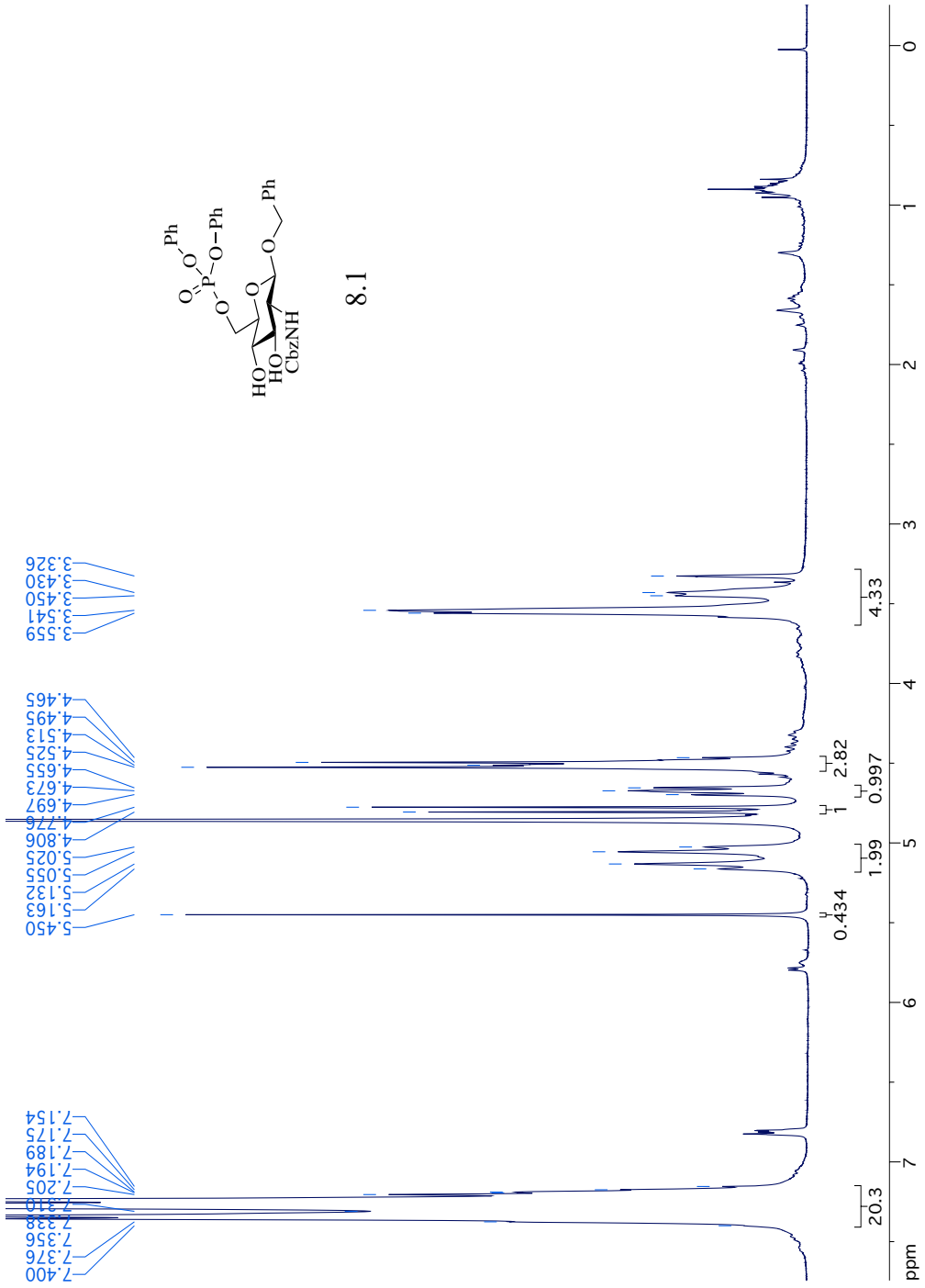
SD 7. ^{31}P-NMR of N-acetylglucosamine-6-phosphate



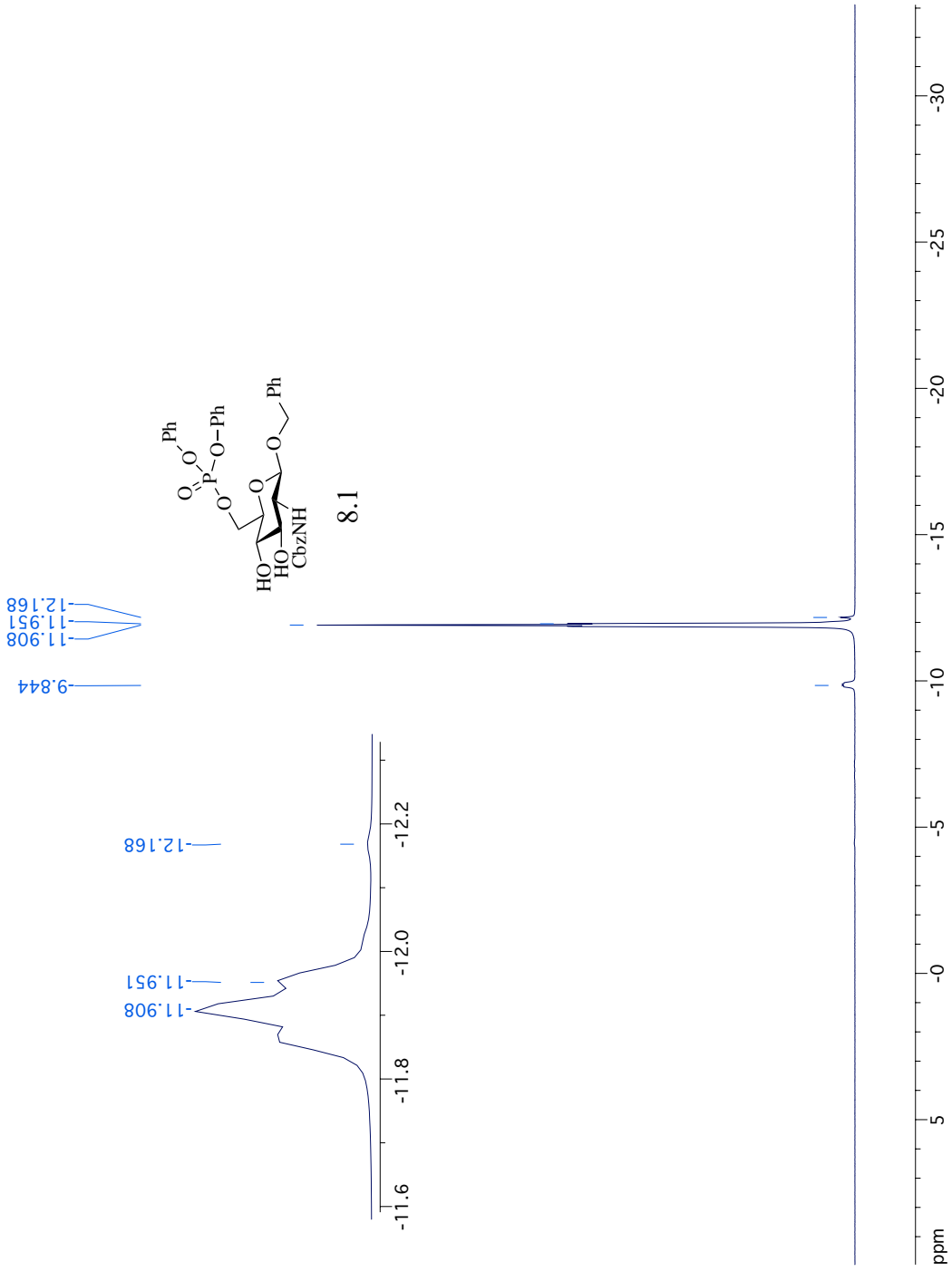
SD 8. ^{1}H-NMR of 2-carbobenzyloxyamino-2-deoxy-D-glucopyranose



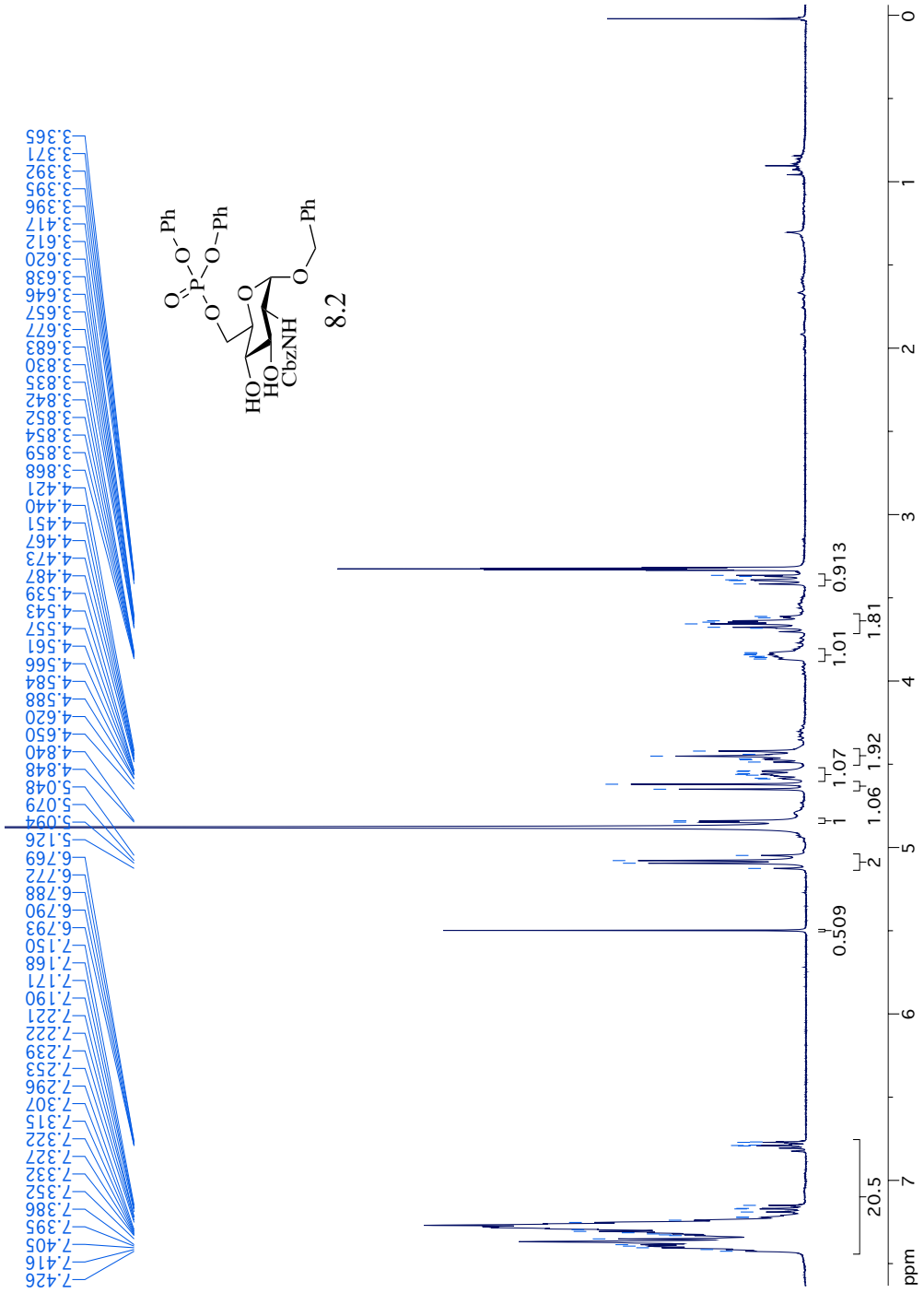
SD 9. ^{1}H-NMR of O-benzyl-2-carbobenzyloxyamino-2-deoxy-D-glucopyranose



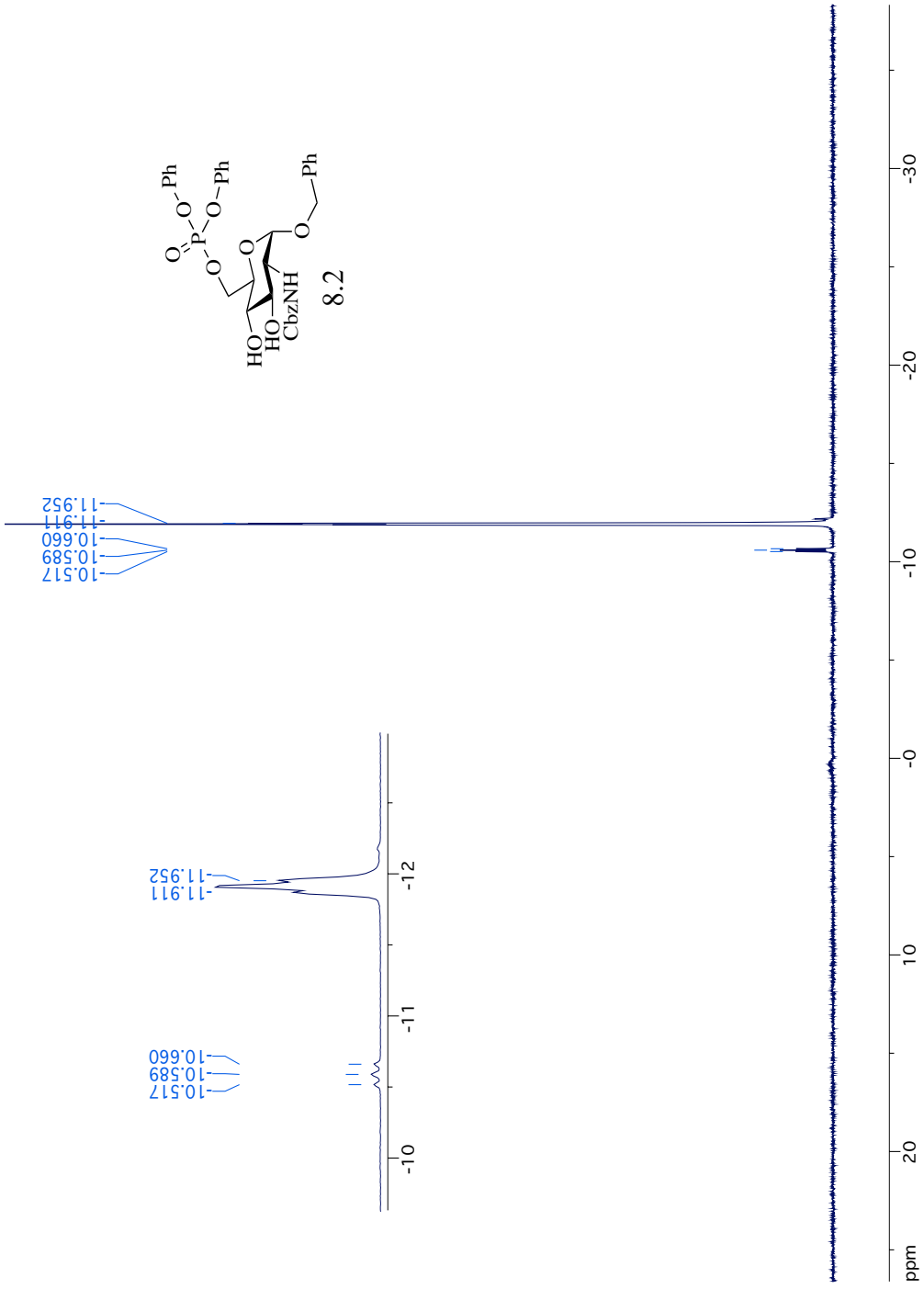
SD 10. ^{1}H-NMR of β-O-benzyl-6-diphenoxyphosphoryl-2-carbobenzyloxyamino-2-deoxy-D-glucopyranose



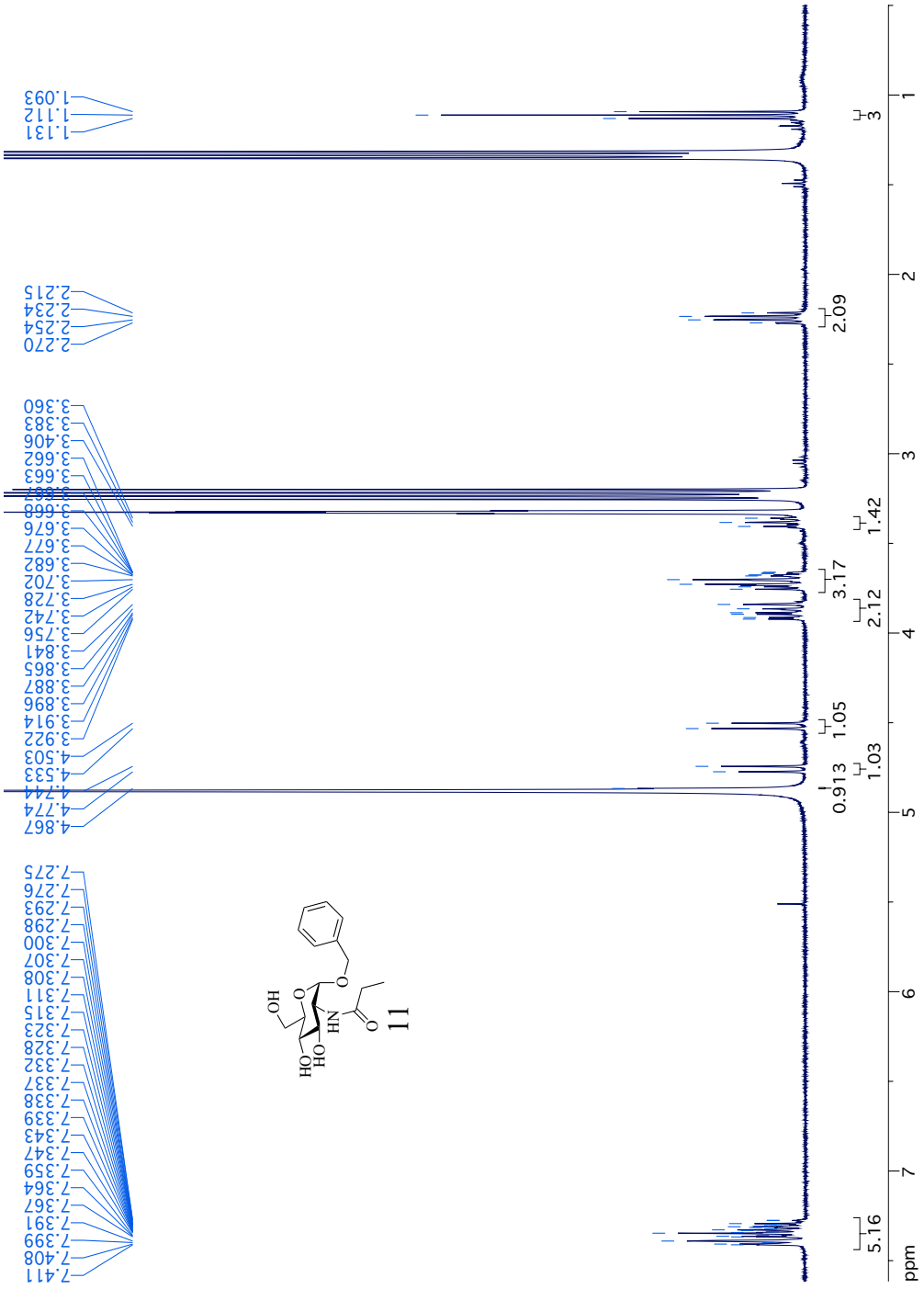
SD 11. ^{31}P-NMR of β-O-benzyl-6-diphenoxyphosphoryl-2-carbobenzyloxyamino-2-deoxy-D-glucopyranose



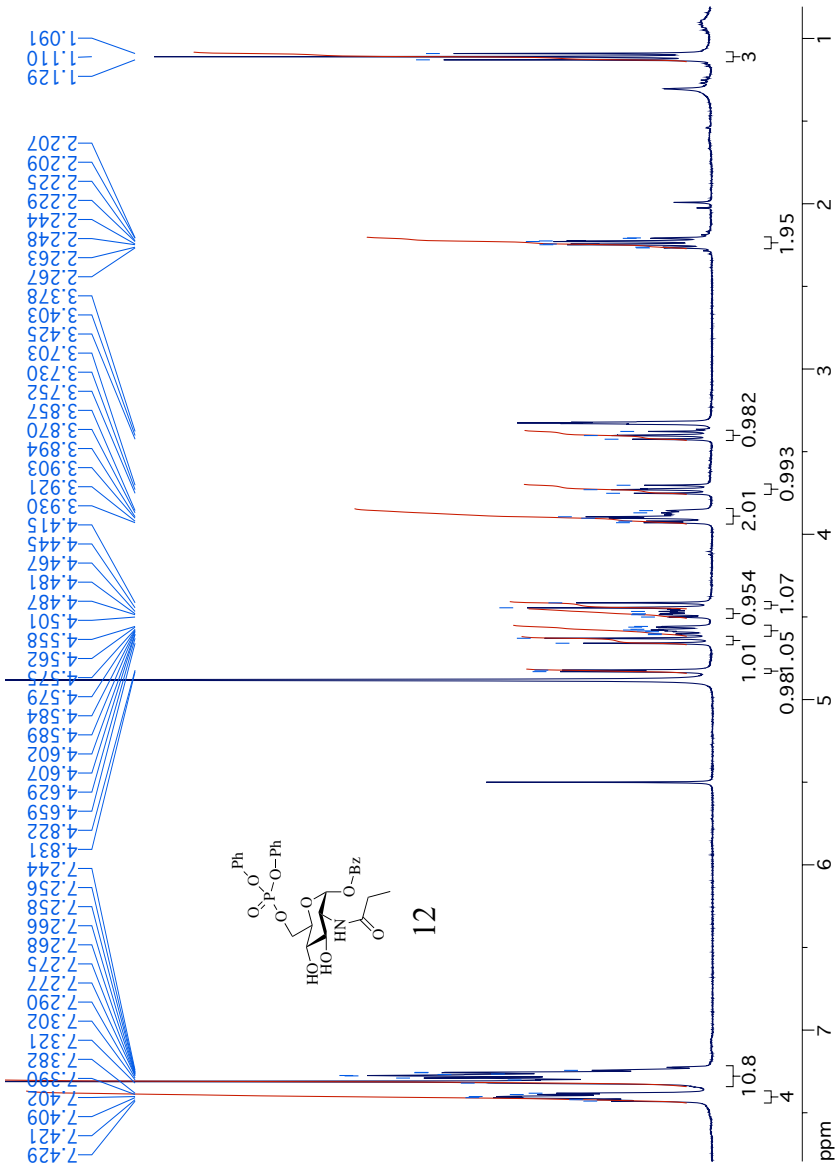
SD 12. ^{1}H-NMR of α-O-benzyl-6-diphenoxyphosphoryl-2-carbobenzyloxyamino-2-deoxy-D-glucopyranose



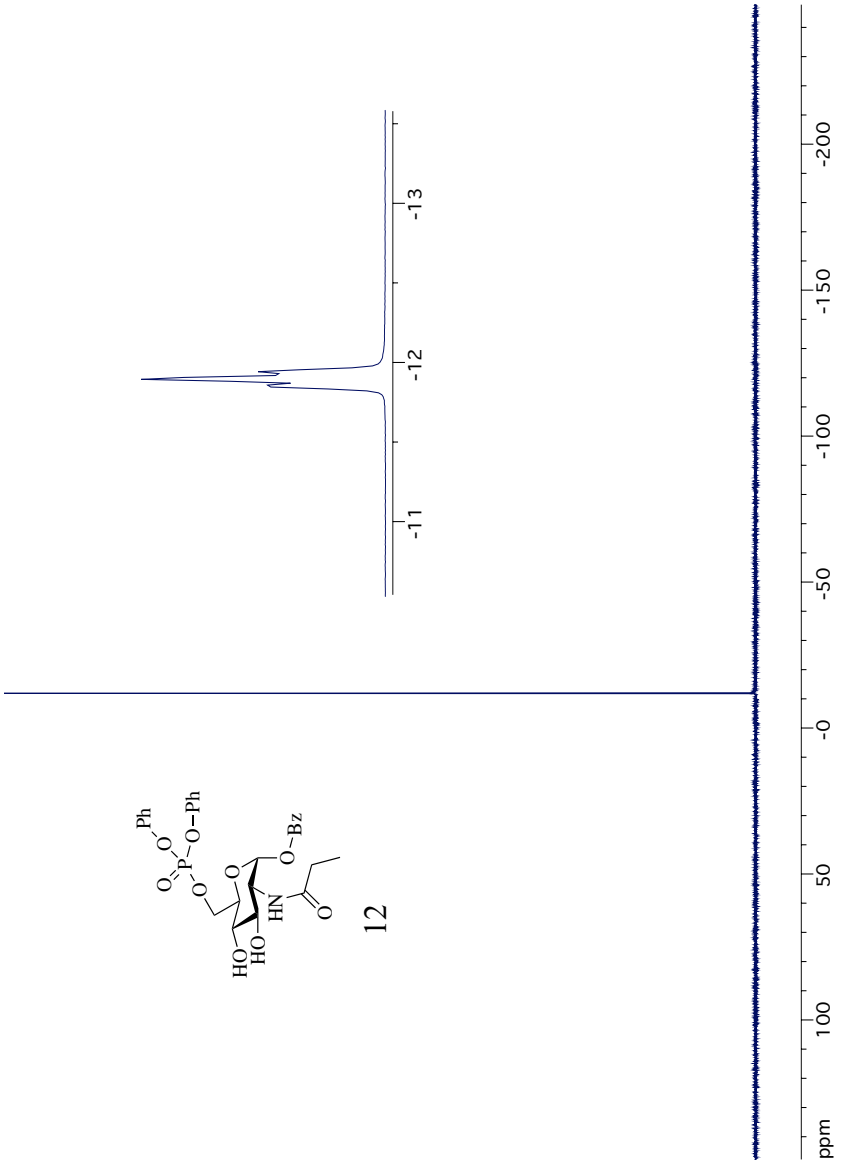
SD 13. ^{31}P-NMR of α-O-benzyl-6-diphenoxyphosphoryl-2-carbobenzyloxyamino-2-deoxy-D-glucopyranose



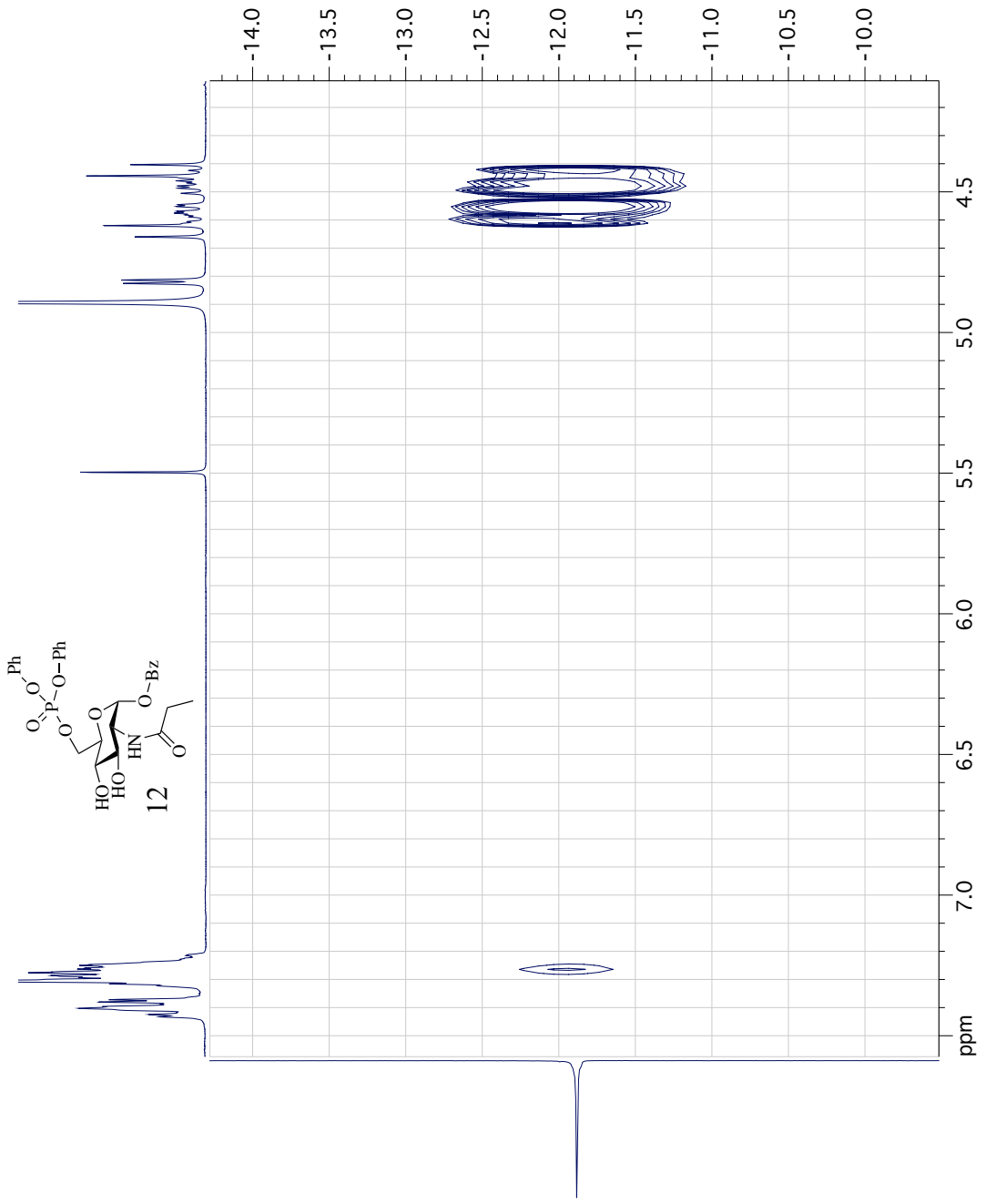
SD 14. ^{1}H-NMR of O-benzyl-N-propylglucosamine



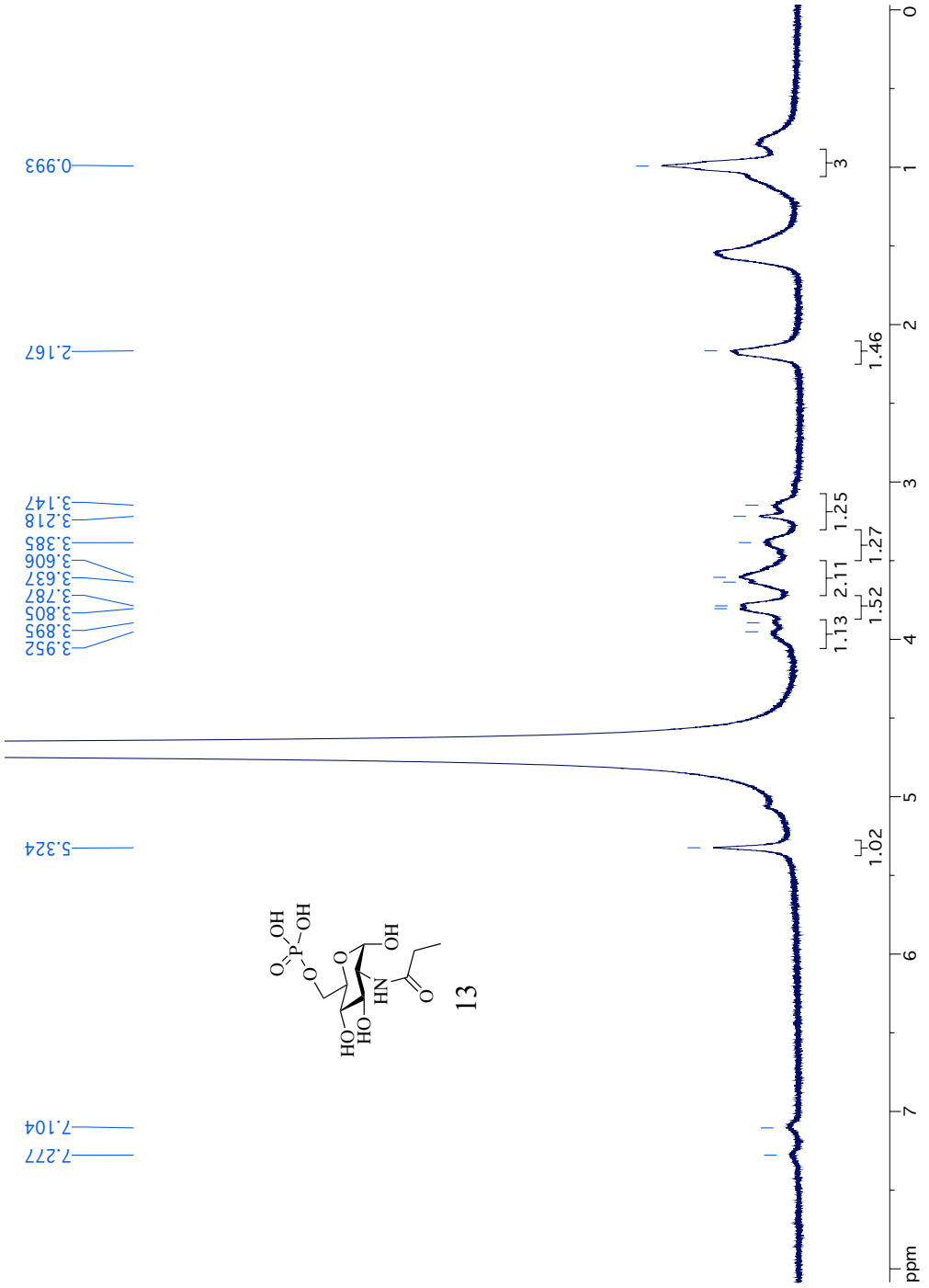
SD 15. ^{1}H-NMR of O-benzyl-6-diphenoxyphosphoryl-2-propylamino-2-deoxy-glucopyranose



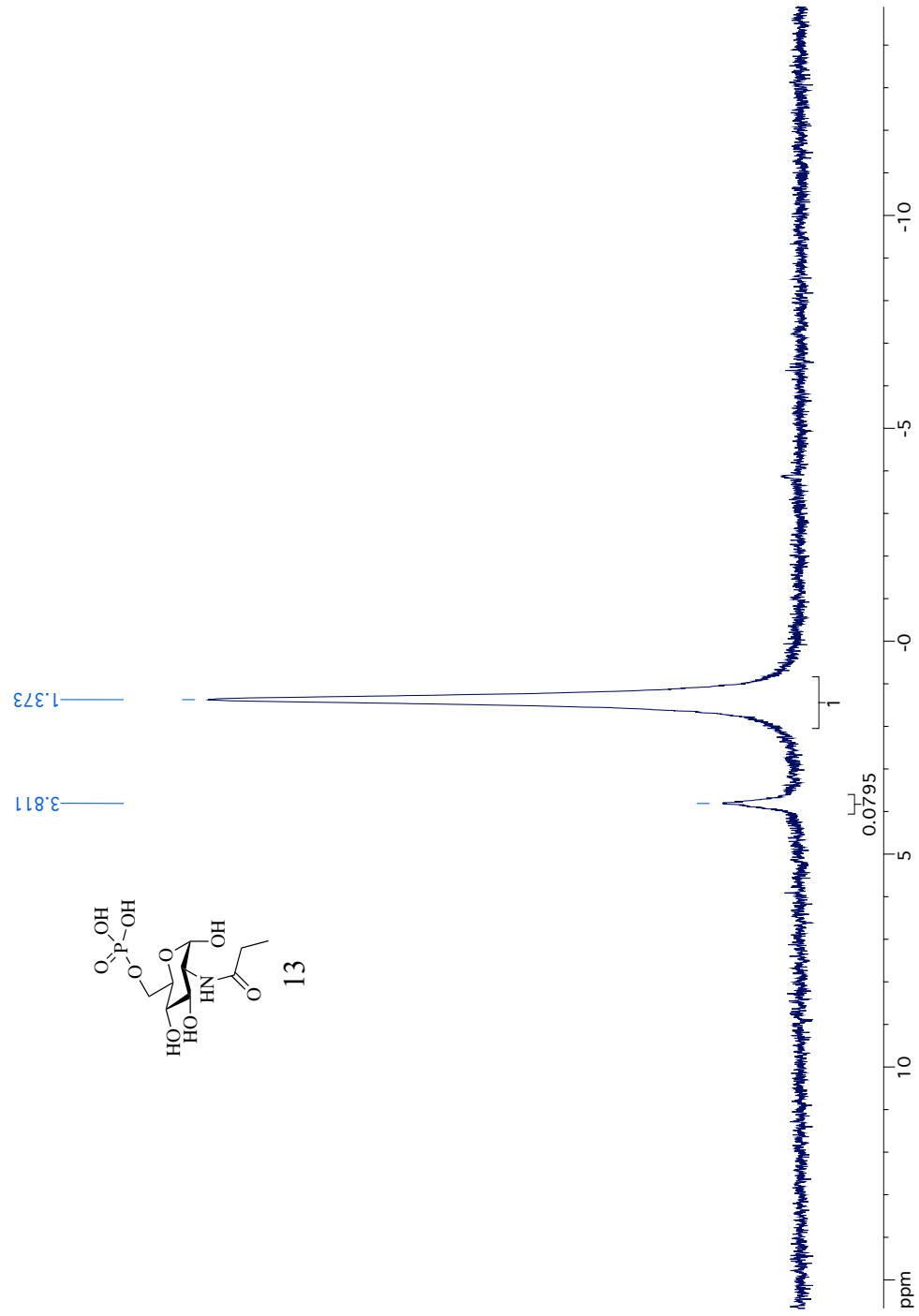
SD 16. ^{31}P-NMR of O-benzyl-6-diphenoxyphosphoryl-2-propylamino-2-deoxy-glucopyranose



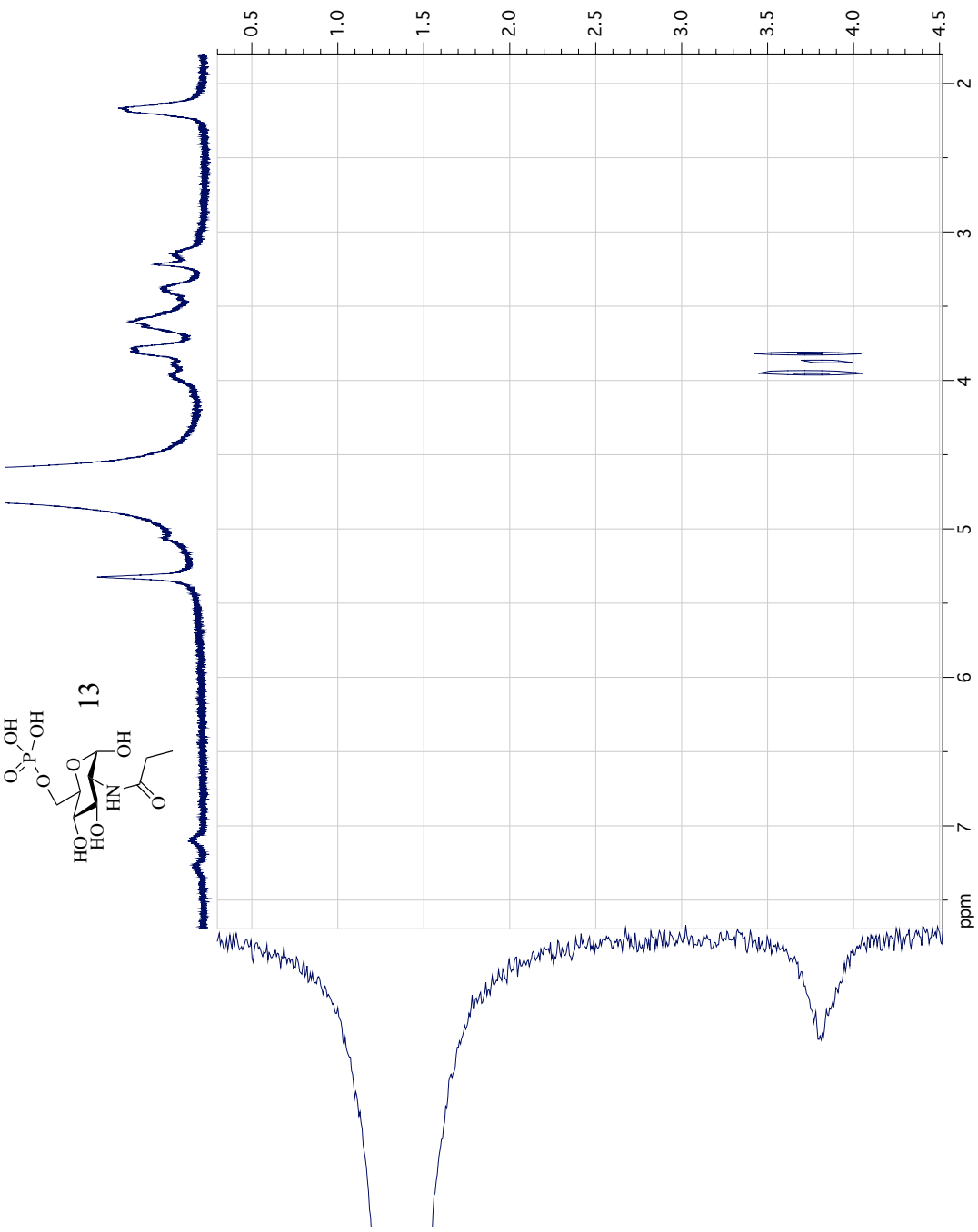
SD 17. ^{1}H-^{31}P HMBC of O-benzyl-6-diphenoxyphosphoryl-2-propylamino-2-deoxy-glucopyranose



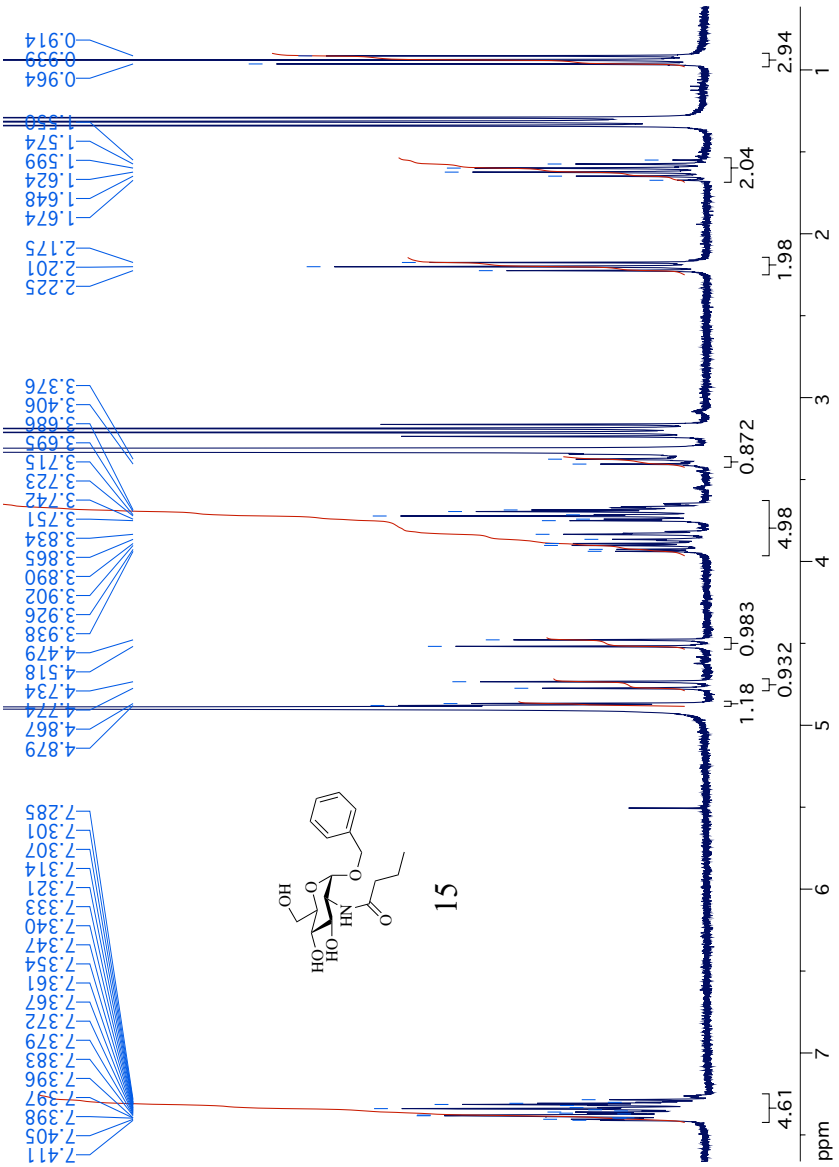
SD 18 . ^{1}H-NMR of N-propylglucosamine-6-phosphate



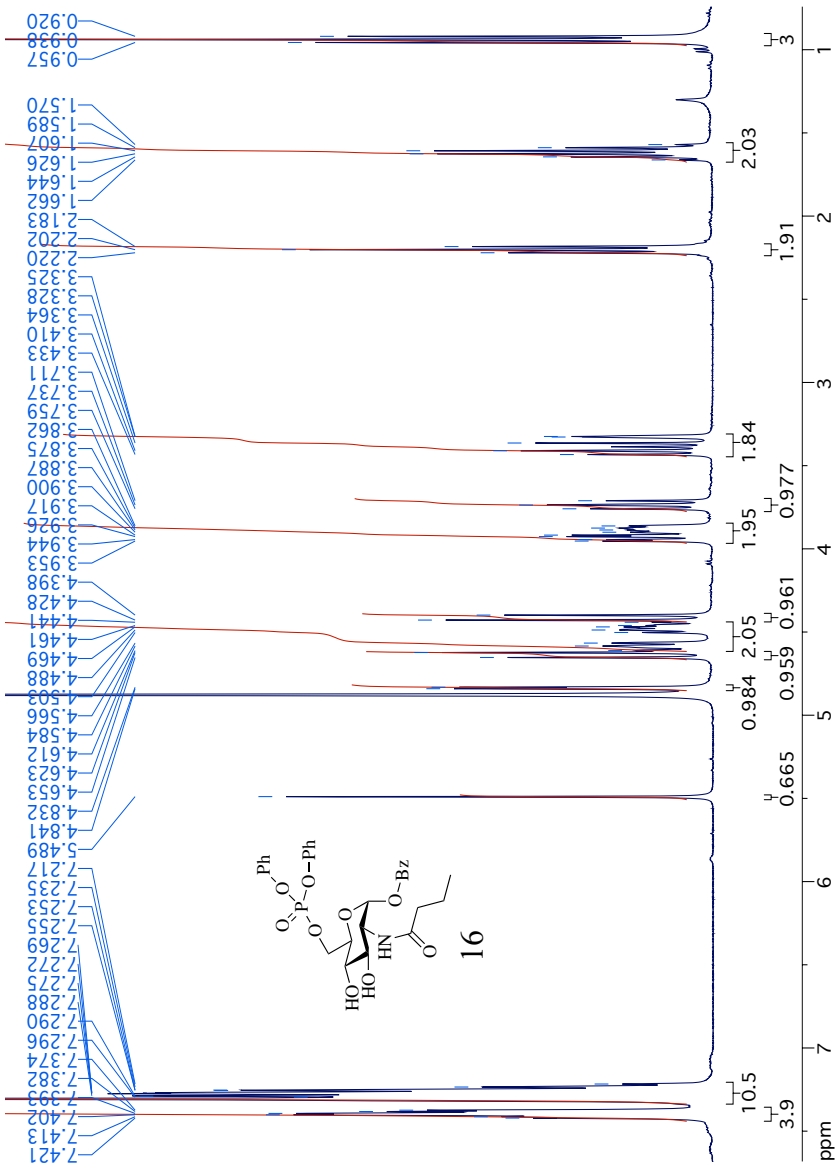
SD 19. ^{31}P-NMR of N-propylglucosamine-6-phosphate



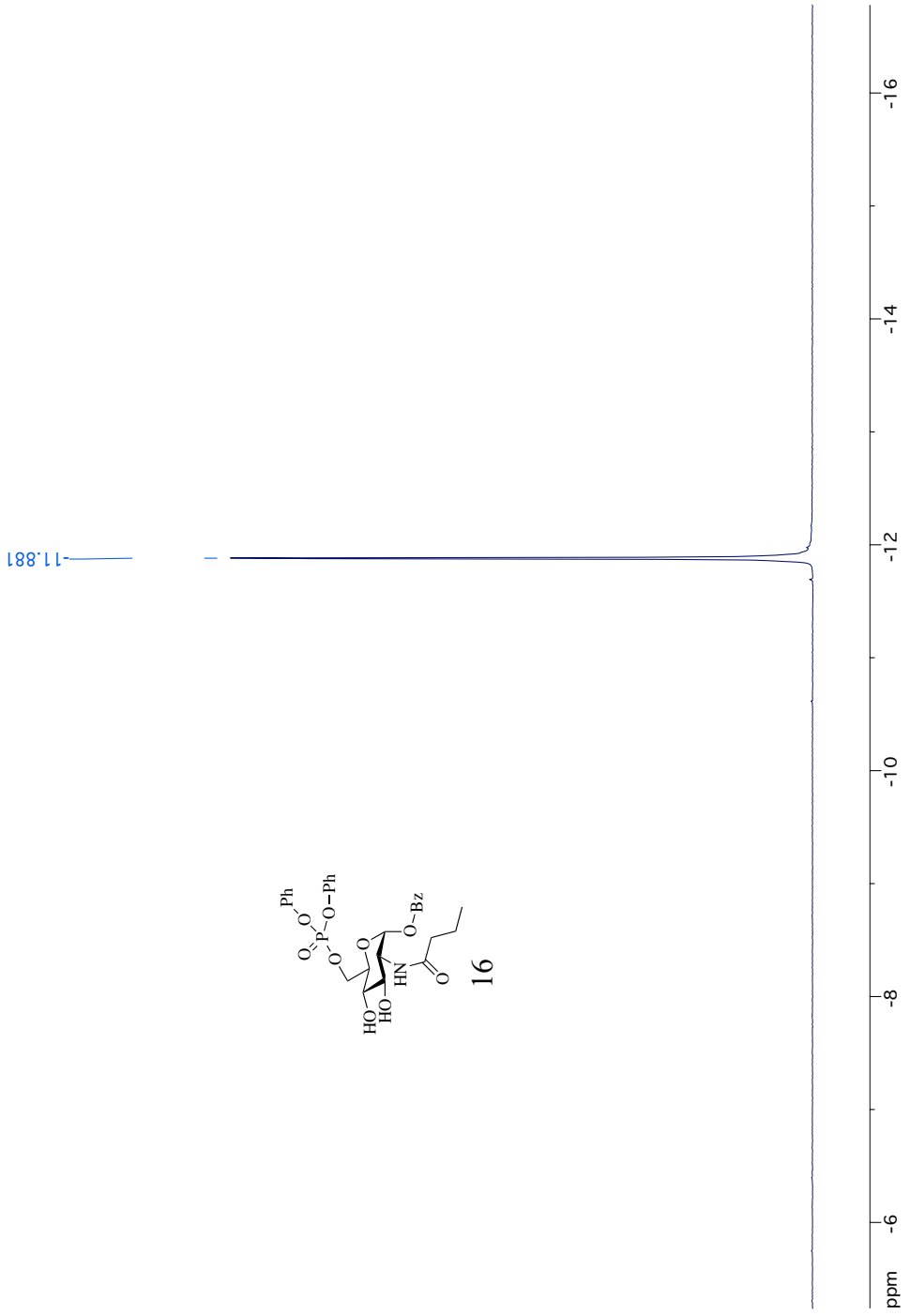
SD 20. ^{1}H-^{31}P HMBC of N-propylglucosamine-6-phosphate



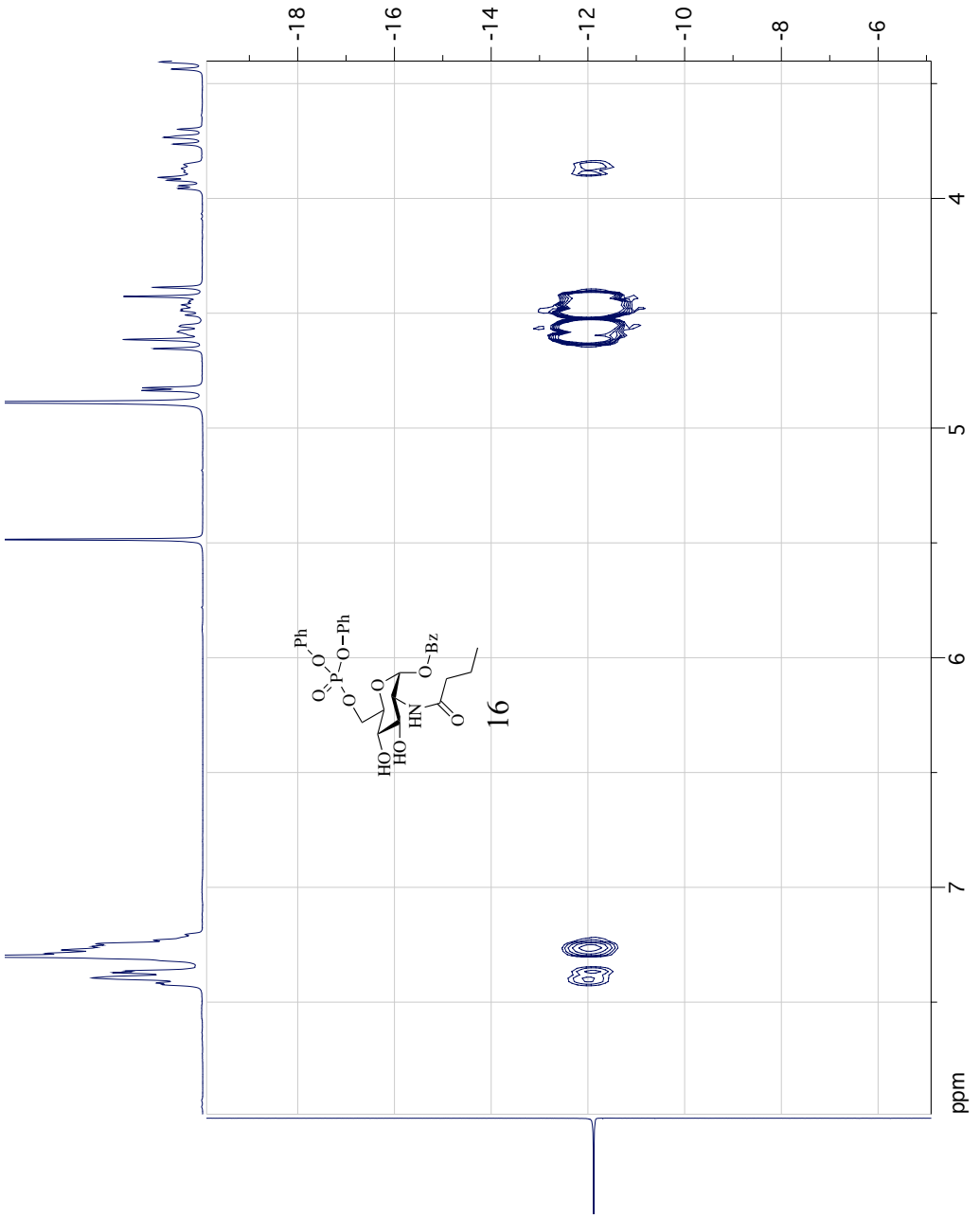
SD 21. ^{1}H-NMR of O-benzyl-2-butylamino-2-deoxy-glucopyranose



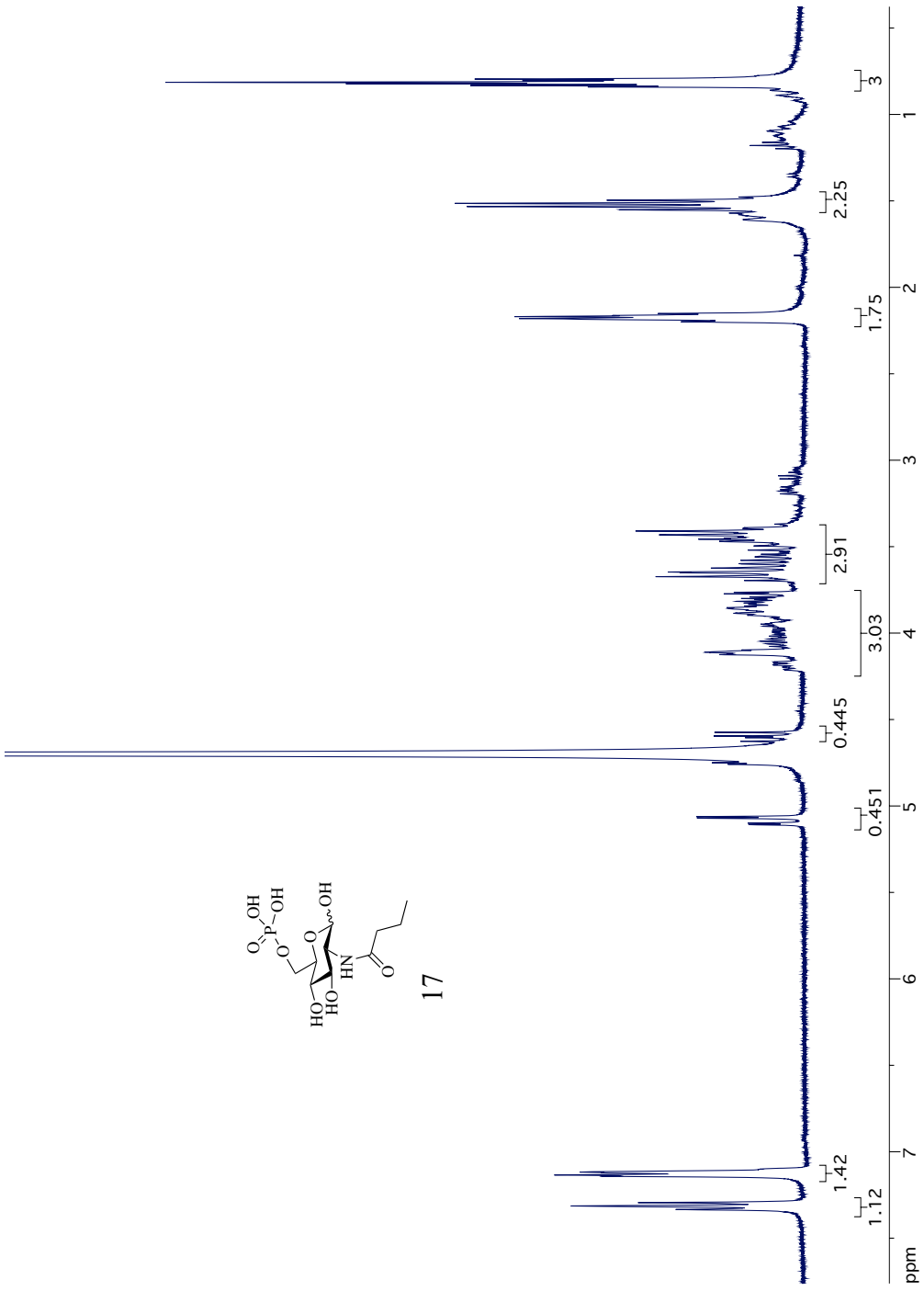
SD 22. ^{1}H-NMR of O-benzyl-6-diphenoxyphosphoryl-2-butylamino-2-deoxy-glucopyranose



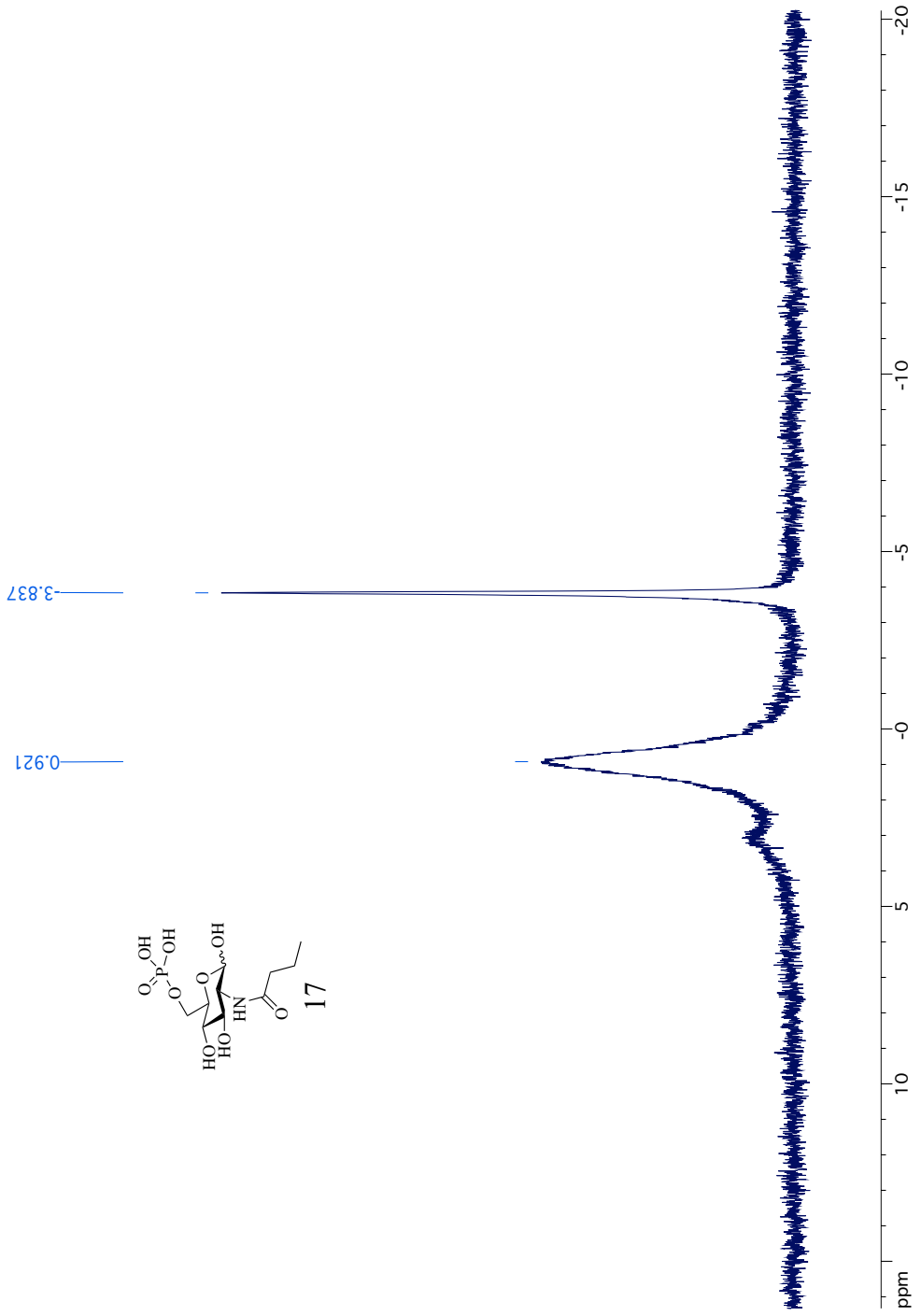
SD 23. ^{31}P-NMR of O-benzyl-6-diphenoxyphosphoryl-2-butylamino-2-deoxy-glucopyranose



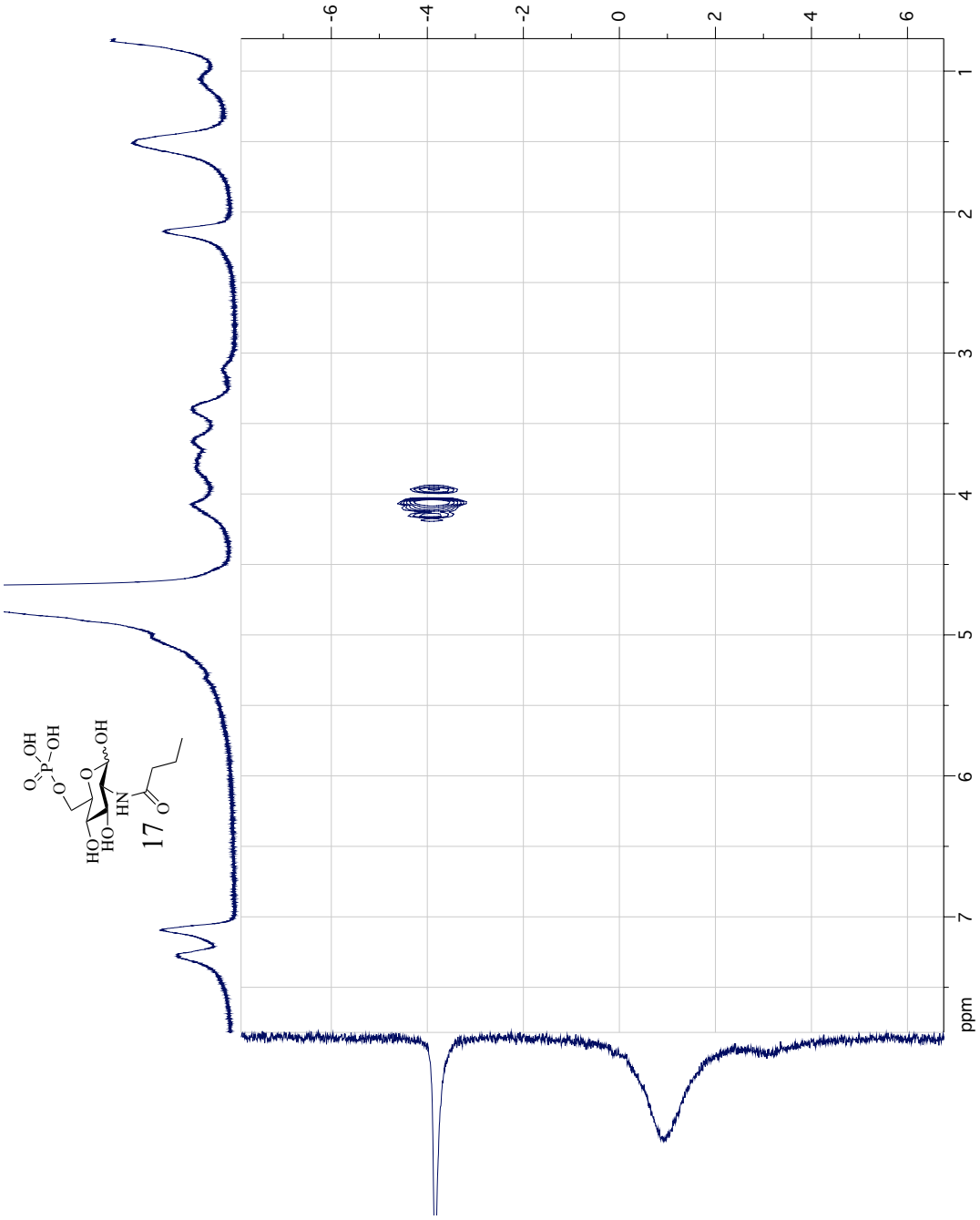
SD 24. ^{1}H-^{31}P HMBC of O-benzyl-6-diphenoxyphosphoryl-2-butylamino-2-deoxy-glucopyranose



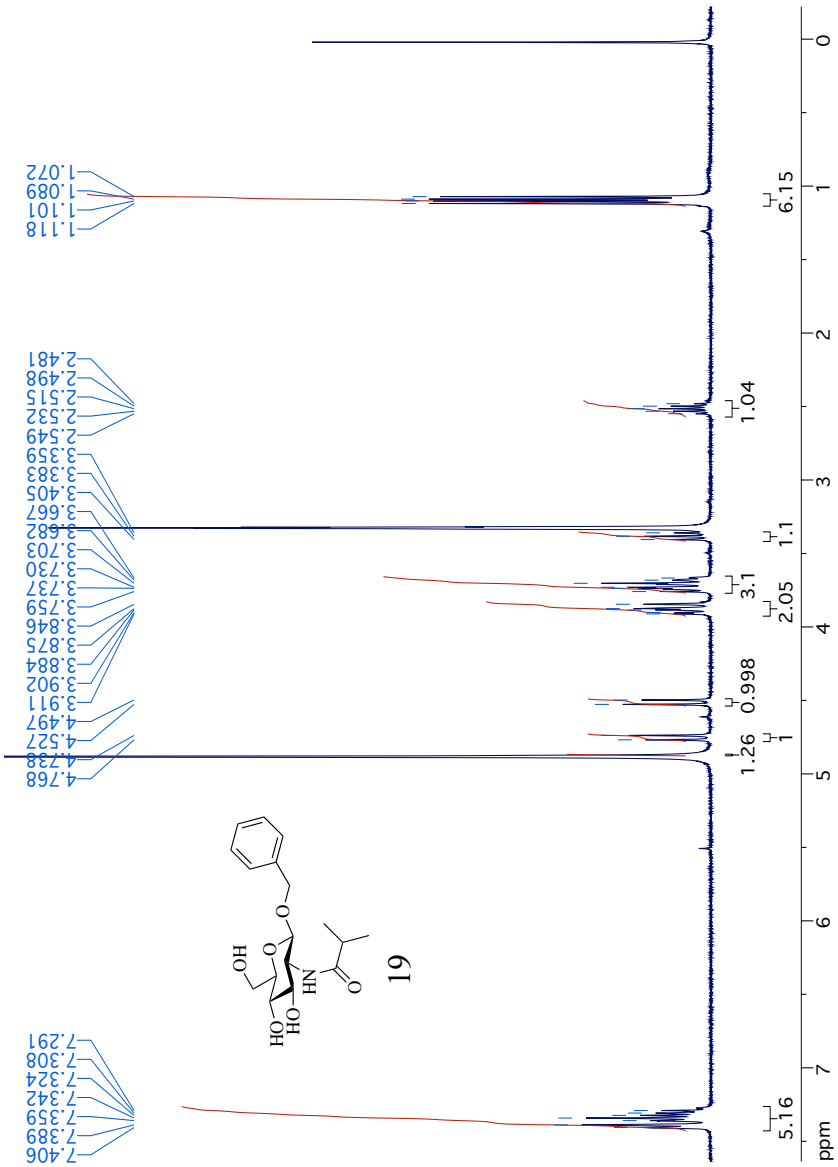
SD 25. ^{1}H-NMR of N-*n*-butylglucosamine-6-phosphate



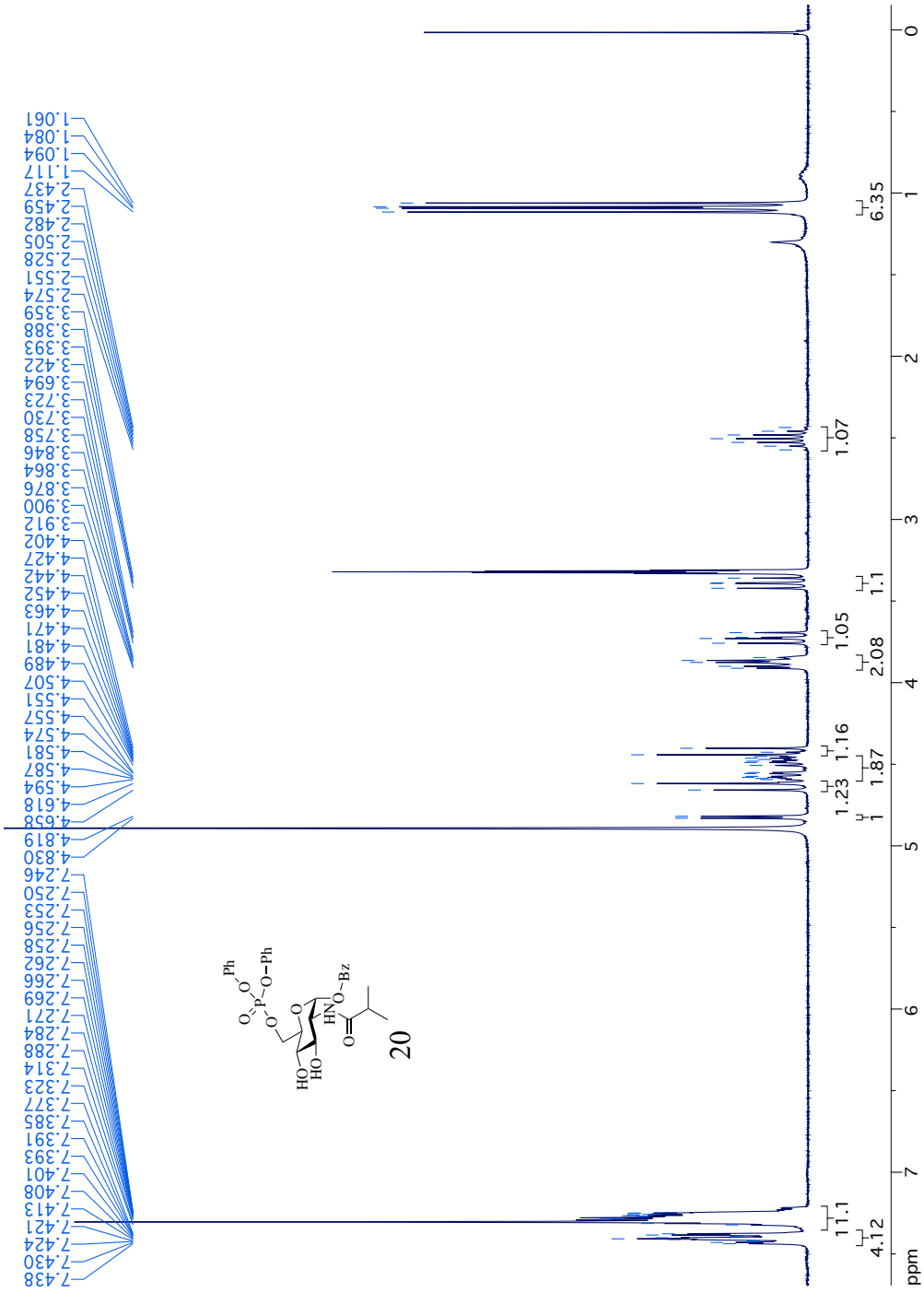
SD 26. ^{31}P-NMR of N-*n*-butylglucosamine-6-phosphate



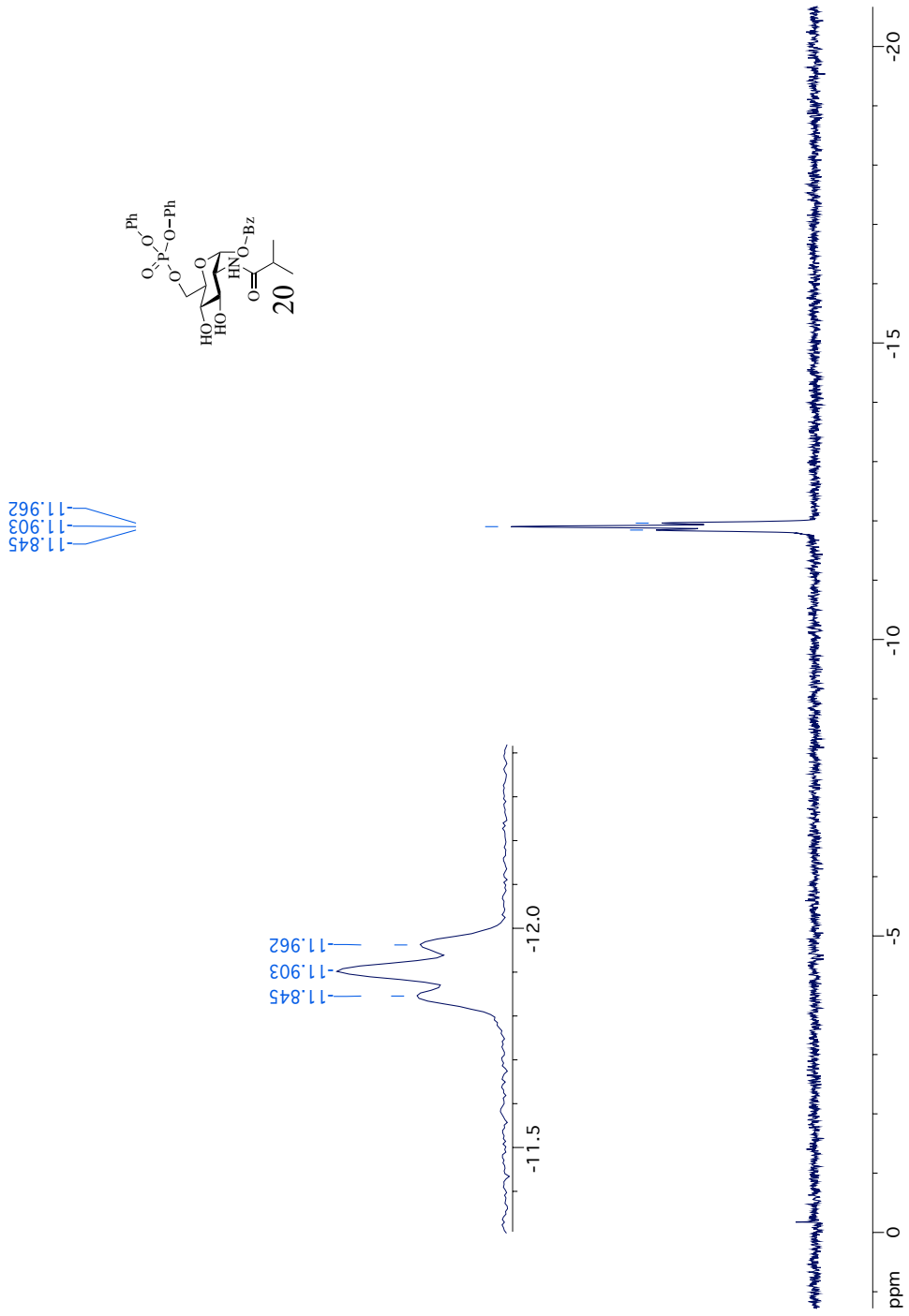
SD 27. ^{1}H-^{31}P HMBC of N-*n*-butylglucosamine-6-phosphate



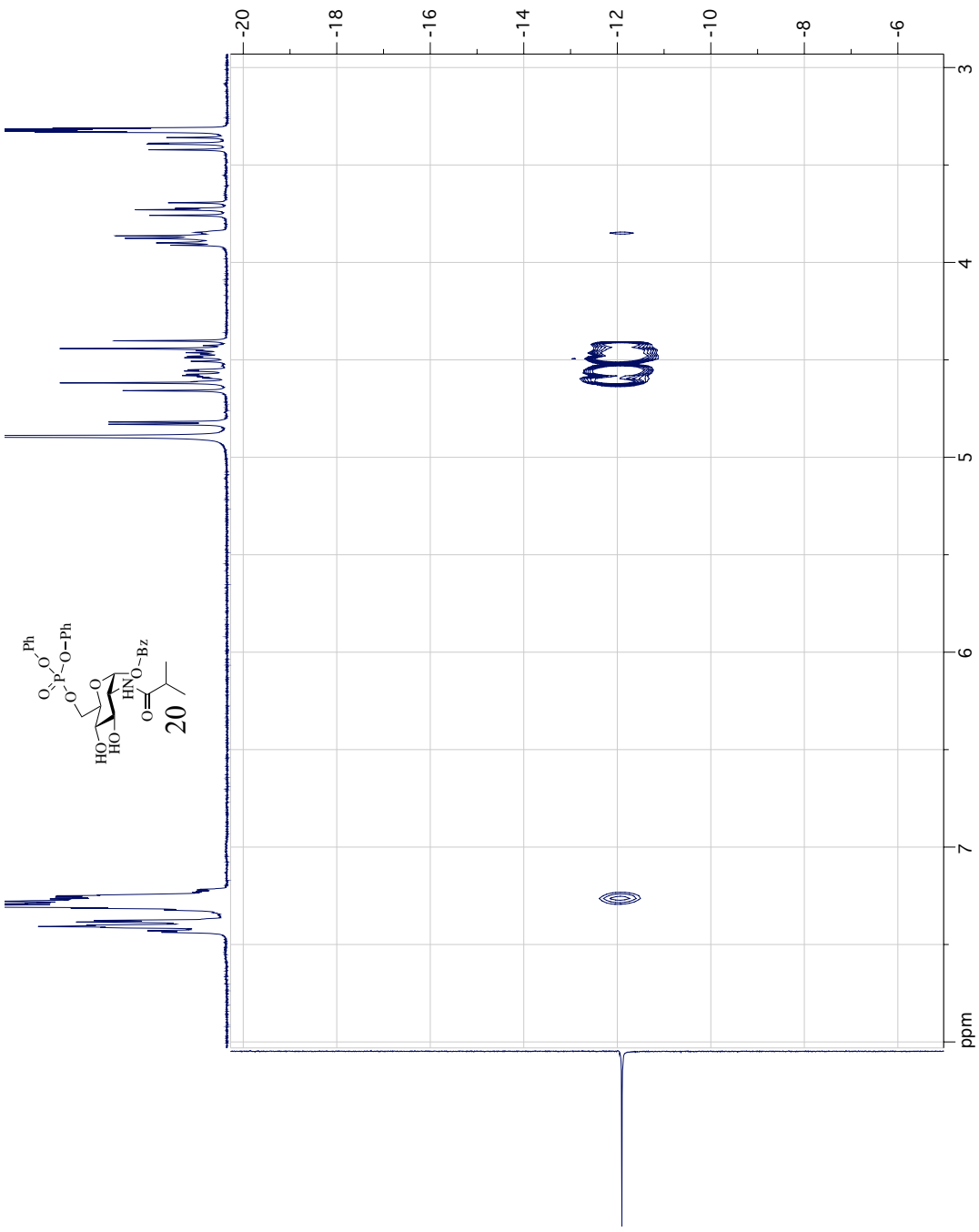
SD 28. ^{1}H-NMR of O-benzyl-2-*i*-butylamino-2-deoxy-D-glucopyranose



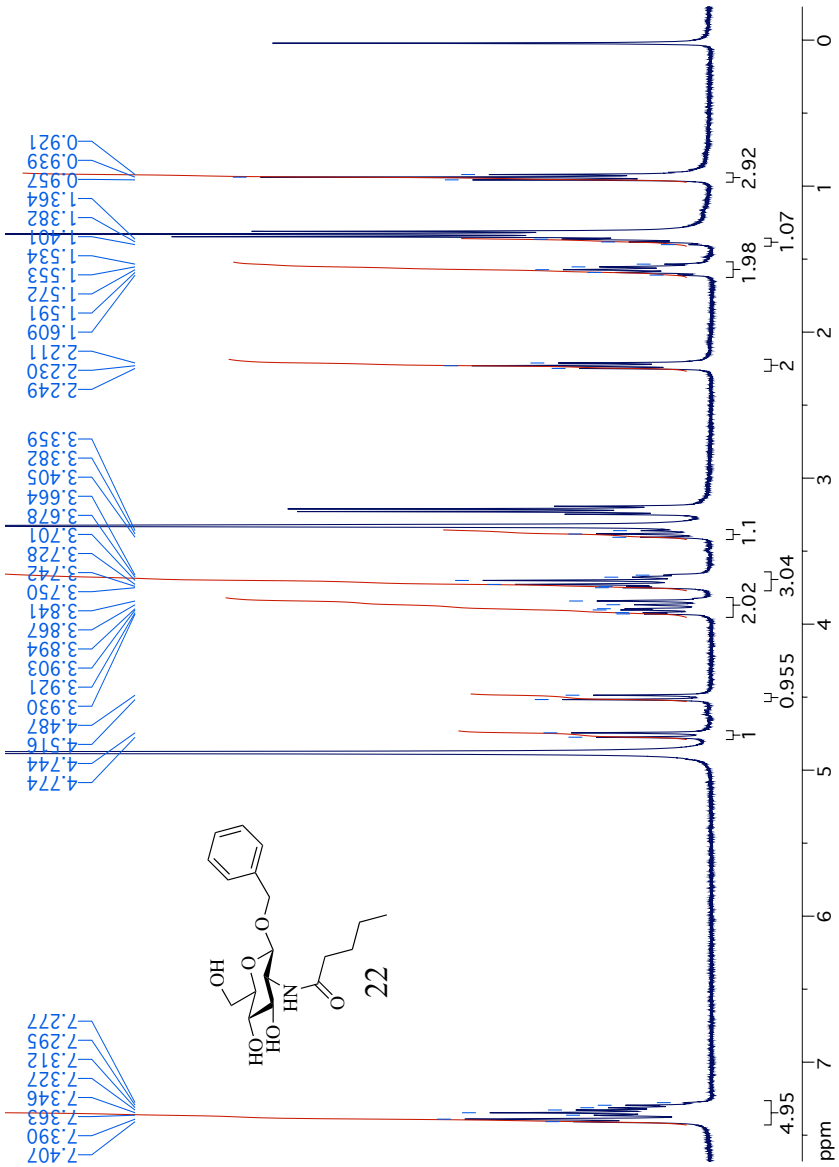
SD 29. ^{1}H-NMR of O-benzyl-6-diphenoxyphosphoryl-2-*i*-butylamino-2-deoxy-D-glucopyranose



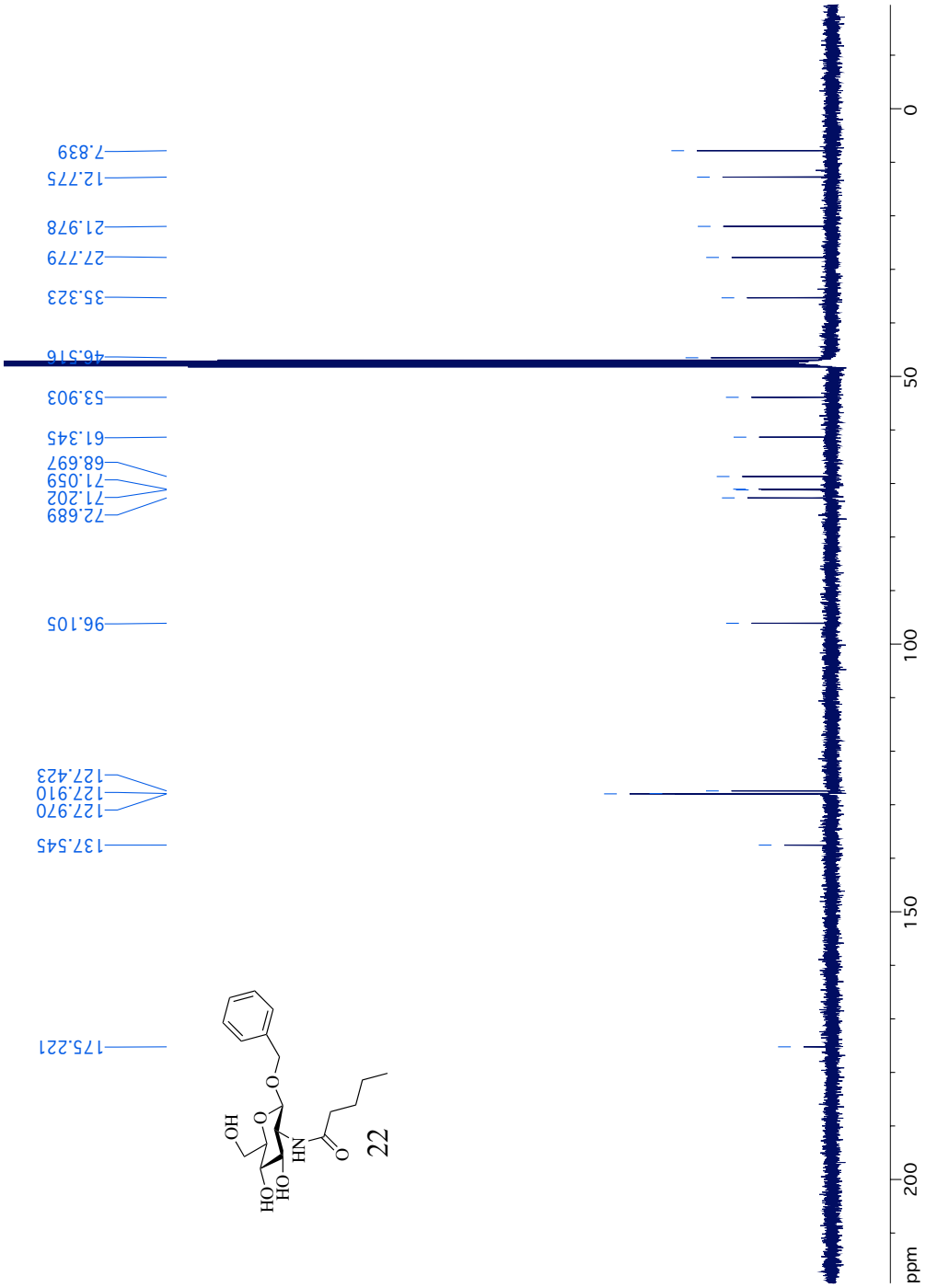
SD 30. ^{31}P-NMR of O-benzyl-6-diphenoxyphosphoryl-2-*i*-butylamino-2-deoxy-D-glucopyranose



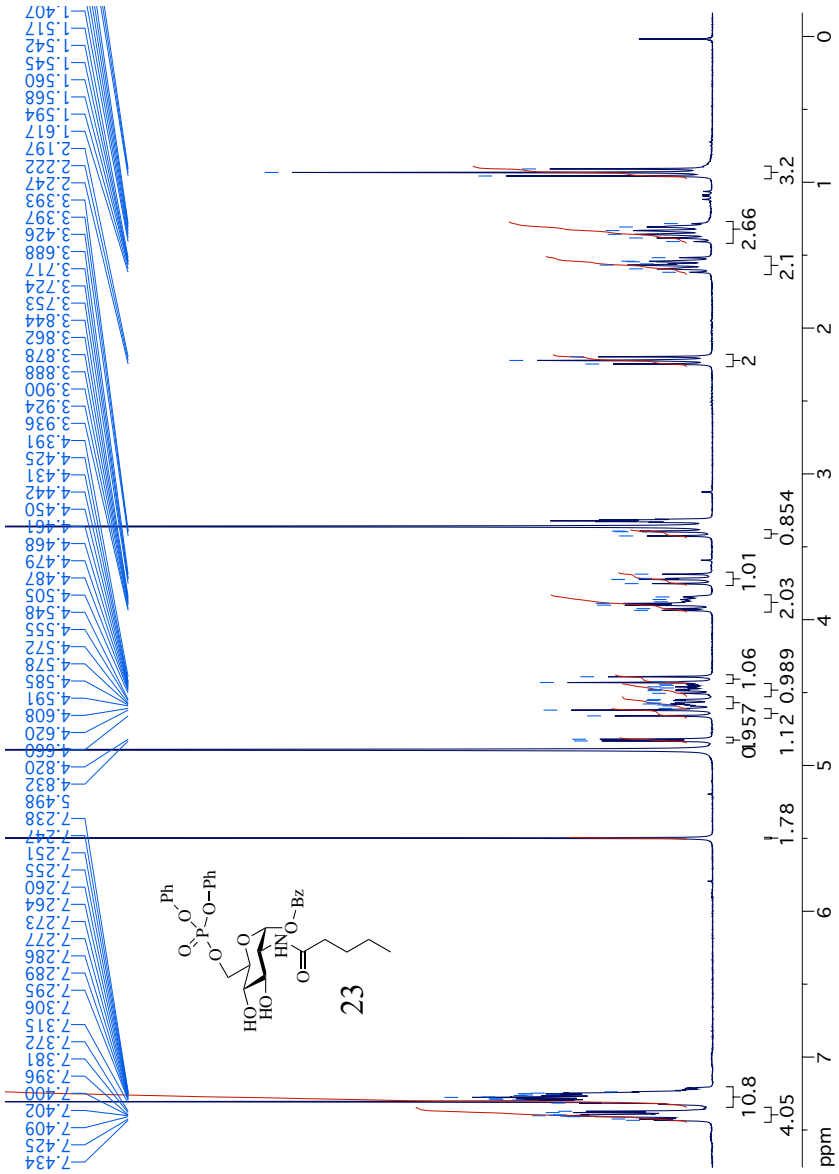
SD 31. ^{1}H-^{31}P HMBC of O-benzyl-6-diphenoxyphosphoryl-2-*i*-butylamino-2-deoxy-D-glucopyranose



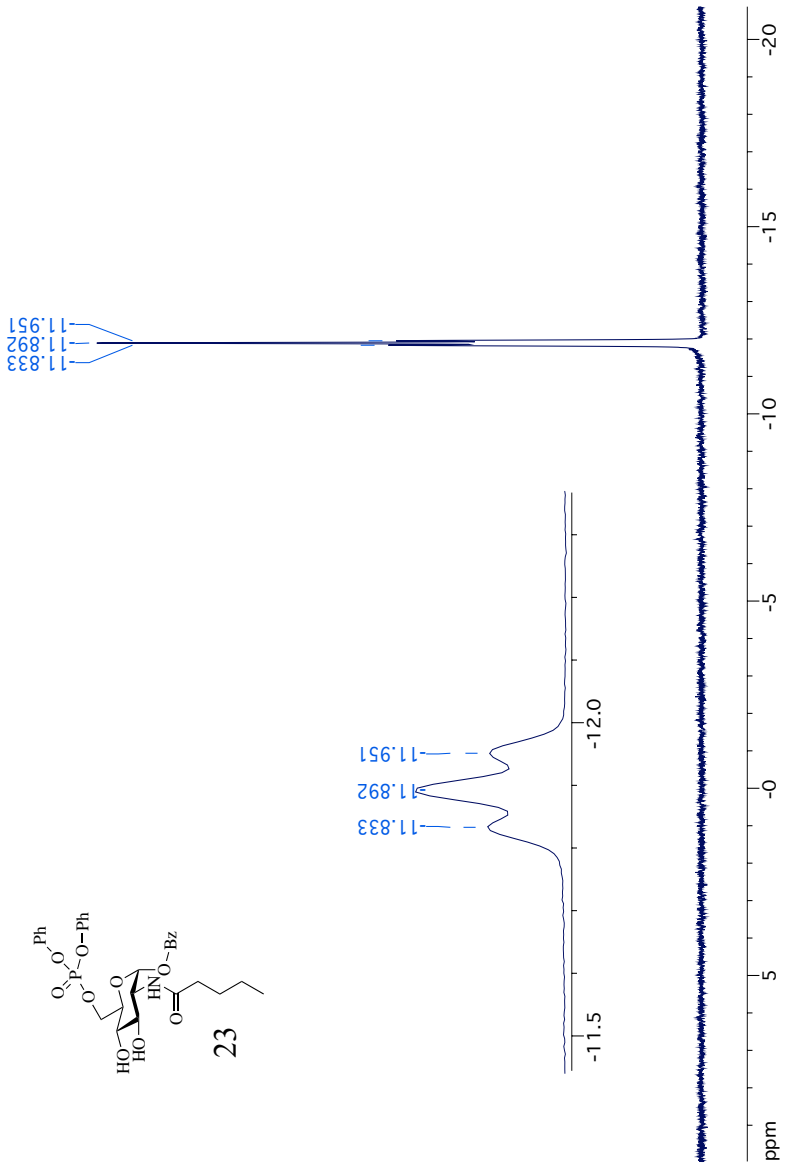
SD 32. ^{1}H-NMR of O-benzyl-2-valeroylamino-2-deoxy-D-glucopyranose



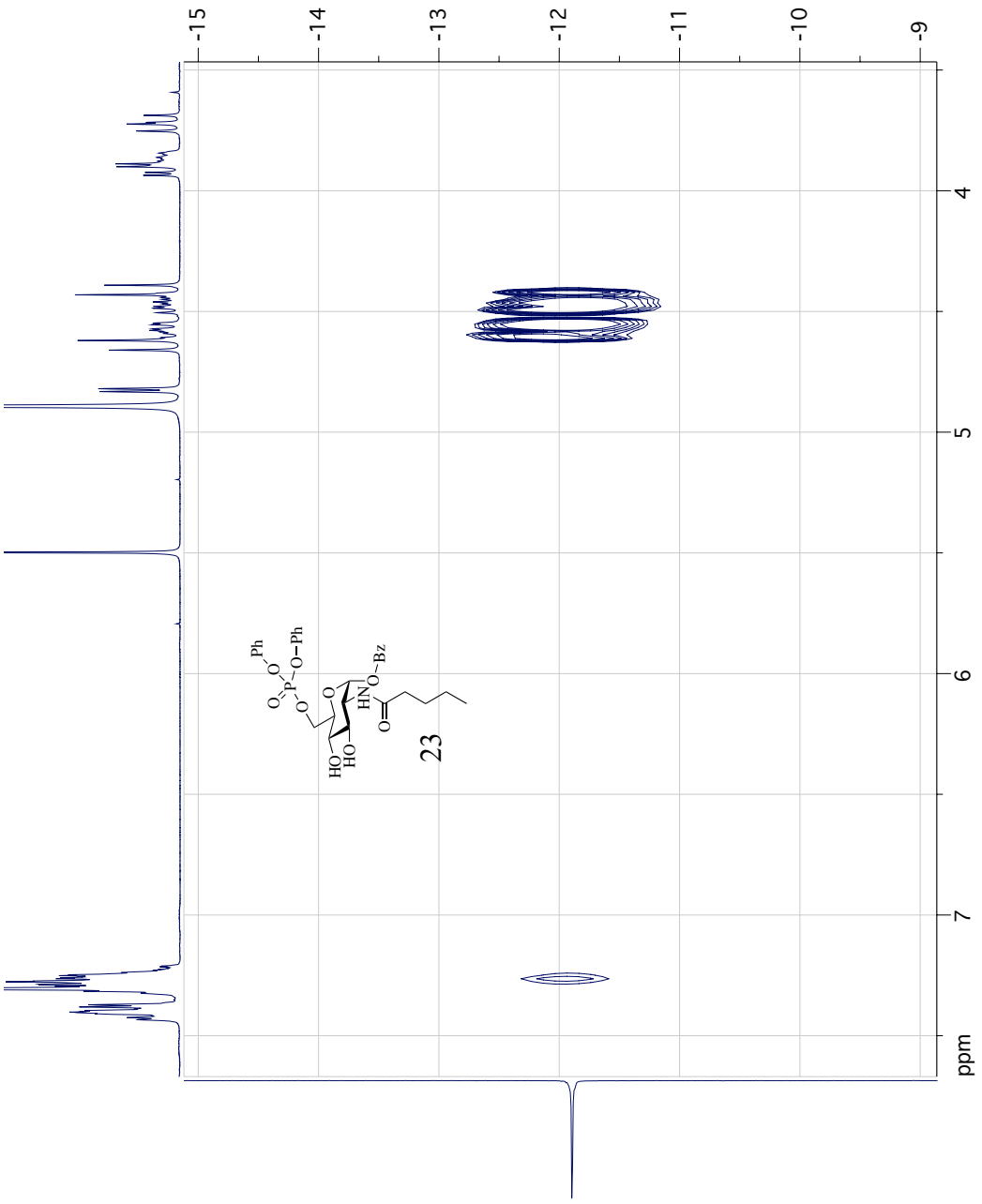
SD 33. ^{13}C-NMR of O-benzyl-2-valeroylamino-2-deoxy-D-glucopyranose



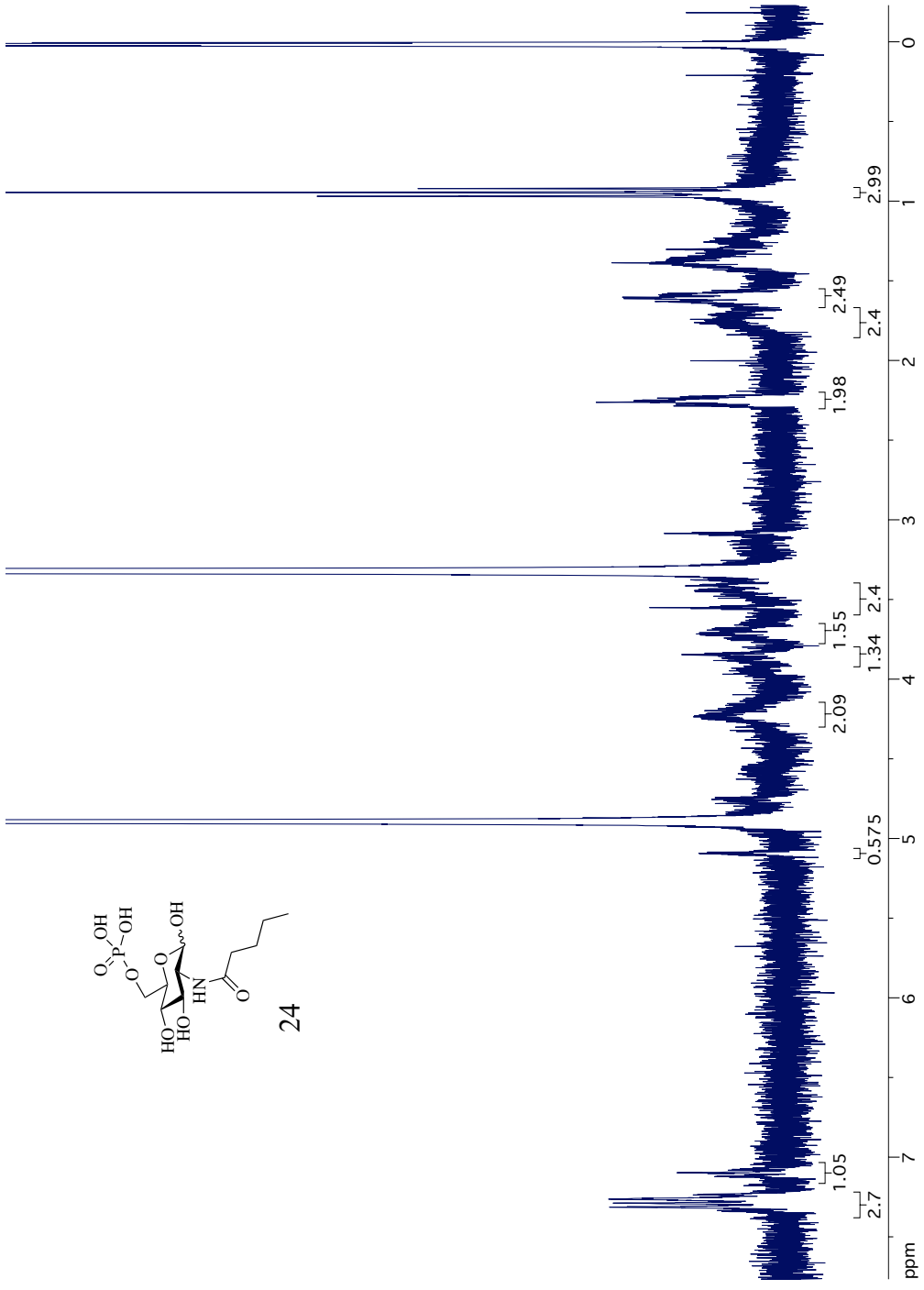
SD 34. ^{1}H-NMR of O-benzyl-6-diphenoxyphosphoryl-2-valeroylamino-2-deoxy-D-glucopyranose



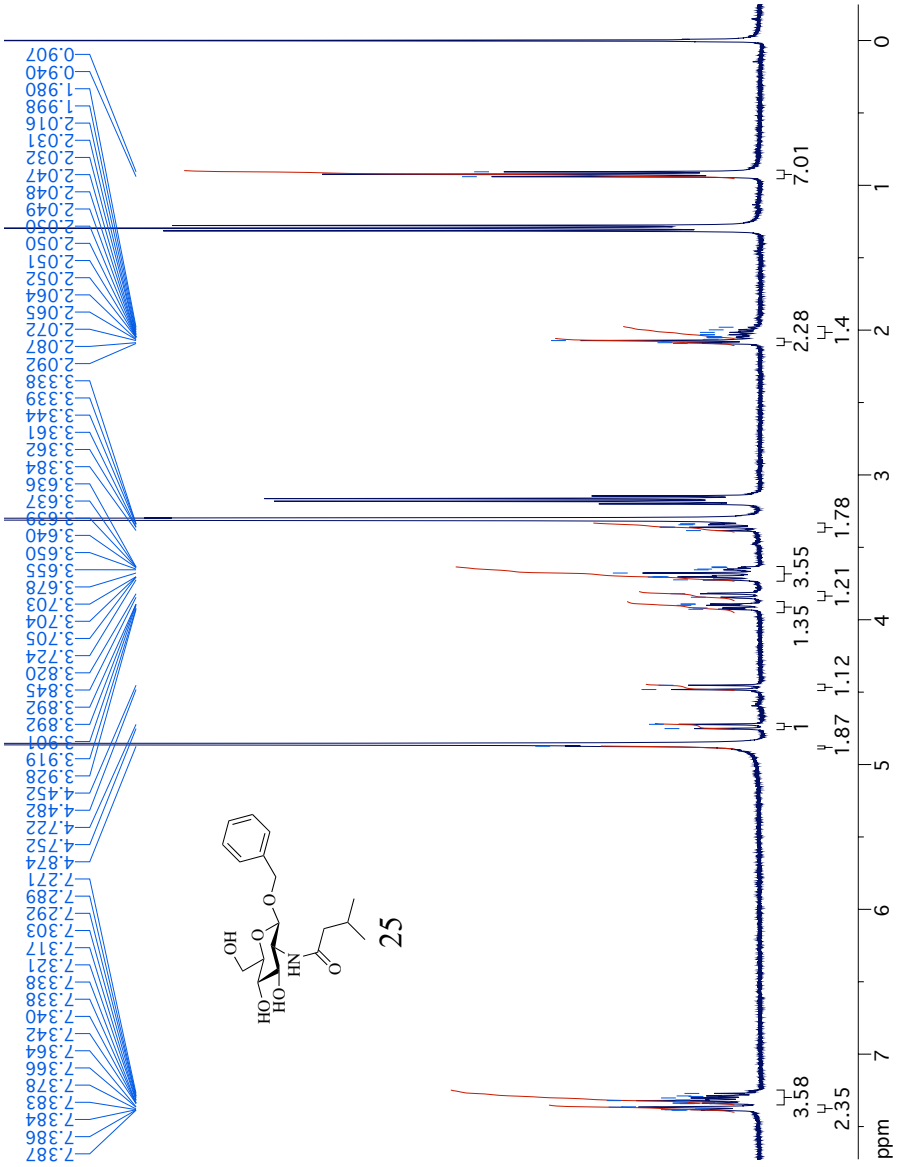
SD 35. ^{31}P-NMR of O-benzyl-6-diphenoxyphosphoryl-2-valeroylamino-2-deoxy-D-glucopyranose



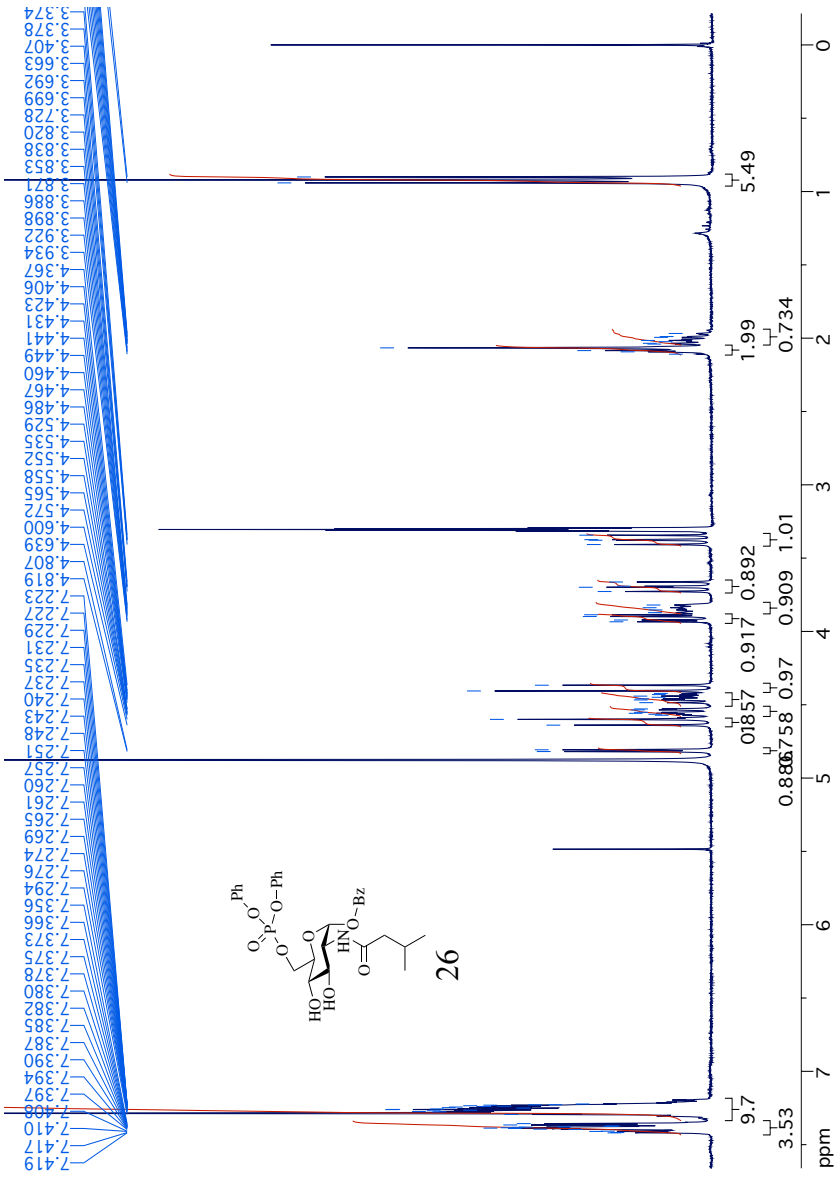
SD 36. ^{1}H-^{31}P HMBC of O-benzyl-6-diphenoxyphosphoryl-2-valeroylamino-2-deoxy-D-glucopyranose



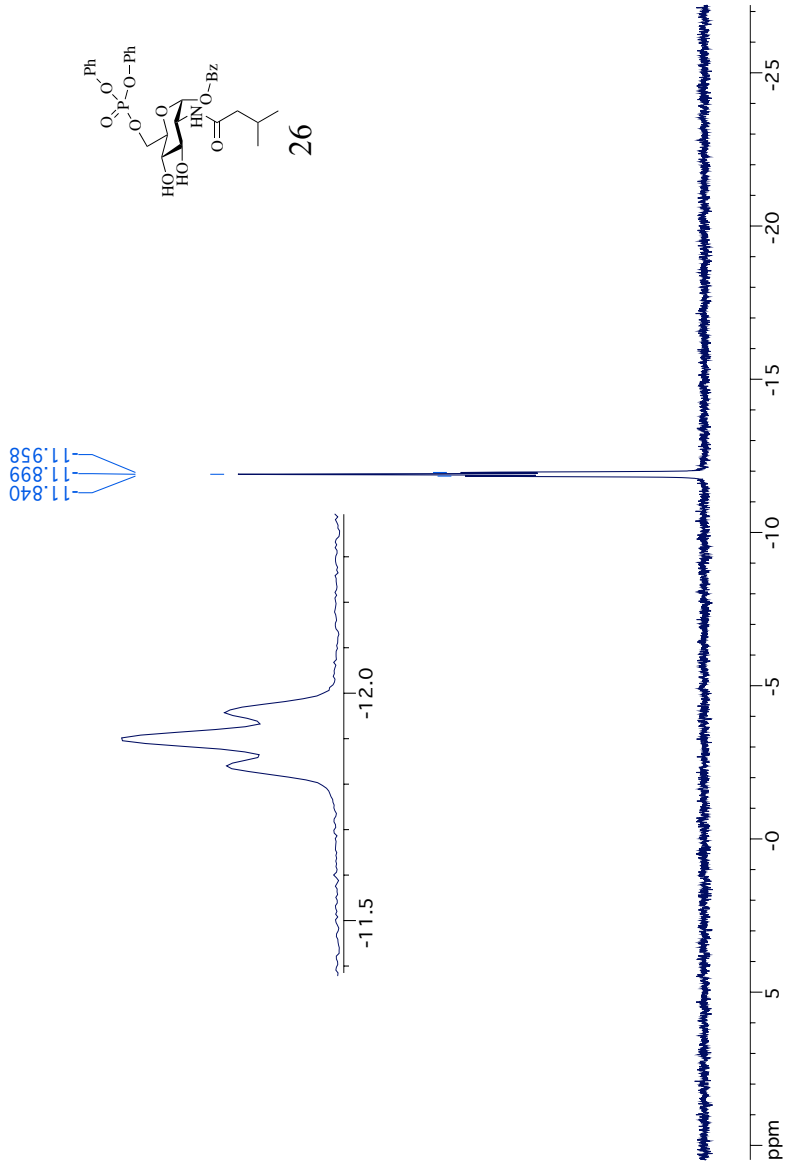
SD 37. ^{1}H-NMR of N-*n*-valeroylglucosamine-6-phosphate



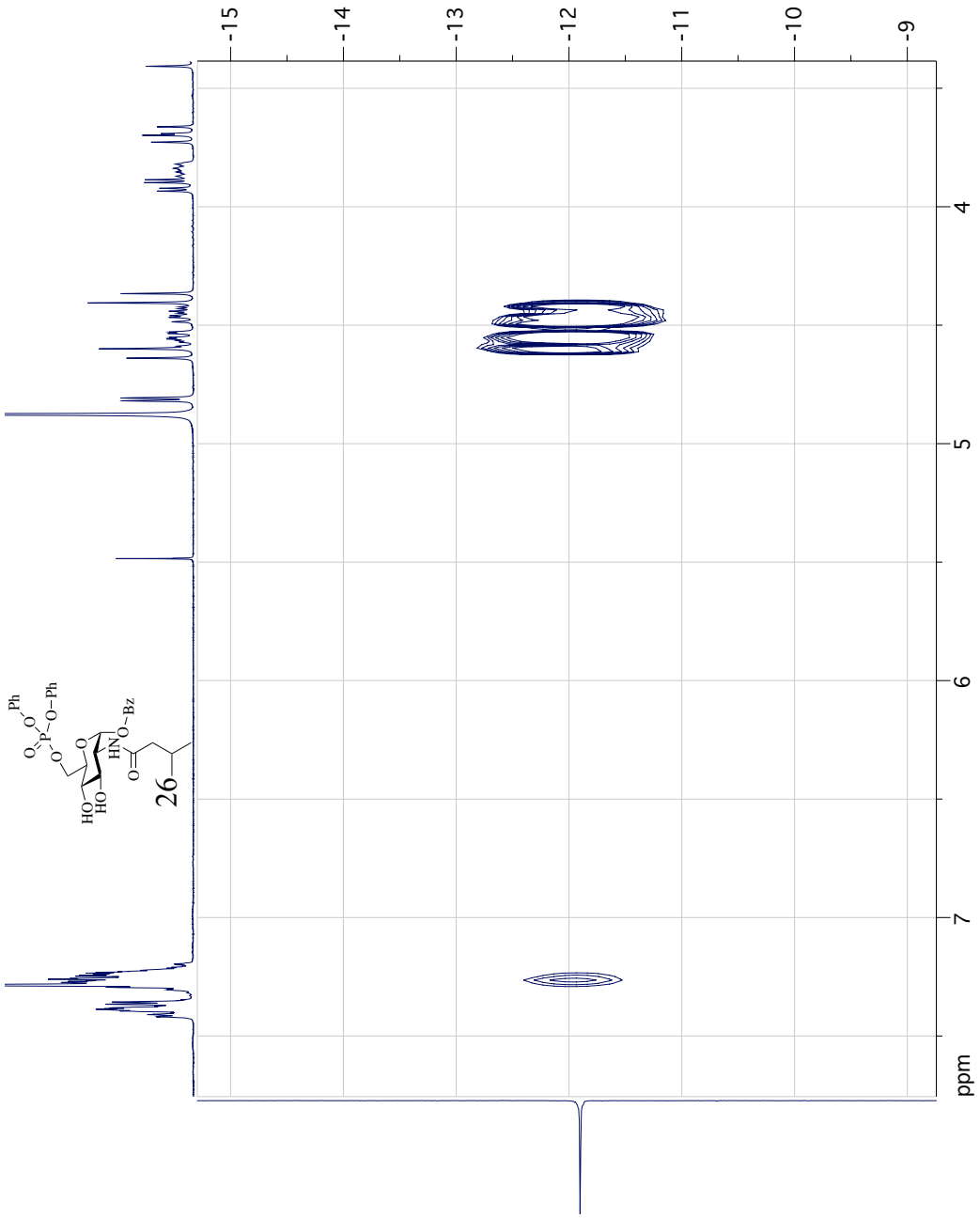
SD 38. ^{1}H-NMR of O-benzyl-2-*i*-valeroylamino-2-deoxy-glucopyranose



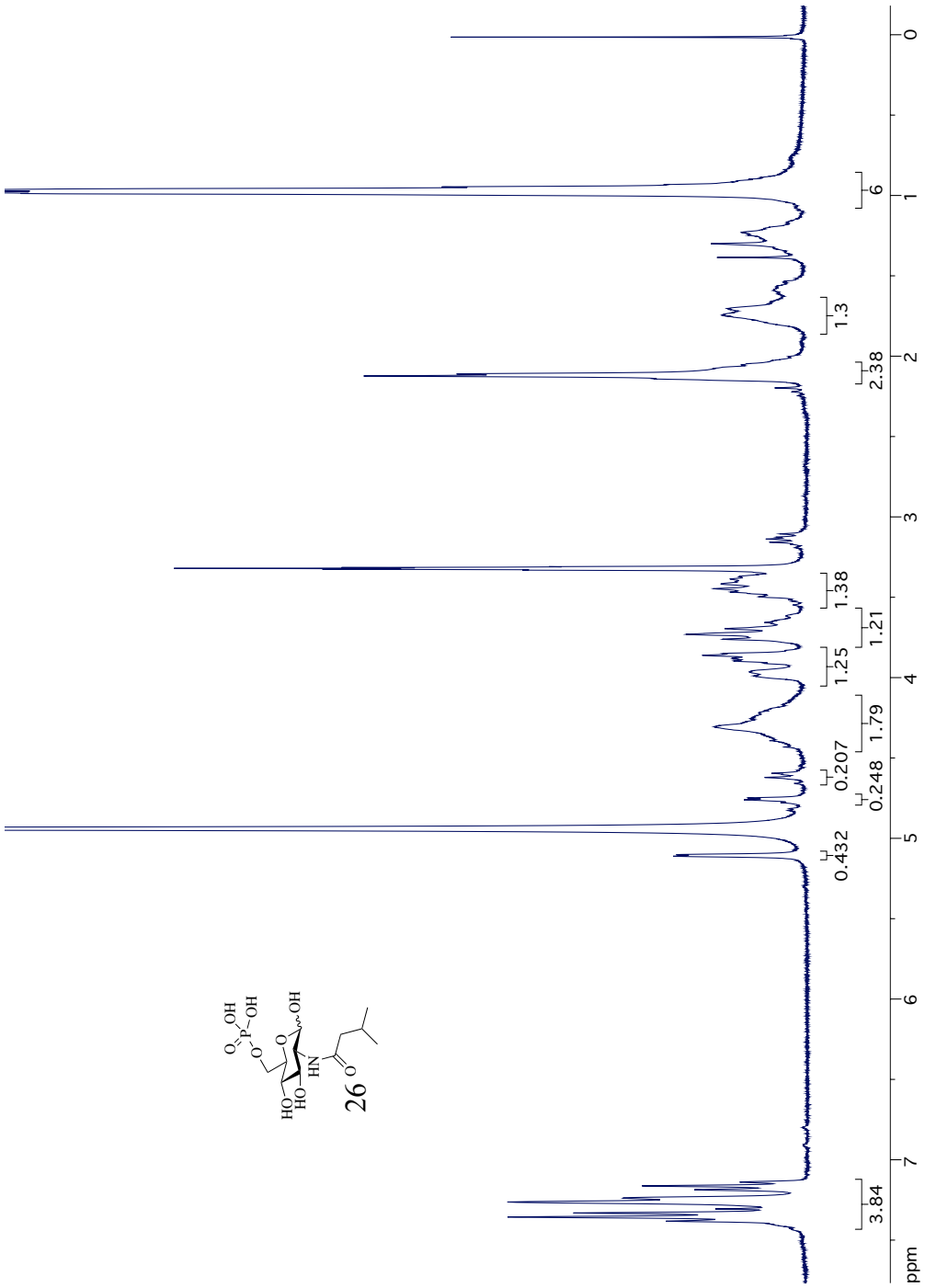
SD 39. ^{1}H-NMR of O-benzyl-6-diphenoxyphosphoryl-2-*i*-valeroylamino-2-deoxy-glucopyranose



SD 40. ^{31}P-NMR O-benzyl-6-diphenoxyphosphoryl-2-*i*-valeroylamino-2-deoxy-glucopyranose



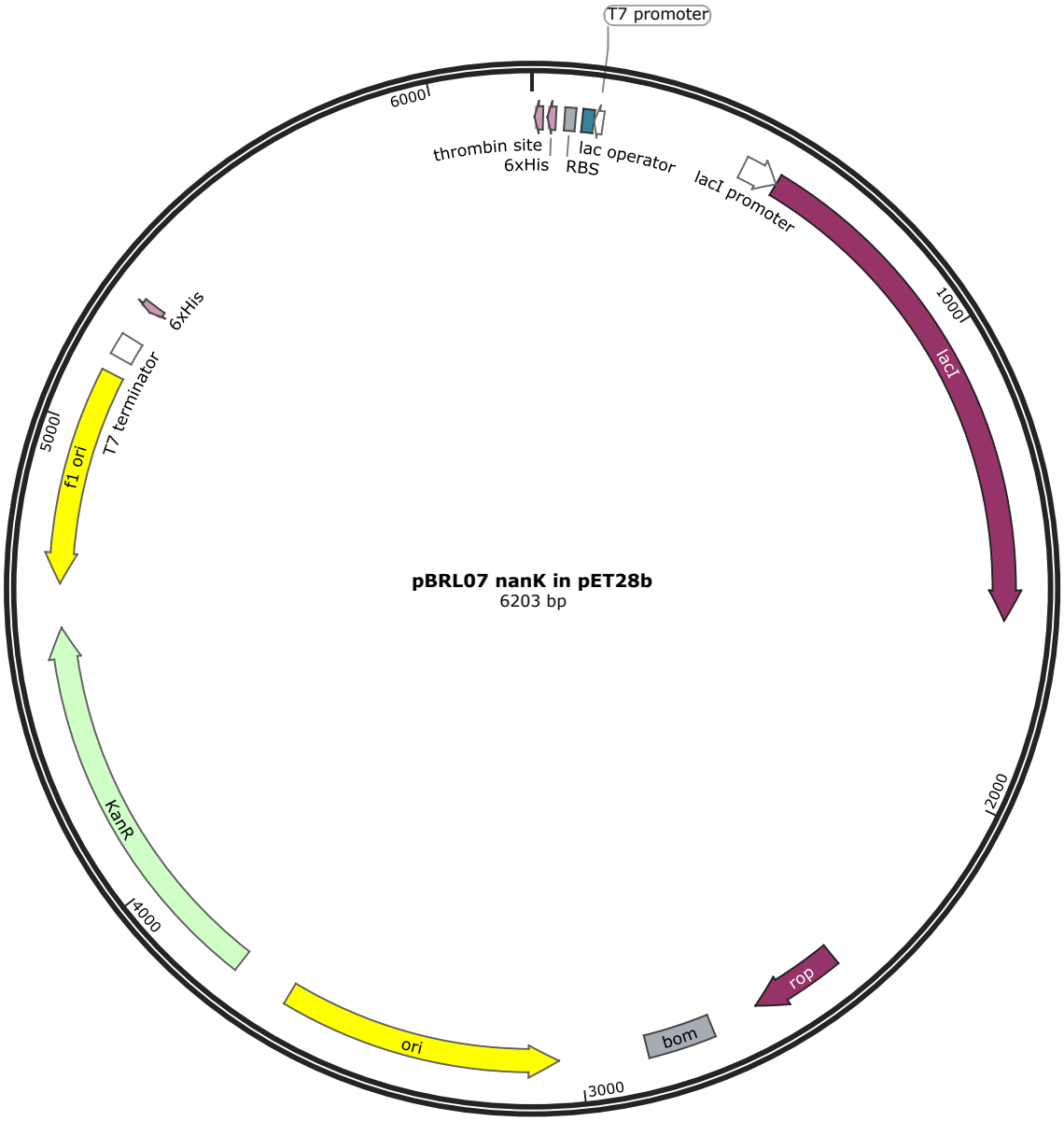
SD 41. ^{1}H-^{31}P HMBC of O-benzyl-6-diphenoxyphosphoryl-2-*i*-valeroylamino-2-deoxy-glucopyranose



SD 42. ^{1}H-NMR of N-*i*-valeroylglucosamine-6-phosphate

3.3.5 Plasmids maps

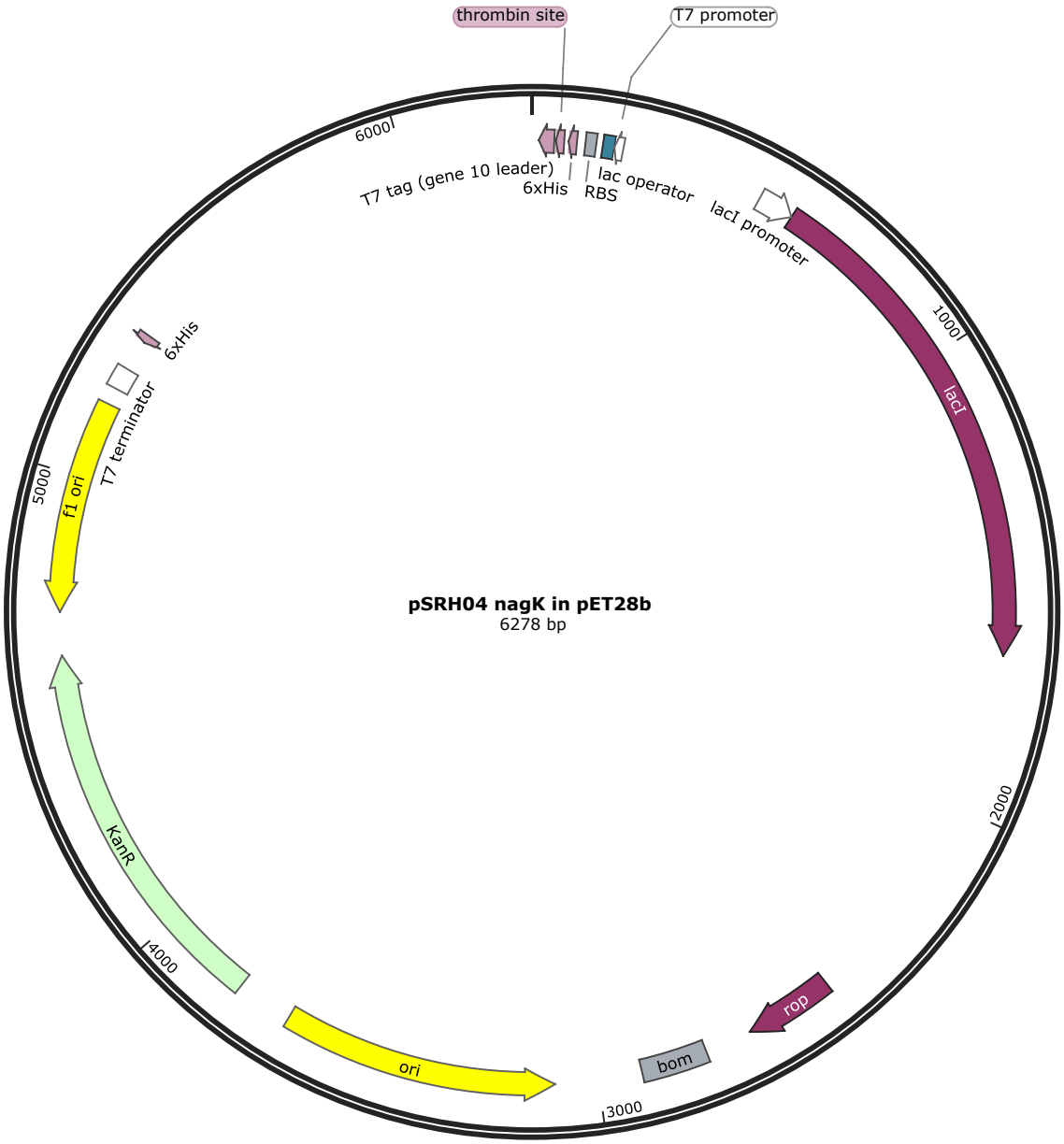
Sequence: pBRL07 nanK in pET28b.dna (Circular / 6203 bp)
Features: 14 visible, 14 total



Printed from SnapGene®: Oct 11, 2016 4:08 PM Page 1

Plasmid map of pBRL07, nanK in pET28b

Sequence: pSRH04 nagK in pET28b .dna (Circular / 6278 bp)
Features: 15 visible, 15 total



Plasmid map of pSRH04, nagK in pET28b

Chapter 4. Study of the epimerization reaction catalyzed by NanE

4.1 Introduction

Sialic acid metabolism has been well studied both in eukaryotic and prokaryotic cells and common between the two pathways is the reliance on N-acetylmannosamine (ManNAc) as a key building block[1–5]. Three different 2-epimerases (the first two enzymes are found in eukaryotic cells and the last one is restricted to bacteria[6]) are associated with ManNAc metabolism: UDP-N-acetylglucosamine 2-epimerase (EC 5.1.3.14) that catalyzes the transformation of UDP-N-acetylglucosamine into UDP-N-acetylmannosamine; GlcNAc 2-epimerase (EC 5.1.3.8) that catalyzes the conversion of GlcNAc into ManNAc; and the GlcNAc-6-P 2-epimerase (EC 5.1.3.9) that catalyzes the epimerization between GlcNAc-6-P and ManNAc-6-P[1,6] (table 4.1).

Table 4. 1. 2-epimerases found in eukaryotic and prokaryotic organisms that are related to ManNAc metabolism[6].

EC 5.1.3.14	UDP-GlcNAc ⟶	UDP + ManNAc	Eukaryotic
EC 5.1.3.8	GlcNAc ⇌	ManNAc	Eukaryotic
EC 5.1.3.9	GlcNAc-6-P ⇌	ManNAc-6-P	Prokaryotic

The eukaryotic epimerase EC 5.1.3.8 works on non-phosphorylated sugars, which could make it a good candidate to be incorporated into *E. coli* due to the fact that the sialic acid synthesis reaction involves non phosphorylated substrates. However, prokaryotic organisms uptake sugars using the PTS system, which phosphorylate the sugar in the transfer process. This means that all sugars in bacteria will be phosphorylated at some

point. For this reason, this eukaryotic epimerase is not suitable for the biosynthesis of sialic acid.

Lundgren *et al.*[7] showed that the anabolic enzymes in the sialic acid biosynthesis route are highly substrate specific, making them not suitable for sialic analogs biosynthesis. GlcNAc-6-P epimerase (NanE) is the first step of sialic acid catabolic pathway (fig 4.1). This enzyme catalyzes the epimerization of ManNAc-6-P into GlcNAc-6-P (fig 4.2). This enzymatic reaction is reversible, which makes it a possible access point for the use of N-acylglucosamines as precursors for sialic acid analogs production.

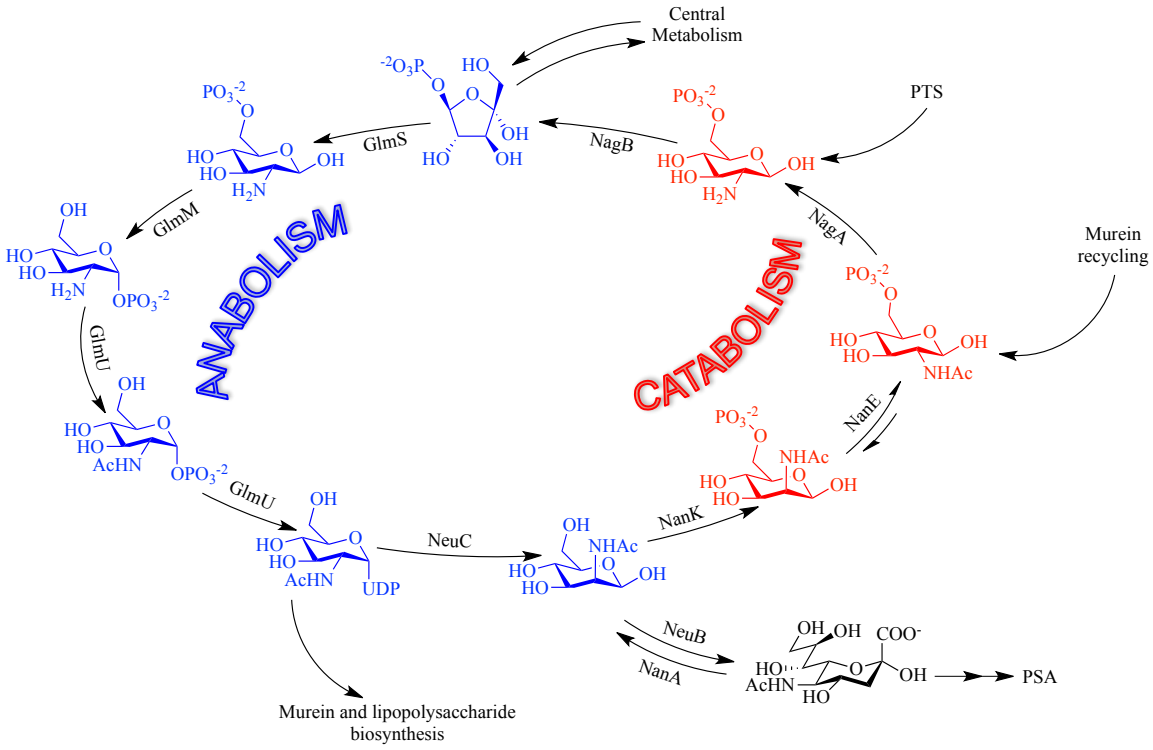


Figure 4. 1. Sialic acid metabolic pathways representation.

NanE catalyzes the conversion of N-acetylmannosamine-6-phosphate into N-acetylglucosamine-6-phosphate (fig 4.2). The equilibrium in this reaction is displaced

towards GlcNAc-6-P, however, based on the microscopic reversibility principle, if the enzyme is incubated with the product, it should produce ManNAc-6-P, albeit in low levels.

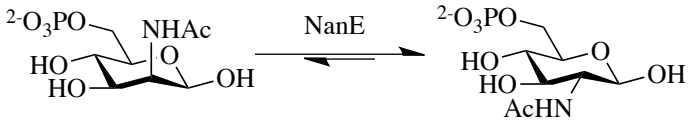


Figure 4. 2. Epimerization reaction catalyzed by NanE.

The mechanism for this enzymatic reaction is shown in figure 4.3. The reaction starts with opening of the pyranose ring, which is followed by a base deprotonating the C2 hydrogen to produce the enolate intermediate[8]. Protonation of the enolate occurs from the opposite face, placing the substituent into the opposite position. After reformation of the pyranose ring, the acetamide group is located in axial position[8].

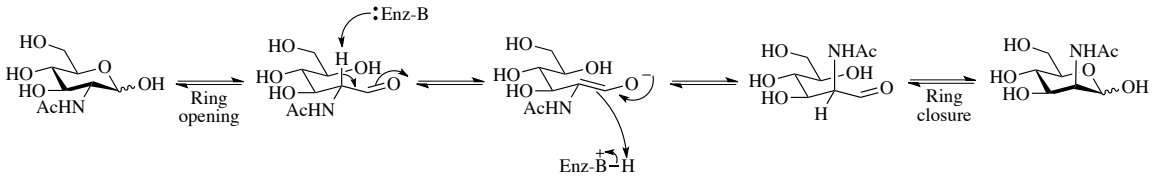


Figure 4. 3. Epimerization mechanism of NanE.

4.2 Results and Discussion

It is no secret that the analysis of carbohydrates is a challenging task. Sugars lack of chromophores makes them invisible to UV detectors and the highly polar nature prevent them to be retained in reverse phase resins or, on the other hand, they are strongly retained in normal phase chromatography stationary phases.

Epimers are stereoisomers that differ in the configuration of one of the stereogenic centers. This small difference confers very similar physical-chemical properties; however, the chemical environment is different.

NMR provides high-resolution information of the chemical environment around each NMR active nuclei. It is thus able to discriminate between the two epimers due by relying on the difference in the chemical shifts of specific protons in the analytes. We therefore selected NMR as the detection and quantification tool to analyze NanE-mediated epimerization of ManNAc.

Before the analysis of the enzymatic epimerization reaction, we had to develop the NMR detection and quantification method. The first problem to be addressed was the presence of water in the sample. Due to the aqueous mixture where the reaction was performed, a very strong water peak was interfering with the sample signals (fig. 4.4). For this reason, NMR spectrum was collected via presaturation. In this method, water resonance was saturated by a long low-energy pulse (2 seconds in our case), then, a short high-energy pulse is applied, and the spectrum is collected. D_2O was added to a final concentration of 10% to help to lock the signal.

Figure 4.5 shows the H^1-NMR of GlcNAc with water signal suppression. We can see how the intensity of the water signal decreases considerably allowing seeing the C-H signals of the pyranose ring (shift 3.2 – 3.8) and one of the anomeric protons signals (δ 5.0).

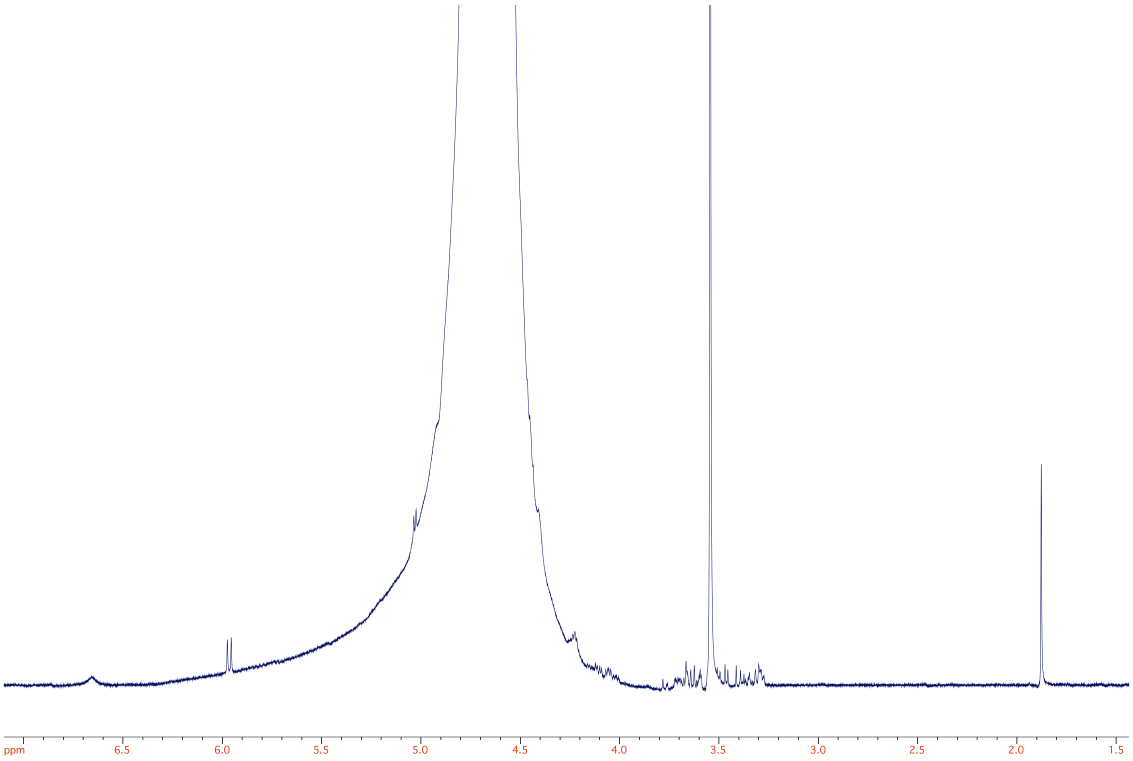


Figure 4. 4. Proton NMR of GlcNAc.

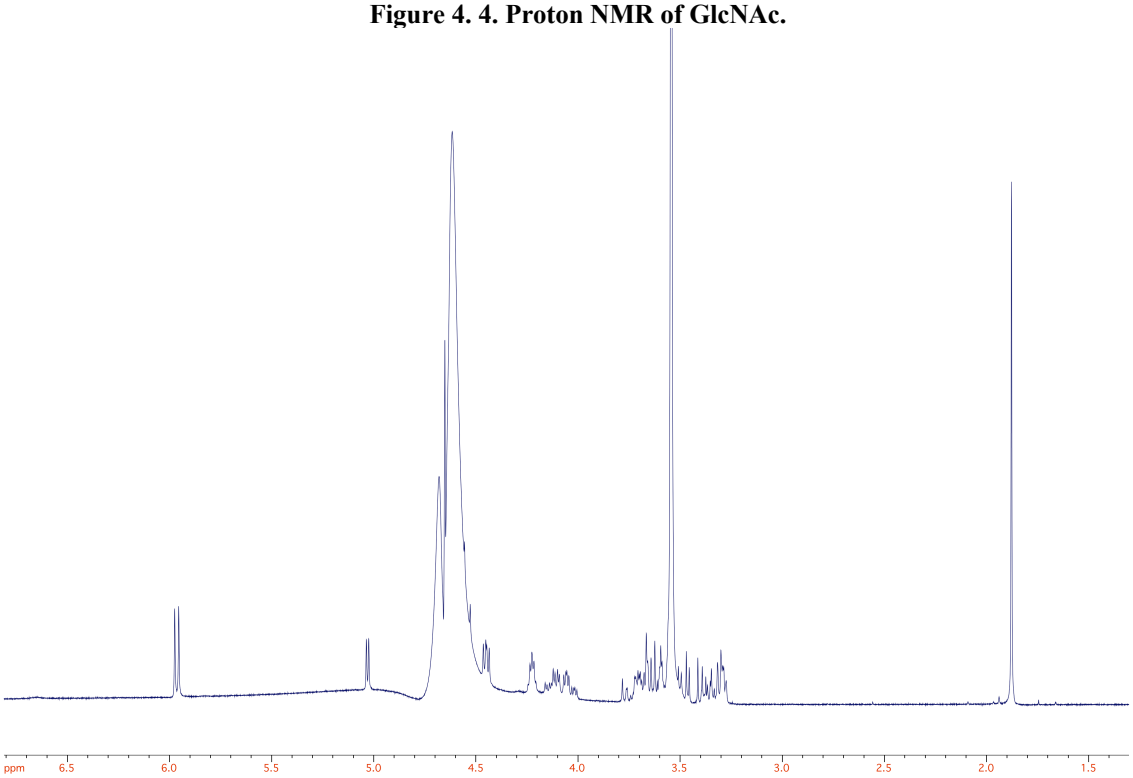


Figure 4. 5. Proton NMR of GlcNAc with water suppression.

Once solved the problem with the intensity of the water signal, the next step was to determine if we were able to differentiate between GlcNAc and ManNAc. Sugars were dissolved in Tris saline buffer pH 7; NagK was added and the mixture and ATP was used as phosphate source for the kinase; this kind of solution resembles the conditions for the enzymatic reaction. In addition, D_2O was used to lock the signal in the NMR. Spectra were collected in a 300 MHz NMR.

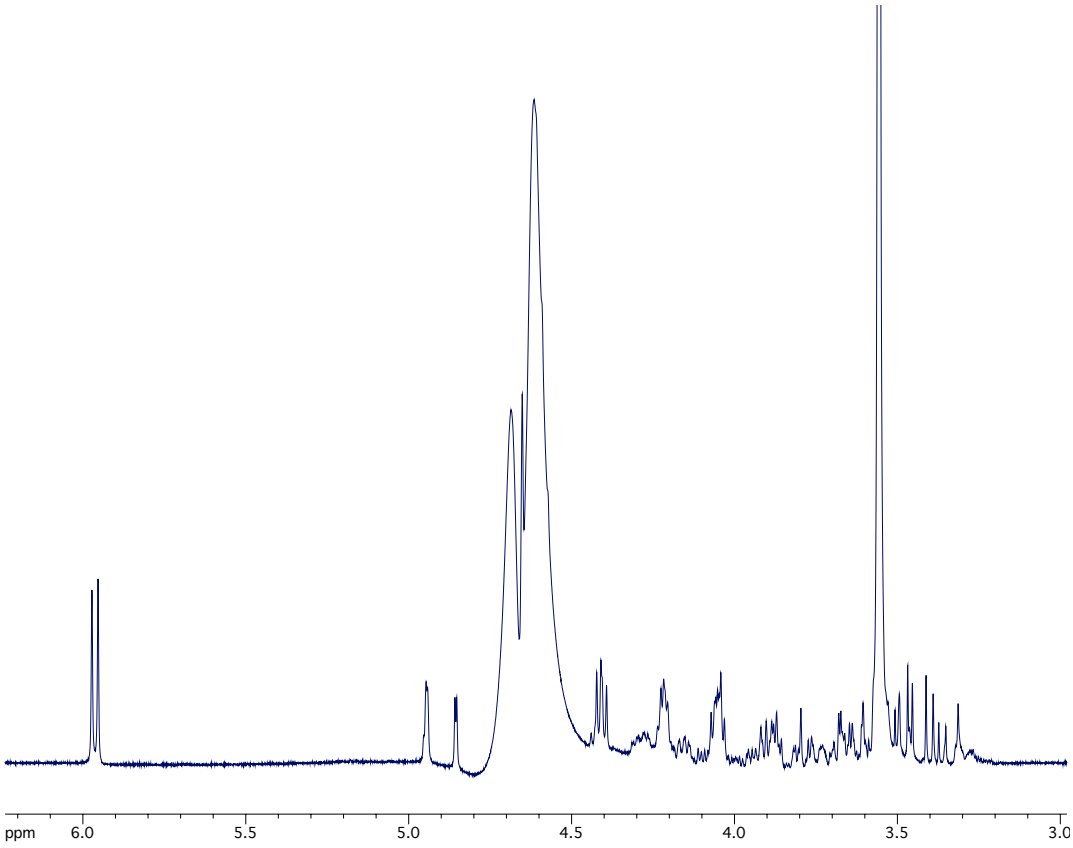


Figure 4. 6. Proton NMR of ManNAc with water suppression.

Having addressed the concern of dynamic range of the NMR assay, we investigated the ability to distinguish GlcNAc from ManNAc by this approach. Both aminosugars were analyzed by NMR under the same experimental conditions; figures 4.5 and 4.6 show proton

NMR spectra with water suppression of GlcNAc and ManNAc respectively. An important region to pay attention is between 4.8 – 5.1 ppm. In fig 4.5 (GlcNAc) we see only one signal at 5.1 ppm, on the other hand, in ManNAc NMR (fig. 4.6), the signal at 5.1 is absent but there are two signals with a shift of 4.87 and 4.96 ppm. This main aspect allows distinguishing between GlcNAc and ManNAc easily.

In order to investigate the sensitivity of the assay, the sugars were tested in two different concentrations in this experiment. Sugars were mixed to a final concentration of 10 mM and 1 mM for the 1:10 ratio and in 10 mM and 0.1 mM for the 1:100 ratio.

Figure 4.7 shows the H^1-NMR for the 1:10 ratio. The red spectra correspond to the mixture GlcNAc:ManNAc 1:10 and the blue spectra corresponds to the inverted ratio (GlcNAc:ManNAc 10:1). Careful examination in the region between 4.8 ppm and 5.2 ppm, shows the presence of two different signals. The doublet at 5.0 corresponding to the GlcNAc, and another one around 4.9 corresponding to the ManNAc. When both spectra are compared, is evident that the intensity of the signals varies with the sugar concentration in the sample. In other words, when the GlcNAc concentration increases, the intensity of the doublet at 5.0 increases. Based on these results we conclude that a 10% conversion in the epimerization reaction would be detectable using this method, making it appropriate for qualitative characterization of NanE-mediated isomerization.

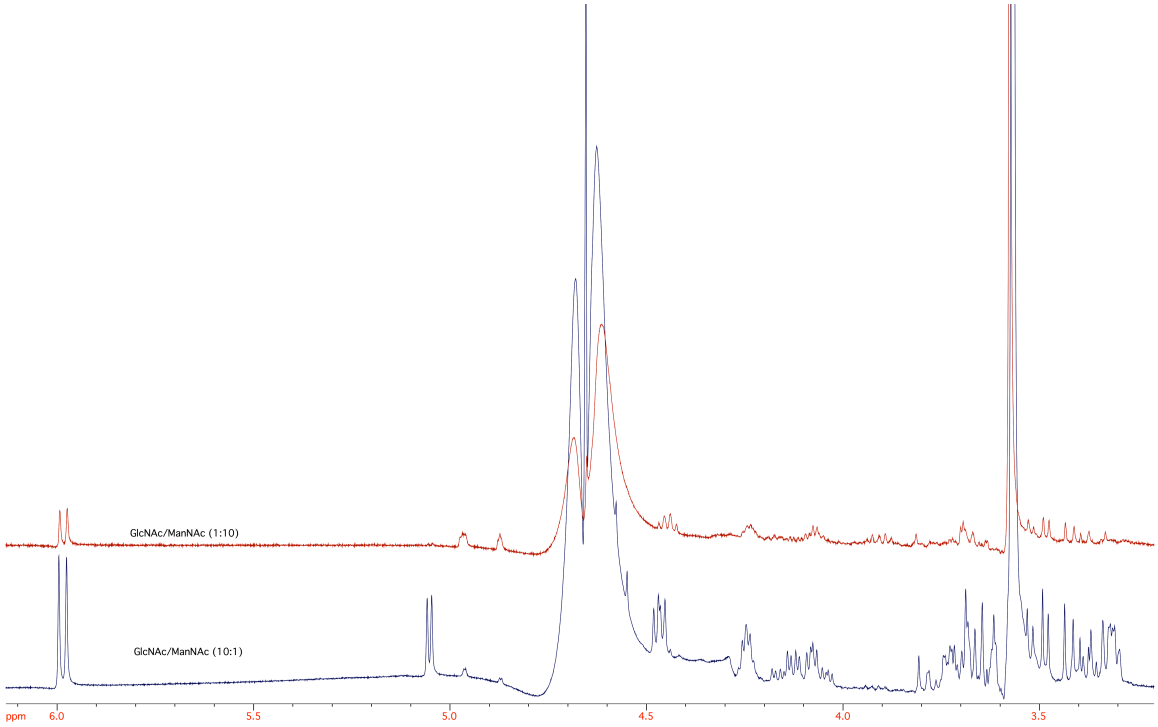


Figure 4. 7. Proton NMR of ManNAc and GlcNAc with a concentration ratio of 1:10.

The next step was to determine if a 1% conversion could be in the detection limit of the experiment. For this, a mixture of GlcNAc and ManNAc in a concentration ratio of 1:100 (0.1:10 mM) was prepared and analyzed by NMR. Figure 4.8 shows the H^1-NMR of the mixture of sugars. However, when the region of 4.8-5.2 ppm is analyzed, only one of the set of signals is clearly observed. From this data we can conclude that 1% conversion of GlcNAc into ManNAc would not be detectable and in addition this analytical method would not be appropriate for quantitative evaluation of conversion less than 10%.

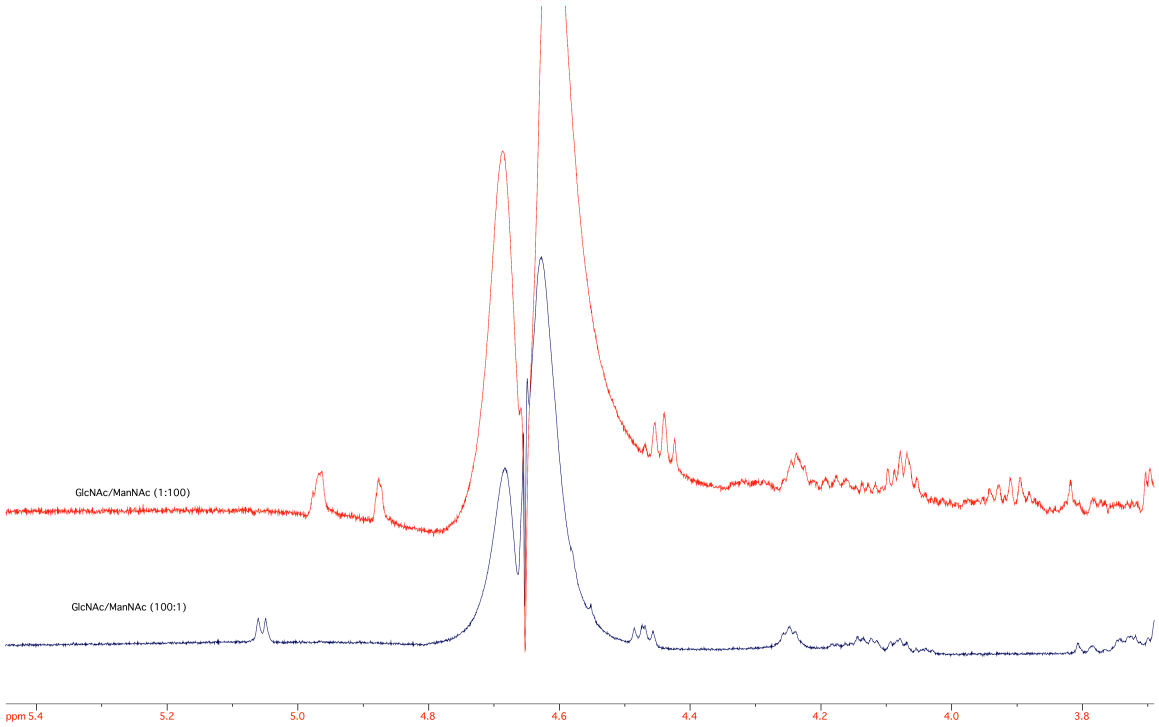


Figure 4. 8. Proton NMR of ManNAc and GlcNAc with a concentration ratio of 1:100.

Once the analytical method was developed, the epimerization reaction was performed and studied using NMR.

NanE was overexpressed on *E.coli* BL21 (D3) using the plasmid pBRL06 (annex). Dr. Lundgren, as part of his doctorate project, built the plasmid pBRL06 in our lab. After the overexpression procedure, the enzyme was isolated using nickel affinity chromatography and dialyzed against saline phosphate buffer 100 mM, pH 7.4.

For the NMR experiment, NanE was mixed with a solution containing 10 mM ManNAc (natural substrate). Another test was set with GlcNAc (epimerization product). In the first case, we expected that ManNAc will be transformed into GlcNAc and the doublet at 5.0 should appear. On the other hand, due to the equilibrium, if we use GlcNAc as substrate for NanE, the ManNAc signals around 4.9 should appear.

Figure 4.9 shows the result of this experiment. The red spectrum shows the epimerization reaction when ManNAc is used as substrate. The result was consistent with our expectation that the doublet at 5.0 should appear as the enzymatic reaction progresses. This data confirms that the enzyme is active and that we can detect isomerization of ManNAc into GlcNAc. On the other hand, the experiment with GlcNAc as substrate shows the doublet at 5.0, without any evidence of the ManNAc signals. This result does not mean that the reverse reaction is not occurring. Our analytical method development work shows that lower limit of detection with the NMR method is higher than 1 mM but lower that 10 mM, thus we can only conclude that NanE did not produce sufficient ManNAc to be detected in our assay.

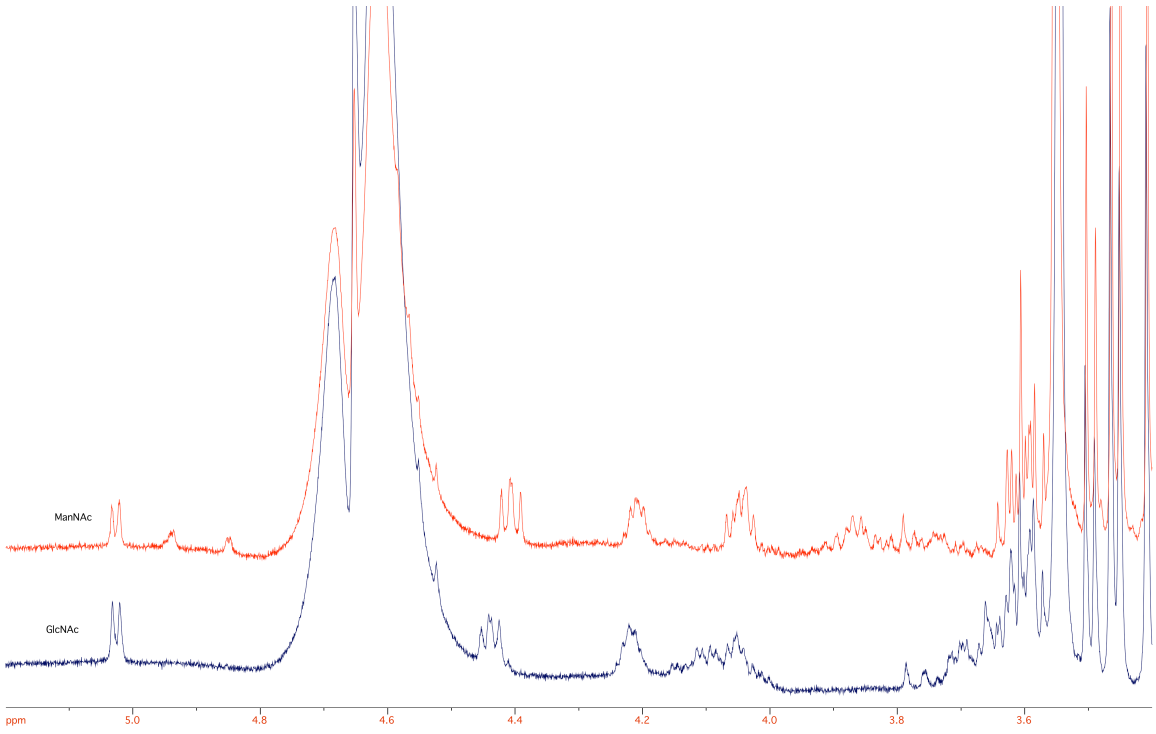


Figure 4. 9. Proton NMR of NanE epimerization reaction using ManNAc (red) and GlcNAc (blue) as substrates.

An important point to discuss is the incubation period used in the reaction. We incubated the reaction at 37°C during a 4 hours period. This period of time is consistent with the results obtained by Brigham *et al*.[9]; they found that after 5 hours, the enzymatic reaction ManNAc to GlcNAc reaches steady state in a ratio of 1:3 ManNAc : GlcNAc. This ratio is coherent if we consider that thermodynamically, GlcNAc is favored with an equilibrium ratio of 1:3.9[8,10] .

When Brigham *et al.* tested the reverse reaction, the epimerization after 18 hours was poor, obtaining an equilibrium ratio of 10:1 GlcNAc:ManNAc[9]. This result suggests that we should be able to detect the epimerization from GlcNAc to ManNAc using our NMR method after 18 hours of reaction. The experiment was set but we were not able to detect any ManNAc after an overnight reaction.

How can we explain these results? We know NanE is active because we could produce GlcNAc from ManNAc, and it is reported that after 18 hours we should get a 1:10 ratio when GlcNAc is used as only substrate and we demonstrated that our NMR experiment can detect those concentrations. The answer to these questions lays in the results obtained in chapter 3, where we demonstrated that kinases are highly substrate-specific. In this experiment, the kinase used was N-acetyl-Mannosamine-6-phosphate epimerase (NanK) in all our experiments, whereas we should change to NagK for the reverse reaction. To conclude, NanE catalyzed the conversion from ManNAc to GlcNAc. In order to use the catabolic pathway of sialic acid to incorporate N-acylglucosamines and be able to produce sialic acid analogs, is necessary the epimerization of these N-acylglucosamines substrates to N-acylmannosamines. In this chapter we worked on developing an assay to evaluate the ability of N-acylGlcN to be converted to N-acyl ManN

by NanE. We were able to show that our assay could readily detect NanE–mediated production of GlcNAc from ManNAc, however we could not detect the reverse reaction when the enzyme was incubated with GlcNAc due to the substrate specificity of the kinase used. Nevertheless, the reverse reaction is likely occurring. Based on Le Chatelier principle and microscopic reversibility the reverse reaction will occur and in the presence of NanE should generate the equilibrium mixture of the two sugars. Le Chatelier's principle, the high limit of detection of NMR analysis suggest that the reaction could occur, but in a concentration too low to be detected by NMR. For this reason, the sialic acid analogs production was the next recommended experiment.

References

1. Ferrero, M. Á. *et al.* Purification and characterization of GlcNAc-6-P 2-epimerase from Escherichia coli K92. *Acta Biochim. Pol.* **54**, 387–399 (2007).
2. Tanner, M. E. The enzymes of sialic acid biosynthesis. *Bioorg. Chem.* **33**, 216–28 (2005).
3. Chou, W. & Hinderlich, S. Sialic acid biosynthesis: stereochemistry and mechanism of the reaction catalyzed by the mammalian UDP-N-acetylglucosamine 2-epimerase. *J. Am. Chem. Soc.* **125**, 2455–2461 (2003).
4. Maru, I., Ohnishi, J., Ohta, Y. & Tsukada, Y. Why is sialic acid attracting interest now? Complete enzymatic synthesis of sialic acid with N-acylglucosamine 2-epimerase. *J. Biosci. Bioeng.* **93**, 258–265 (2002).
5. Viswanathan, K. *et al.* Engineering Sialic Acid Synthetic Ability into Insect Cells: Identifying Metabolic Bottlenecks and Devising Strategies to Overcome Them.

Biochemistry **42**, 15215–15225 (2003).

6. Ringenberg, M. A., Steenbergen, S. M. & Vimr, E. R. The first committed step in the biosynthesis of sialic acid by *Escherichia coli* K1 does not involve a phosphorylated N-acetylmannosamine intermediate. *Mol. Microbiol.* **50**, 961–975 (2003).
7. Lundgren, B. R. & Boddy, C. N. Sialic acid and N-acyl sialic acid analog production by fermentation of metabolically and genetically engineered Escherichia coli †. *Org. Biomol. Chem.* 1903–1909 (2007). doi:10.1039/b703519e
8. Samuel, J. & Tanner, M. E. Mechanistic aspects of enzymatic carbohydrate epimerization. *Nat. Prod. Rep.* **19**, 261–277 (2002).
9. Brigham, C. *et al.* Sialic acid (N-acetyl neuraminic acid) utilization by *Bacteroides fragilis* requires a novel N-acetyl mannosamine epimerase. *J. Bacteriol.* **191**, 3629–3638 (2009).
10. Luchansky, S. J., Yarema, K. J., Takahashi, S. & Bertozzi, C. R. GlcNAc 2-epimerase can serve a catabolic role in sialic acid metabolism. *J. Biol. Chem.* **278**, 8035–8042 (2003).

4.3. Experimental Section

4.3.1. Plasmid Production

One strain of chemical competent *E. coli* BL21 was transformed with plasmid pBRL06 that carries nanE. The strain was placed on a glass test tube and incubated on ice for 5 min. 5 µL of plasmid solution were placed in the bacterial suspension and incubated for 30 min on ice.

After the incubation period the test tube was placed in a 42 °C water bath for 45 s and returned to the ice bath for 2 min. 1 mL of LB broth was added to the tube and it was were incubated at 37 °C for half an hour; then, the bacterial suspension was inoculated on LB plates supplemented with kanamycin and incubated overnight at 37 °C.

Single colonies were taken and inoculated in 5 mL of LB broth supplemented with kanamycin and incubated for 18 h. After the incubation time, the cells were spun down and the plasmids were extracted using the P1, P2, P3 protocol.

4.3.2 Gene expression and protein extraction from E. coli

To produce the seed culture one single colony of the transformants was used to inoculate 10 mL of LB broth supplemented with kanamycin. The culture was incubated at 37 °C and 200 rpm overnight. Next day, 400 mL of fresh LB were supplemented with 400 µL of 1000 × kanamycin solution and inoculated with the seed culture. The bacterial culture was incubated at 37 °C with shaking until OD_{600} reaches 0.3-0.6.

When the proper optical density was reached, IPTG 1M was added to a final concentration of 0.5 mM and the culture was placed in the incubator at 30 °C with shaking.

After 12 to 18 hours of incubation, the culture was centrifuged at 5000 × g for 20 min. at 4 °C. Cells were resuspended in 50 mL of lysis buffer[11] and placed in an ice bath. The cells were lysed by sonication using 5 pulses of 20 s with 30 s intervals of cooling.

The cell suspension was centrifuged at 9000×g for 1 h and the supernatant was transferred to a clean falcon tube.

The recombinant proteins were purified by Ni affinity chromatography.

4.3.3 Nickel purification for 6X-His tagged proteins

For our lab purposes, proteins are generally expressed using pET-vectors (Novagen). We use Ni-NTA Superflow (Qiagen) for the resin, and the amount of this resin to the lysate containing the tagged protein should be determined empirically.

The lysate from the protein production was incubated with nickel resin, Ni-NTA superflow, for 1-4 hours at 4 °C with gentle rocking/shaking. After the incubation, the mixture was poured into a column and the flow-through was collected and identified. The column was washed with 10× resin volume using elution buffer[12] 0 mM imidazole and the fraction was collected and labeled 0 mM. Then the column was wash/eluted with increasing concentration of imidazole (10× 20 mM, 5× 100 mM twice and 5× 250 mM twice). The fractions were kept in ice to prevent protein denaturation and the fractions were analyzed by SDS-PAGE.

---

[11] Lysis buffer = 100 mM sodium phosphate, 300 mM NaCl, 10% (v/v) glycerol, 1 mg/mL lysozyme, 1 µg/mL pepstatin A, 1—2 µg/mL leupeptin, pH 8.0

[12] Solutions: 100 mM Tris, 300 mM NaCl, X mM imidazole (X = 0, 20, 100, 250), pH 7.4

Adjust to pH 7.4 after the addition of imidazole.

4.3.4. Proton NMR of a mixture of GlcNAc: ManNAc 10:1 ratio.

In a 1.5 mL Eppendorf tube, 1000 µL reaction was set up under the following conditions: TRIS buffer pH 8 50 mM, GlcNAc 10 mM, ManNAc 1 mM, sodium chloride 10 mM, ATP 12 mM, NagK 0.85 µM and deuterium oxide 10%. The sample was analyzed in a 300 MHz NMR with water peak suppression protocol.

4.3.5. Proton NMR of a mixture of GlcNAc: ManNAc 100:1 ratio.

In a 1.5 mL Eppendorf tube, 1000 µL reaction was set up under the following conditions: TRIS buffer pH 8 50 mM, GlcNAc 10 mM, ManNAc 0.1 mM, sodium chloride 10 mM, ATP 12 mM, NagK 0.85 µM and deuterium oxide 10%. The sample was analyzed in a 300 MHz NMR with water peak suppression protocol.

4.3.6. Proton NMR of a mixture of GlcNAc: ManNAc 1:10 ratio.

In a 1.5 mL Eppendorf tube, 1000 µL reaction was set up under the following conditions: TRIS buffer pH 8 50 mM, GlcNAc 1 mM, ManNAc 10 mM, sodium chloride 10 mM, ATP 12 mM, NagK 0.85 µM and deuterium oxide 10%. The sample was analyzed in a 300 MHz NMR with water peak suppression protocol.

4.3.7. Proton NMR of a mixture of GlcNAc: ManNAc 1:100 ratio.

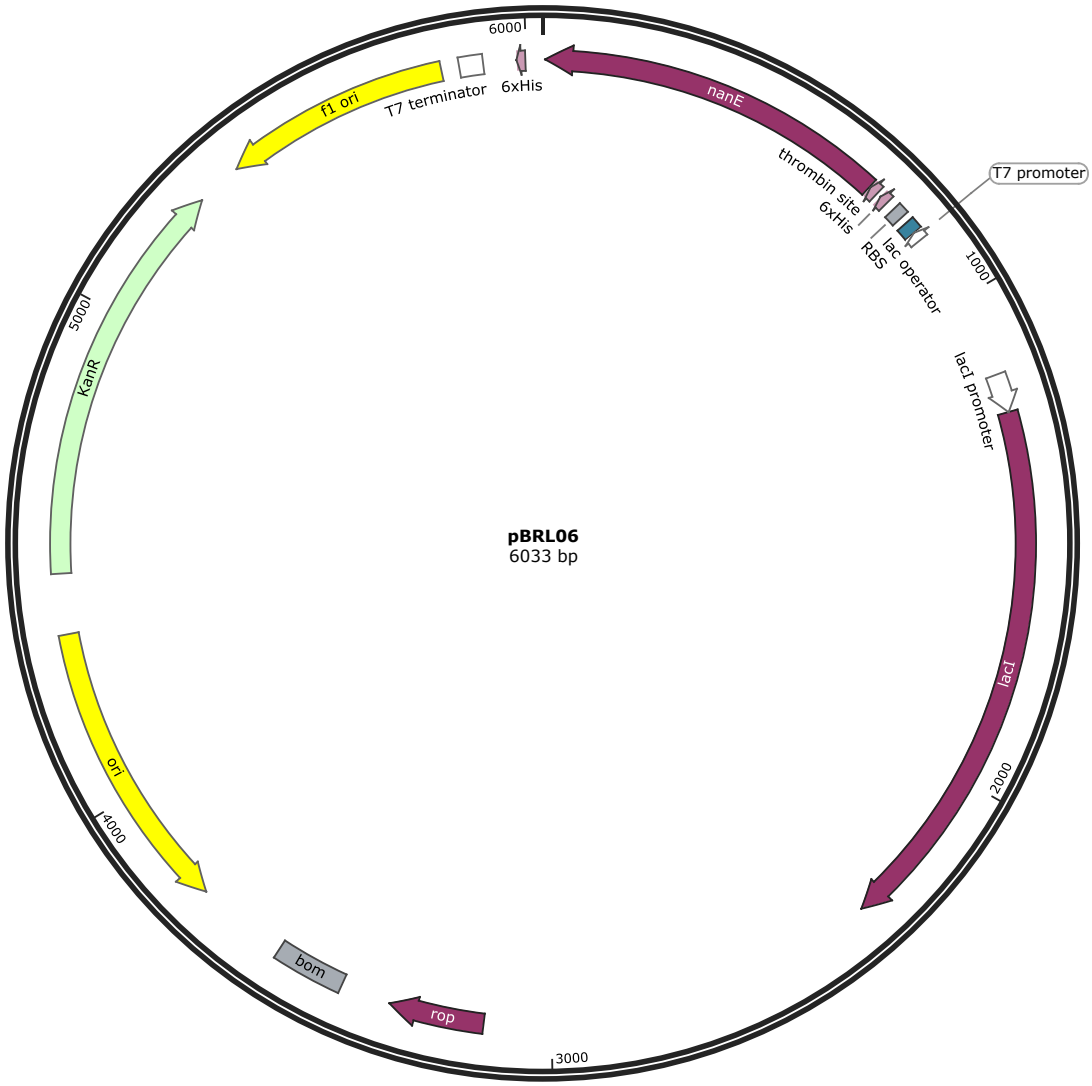
In a 1.5 mL Eppendorf tube, 1000 µL reaction was set up under the following conditions: TRIS buffer pH 8 50 mM, GlcNAc 0.1 mM, ManNAc 10 mM, sodium chloride

10 mM, ATP 12 mM, NagK 0.85 μM and deuterium oxide 10%. The sample was analyzed in a 300 MHz NMR with water peak suppression protocol.

4.3.8. NanE epimerization reaction

In a 1.5 mL Eppendorf tube, 1000 μL reaction was set up under the following conditions: TRIS buffer pH 8 50 mM, sugar substrate 10 mM, sodium chloride 10 mM, ATP 12 mM, NanE 27.5 μM, NanK 0.85 μM and deuterium oxide 10%. The reaction was incubated at 37 °C for four hours and then analyzed in a 300 MHz NMR with water peak suppression protocol.

Sequence: pBRL06dna.dna (Circular / 6033 bp)
Features: 15 visible, 15 total



Printed from SnapGene® Viewer: Jan 30, 2017 4:03 PM Page 1

Plasmid map of pBRL06, *nanE* in pET28b

Chapter 5. In vivo production of sialic acid analogs

5.1 Introduction

Previous chapters have stated the importance of nonulosonic acids for mammal physiology and their wide diversity of roles and functions in the cell as well as how bacteria have evolved to take advantage of the sialic acid present on the cell wall to help them colonize and infect their host.

One specific case of this kind of interaction is the influenza virus. This virus is responsible for the acute respiratory disease commonly known as flu. Every year between 300,000 and 500,000 people die worldwide due to the flu[1–3]. The most effective way to attack influenza is through vaccination. The two mayor glycoproteins on the surface of the virus are hemagglutinin and neuraminidase, which form the main targets for an antibody treatment[2]. Several vaccines have been effectively developed such as split vaccines that are vaccines that contain antigens for both proteins, hemagglutinin and neuraminidase[1].

Another approach against the influenza virus is the use of antiviral drugs. Several of these drugs act on the neuraminidase. This enzyme plays an important role in the life cycle of the virus.

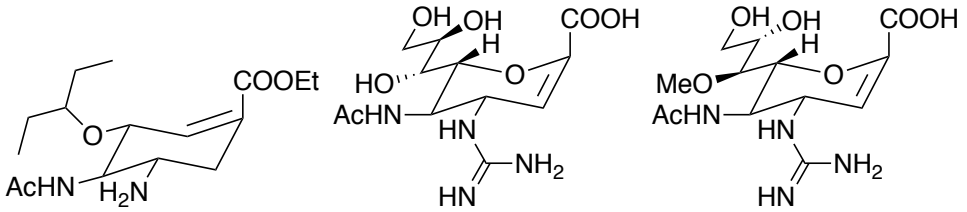
Neuraminidase (NA, EC 3.2.1.18), also called sialidase, is an enzyme that can hydrolyzes α-ketosidic bonds to sialic acids. This enzyme is found in several pathogenic agents such as *Trypanosoma cruzi*, *Clostridium perfringens*, *Streptococcus pneumoniae*, *Vibrio cholerae*, mumps virus, parainfluenza virus, influenza virus among others[4]. It has been suggested that this enzyme helps the pathogen to move across and penetrate sialic acid rich tissues such as the mucosal tissue what makes respiratory track vulnerable[1,5]. NA

is involved in the releasing process of the newly virion particles from the infected cells[1,5]. This allows the new viruses travel and infect healthy cells.

On the other hand, hemagglutinin helps the virion particles get attach to the cell via the recognition of the sialic acid cap on the plasma membrane. This phenomenon triggers the endocytosis process what makes the incorporation of the virus into the cell[1].

As we can see, these two enzymes play a paramount role in the influenza virus life cycle and they share an important characteristic, the affinity for sialic acid. This feature makes sialic acid and important and interesting compound to work on. The modification of the sialic acid moiety can lead to the development of inhibitors of these enzymes making the production of sialic acid analog an important research field.

Currently there are a number of different antiviral drugs on the market that target NA. figure 5.1 shows three of the NA inhibitors. All of these neuraminidase inhibitors are sialic acid analogs. Oseltamivir was approved for use in United States in 1999, it was the first oral administrated NA inhibitor[6]; laninamivir is being researched for the treatment and prophylaxis of Influenzavirus A and Influenzavirus B3 and it is currently in phase III clinical trials[7].



Oseltamivir Zanamivir Laninamivir

Figure 5. 1. Neuraminidase inhibitors marketed as antiviral drugs.

The development of new sialic acid analogs is an important tool against influenza virus and opens a field of research where different modifications on the several nonulosonic

scaffolds could lead to a potent drug not only against viruses but pathogens in general. For this reason, we tested our findings in previous in vivo experiments to see if we could develop a bacterial strain capable to biosynthesize acyl neuraminic acids.

To achieve this goal, we transformed different *E. coli* strains previously developed in our lab to overproduce two enzymes, one involved in the synthesis of sialic acid (NeuB) and the other capable to introduce the *N*-acyl-D-glucosamine into the biosynthesis pathway. These experiments gave us the opportunity to test the stability of *N*-acylglucosamines to NagA and to produce *N*-acyl sialic acid analogs.

5.2 Results and Discussion

Overproduction of nonulosonic acids has become an important tool in drug development due to the fact that this compound family is integral to with multiple biological processes, from brain development in infants to virulence and pathogenicity of bacteria. Several studies show that sialic acids analogues can interfere with the uptake of host's sialic acid from pathogens. For this reason, it is crucial to be able to produce non-natural sialic acid analogues [8].

We developed an experiment that permits us confirm the activity of NanE and NagA in *E. coli*. For our experiment, we transformed two, previously engineered[9], *E. coli* strains, BRL02 (*nanA*$^{-}$ *nanT*$^{+}$) and BRL04 (*nanA*$^{-}$ *nanT*$^{-}$ *nagA*$^{-}$), with the plasmid pBRL30 (*nanE* and *neuB* in a pMGX-A backbone).

In *E. coli* BRL02, the *nanT* gene encoding a membrane transporter protein involved in the uptake of sialic acid into the cell[10], has been knocked out decreasing the possibility of catabolizing the sialic acid produced by the cell and increase the extracellular

concentration of the sialic acid. In addition, the absence of the sialic acid aldolase (encoded by *nanA*) blocks the first step in the catabolic pathway of sialic acid. These features make this strain a suitable option to test the production of the sialic acid analogs. *E. coli* strain BRL04 lacks *nagA* as well as *nanA* and *nanT*. This strain, in combination with BRL02 are ideal to test the impact of deacetylase activity, which can degrade the substrate, on product formation.

Both strains were transformed with pBRL30, a plasmid previously designed in the laboratory[9] that carries two of the enzymes that our hypothesis relies on, N-acetylmannosamine-6-phosphate epimerase (NanE) and N-acetylneuraminic acid synthetase (NeuB). With the overexpression of these two enzymes, we could test the conversion of the *N*-acyl-D-glucosamine into their corresponding *N*-acylneuraminic acids.

Single colonies of the transformed bacteria were cultured in minimum media with 0.3% of glycerol and 0.15% of N-acylglucosamine. After 48 hours of incubation the supernatant was analyzed to determine the presence of the correspondent nonulosonic acid. The use of glycerol allows us to give the bacteria a carbon source that can be used to grow.

It is important to point that the substrates used in this experiment are not phosphorylated, although, all the intermediates in the sialic acid metabolism are phosphorylated at carbon six. In this case, we assume that the phosphotransferase system (PTS) present in *E. coli* would phosphorylate the compounds during the uptake process. This hypothesis is backed on by the results of our in vivo experiment in chapter 2. In that experiment we grew different *E. coli* strains in a culture medium containing different N-acylglucosamines as only carbon source. However passively transported N-acylglucosamines that enters the cell unphosphorylated may be processed by NanE as there

is evidence that *Bacteroides fragilis,* which lacks NanK but possesses a functional epimerase (NanE), can catalyze the epimerization of non-phosphorylated substrates[11,12].

Sialic acid and its analogs are hard to detect due to the lack of chromophores, thus, we derivatized the samples with 1,2-diamino-4,5-methylenedioxybenzene (DMB)[13–15] (fig 5.2). Once the samples were prepared, they were mixed with the DMB working solution and placed in a heating block at 50 °C for 2.5 hours. After the derivatization procedure, the samples were analyzed by HPLC-DAD. Derivatized samples were analyzed using a 1260 Infinity High Performance Liquid Chromatography system (Agilent Technologies) with a Prontosil C18, 5 µm, 125 mm x 4 mm column; flow rate 1.00 mL min^{-1}, λ= 350 nm; mobile phase 84% water, 9% acetonitrile, 7% methanol constant for 12 minutes, followed by a linear gradient of acetonitrile from 9% to 100% over 1 minute until completion.

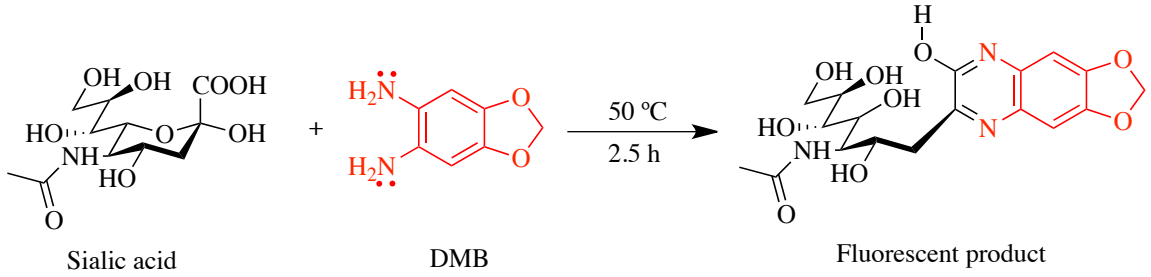


Figure 5. 2. DMB derivatization reaction[13–15]

Sialic acid was used as standard to determine the retention time in the chromatogram. In figure 5.3 we can see a very clear peak corresponding to the DMB derivatized sialic acid standard with a retention time of 4.2 minutes. Due to the different acyl groups we used in the substrates, we expected that the retention time of the sialic acid analogs would change; considering the fact that the acyl groups were less polar than the acetyl group found in sialic acid.

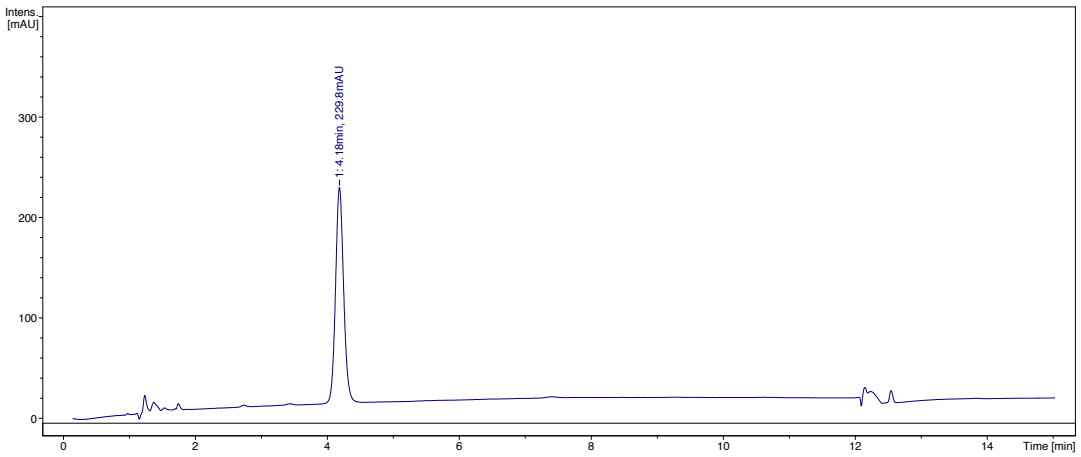


Figure 5. 3. HPLC chromatogram of DMB derivatized sialic acid standard; 1000 µg/mL.

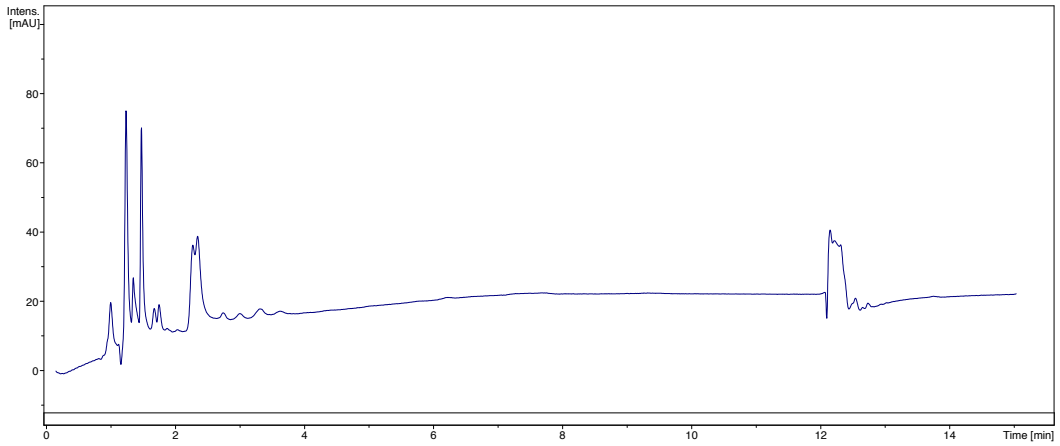


Figure 5. 4. HPLC chromatogram of DMB derivatized sample from feeding experiment on BRL02 transformed with pBRL30 using GlcNAc as substrate.

Figure 5.4 shows the result of our first experiment. This experiment served as a positive control as GlcNAc is the natural precursor for the synthesis of sialic acid. We expected with this test to find a peak at 4 minutes, however, as it is shown, we did not detect any production of sialic acid. This negative result suggests that our NanE based production system was non-functional. We hypothesized that NagA may be hydrolyzing any GlcNAc taken up by the cell converting it back into GlcN-6-P which can enter central

metabolism as Fru-6-P. This could potentially lead to low levels of GlcNAc-6-P in the cells, preventing appreciable accumulation of the required ManNAc derivative. Though we were also aware that the heavy metabolic demands on these highly engineered cell lines frequently lead to loss of genetic elements from the plasmids or mutation of the promoters leading to decreased protein production. Thus interpretation of the direct cause of this negative result was challenging.

When we examined the results of the rest of our tests (appendix#), we also saw that none of the substrates produced sialic acid analogs. It is important to emphasize that we tested only substrates with small acyl groups such as GlcNPr and GlcNnBu and based on the results discussed in chapter 2, these may be substrates of NagA. Thus it is possible the *N*-acyl groups of theses substrates were hydrolyzed during the experiment, preventing analog production.

Although, the experiment provided a negative result, they suggested a path forward, where a similar system lacking NagA could be generated. David Schlachter, another researcher in the lab, transformed BRL04 (strain that has *nagA* knocked out) with pDS3, a plasmid carrying nanE and neuB on pMGX-HisA; this plasmid encodes an N-terminal hexa-histidine tag thus NanE and NeuB would be produced as His-tagged proteins[16]. This feature allowed confirmation of protein expression byWestern blot.

To validate this approach, BRL02, BRL04 and BL21 were transformed with pDS3 plasmid. Cells were grown for 48 h in minimal media and given one feeding 0.5% GlcNAc upon induction with IPTG[9,16]. Samples were taken at 24 and 48 h and the lysates were analyzed by western blot using Anti-His antibodies. Figure 5.5 shows that NanE and NeuB were over expressed in all cell strains what suggests that our previous experiment results

are not due to the lack of enzyme expression, but it could be possible that NagA degraded the substrates tested.

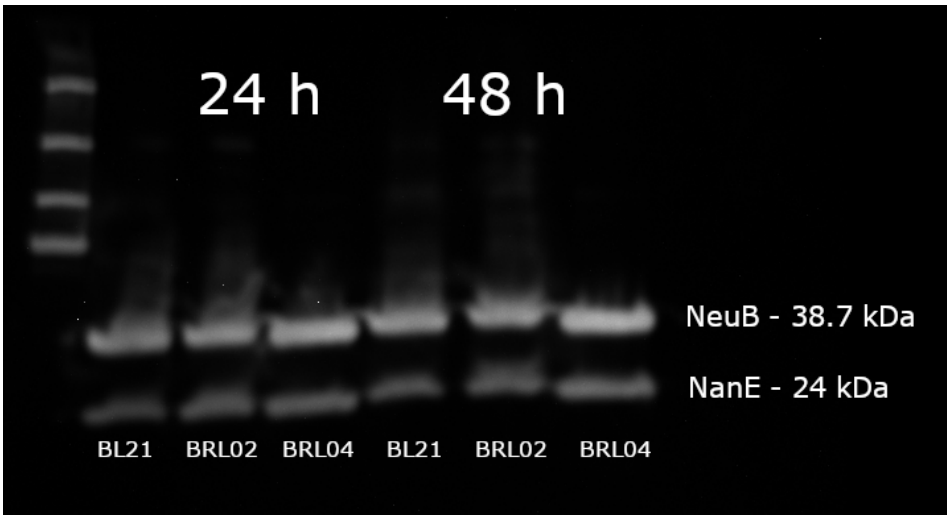


Figure 5. 5. Western blot from BL21, BRL02, and BRL04 lysates. Anti-His antibodies allow to prove that the enzymes are highly expressed in all strains 24 h after induction.

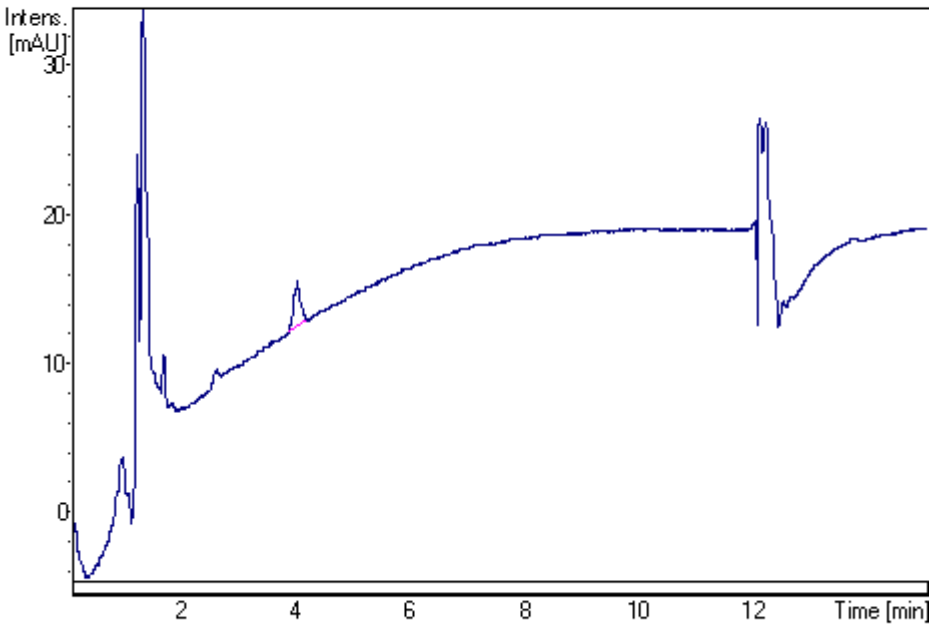


Figure 5. 6. HPLC-DAD chromatogram of DMB derivatized culture broth form feeding experiment with BRL04/pDS3 using GlcNAc as substrate.

After corroborating that NanE and NeuB were expressed in the bacterial systems, the next step was to induce the production of sialic acid to validate the experiment. In this case, BRL04 was transformed with pDS3 and the feeding experiment was performed. Figure 5.6 shows the chromatogram of the DMB derivatized culture broth from the GlcNAc substrate. As seen in figure 5.3, the peak at 4 minutes corresponds to sialic acid what proves that the production of this compound can be achieved in this strain. For this experiment, sialic acid was produced at 29 mg/L.

After corroborating that the *E. coli* BRL04 / pDS3 is capable to produce sialic acid, we tested the production system with different acylglucosamines. Figure 5.7 shows the results obtained from the feeding experiment with GlcNPr. We can see in *E. coli* BRL04 / pDS3 chromatogram a small peak at 7.2 min. This is evidence of the production of the sialic acid analog at 30 mg/L. *E. coli* BL21 / pDS3 was not able to produce any nonulosonic acid what confirms the possibility of substrate degradation by NagA.

Other substrates were tested, and the results are summarized in table 5.1. As we can see, the only compounds that were transformed into their correspondent nonulosonic acids were GlcNPr and GlcNnBu. These results confirm what was found by Lungren *et al*. where they were capable to produce N-butanoyl sialic acid using BRL02 as production system[9]. From these results we can infer that the presence of NagA in the strains BL21 and BRL02 plays an important role for the analog production.

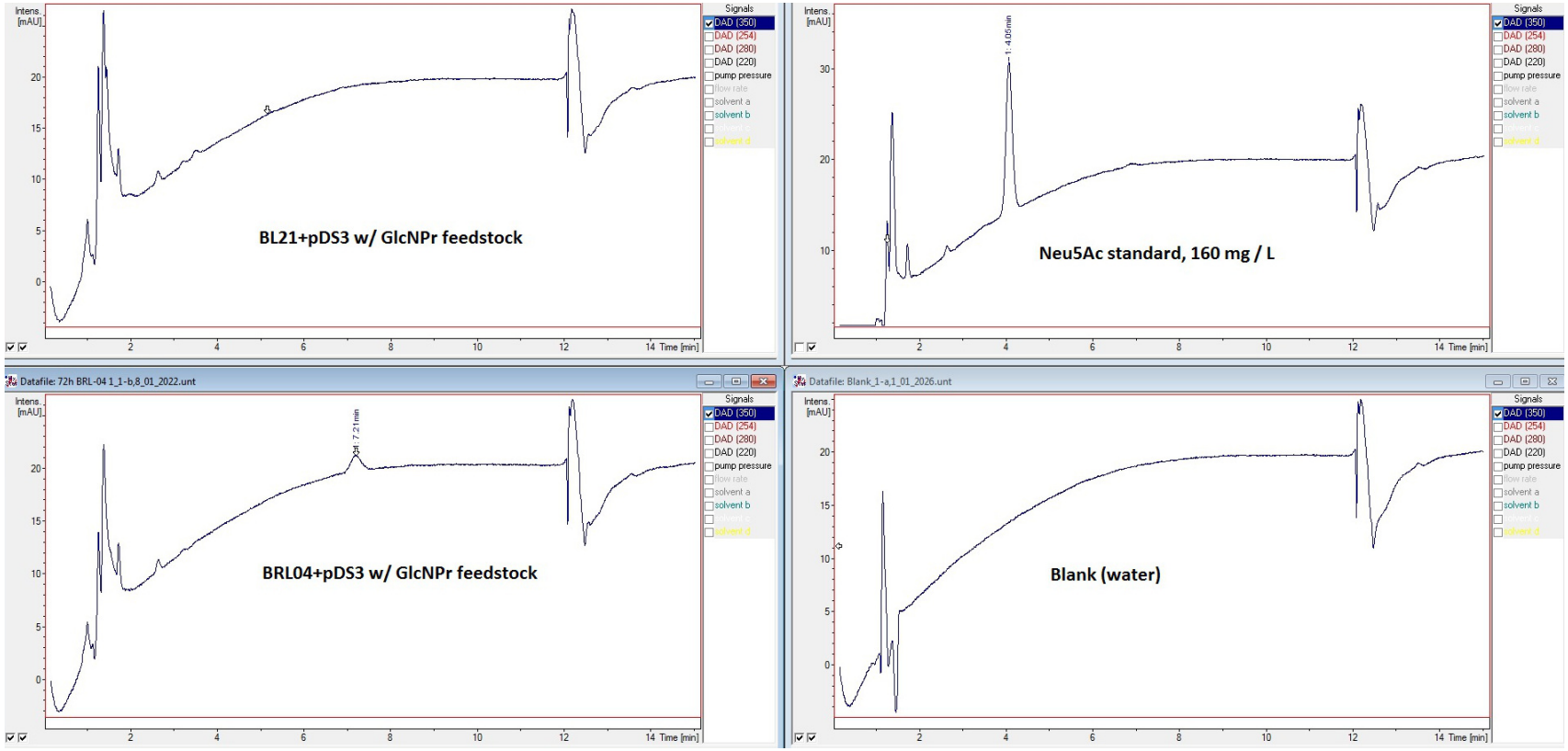


Figure 5. 7. HPLC-UV chromatograms of GlcNPr feeding experiment with BRL04 and BL21, Neu5Ac standard and water as blank.

Table 5. 1. Results of feeding experiment on different *E. coli* strains.

Acylglucosamine	Culture	R.T. (min)	Integration (mAU)	Concentration (mg/L)
GlcNAc	BL21	--	--	--
	BRL02	--	--	--
	BRL04	4.05	121.93	29
GlcNPr	BL21	--	--	--
	BRL02	--	--	--
	BRL04	7.02	123.84	30
GlcNnBu	BL21	--	--	--
	BRL02	--	--	--
	BRL04	7.09	31.33	4
GlcNiBu	BL21	--	--	--
	BRL02	--	--	--
	BRL04	--	--	--
GlcNBz	BL21	--	--	--
	BRL02	--	--	--
	BRL04	--	--	--

Based on the results found in chapter 2, GlcNBz should be converted into the corresponding sialic acid analog, however, we could not detect it. This finding raises a number of questions. For example, was the substrate degraded by the catabolic pathway? Is the epimerase able to convert the glucosamine analog into the mannosamine counterpart? Can the PTS import and phosphorylate the sugar?

This section demonstrated that the production of non-natural sialic acid analogs using a bacterial model is possible but challenging. We were able to produce one analog with a concentration close to the 5-AcNeu acid and another one at a concentration 5 times lower. While experiment confirms that sialic acid analogs can be produced, more questions

arise from the results. In order to be able to reliably obtain N-acyl nonulosonic acids, more research is needed. In particular we should investigate how the rest of the enzymes of the pathway work with the sugar analogs.

However, one major conclusion that can be made is that NagA interferes with the production of the analogs. Although, we found that the deacetylase does not work well with large and bulky acyl groups, we did not test this with other enzymes involved in the metabolic process.

References

1. Ryu, W.-S. Influenza viruses. in *Viral Infections of Humans: Epidemiology and Control* 195–211 (Academic Press, 2017). doi:10.1016/B978-0-12-800838-6.00015-1
2. McAuley, J. L., Gilbertson, B. P., Trifkovic, S., Brown, L. E. & McKimm-Breschkin, J. L. Influenza virus neuraminidase structure and functions. *Front. Microbiol.* **10**, 26–32 (2019).
3. Yamashita, M. *et al.* CS-8958, a prodrug of the new neuraminidase inhibitor R-125489, shows long-acting anti-influenza virus activity. *Antimicrob. Agents Chemother.* **53**, 186–192 (2009).
4. Vavricka, C. J. *et al.* Synthesis of Sulfo-Sialic Acid Analogues: Potent Neuraminidase Inhibitors in Regards to Anomeric Functionality. *Sci. Rep.* **7**, 4–11 (2017).
5. Colman, P. M. Influenza virus neuraminidase: Structure, antibodies, and inhibitors. *Protein Sci.* **3**, 1687–1696 (1994).
6. Agrawal, R., Rewatkar, P. V., Kokil, G. R., Verma, A. & Kalra, A. Oseltamivir: A First Line Defense Against Swine Flu. *Med. Chem. (Los. Angeles).* **6**, 247–251 (2010).
7. Hayden, F. Developing New Antiviral Agents for Influenza Treatment: What Does the Future Hold? *Clin. Infect. Dis.* **48**, S3–S13 (2009).
8. Thorlund, K., Awad, T., Boivin, G. & Thabane, L. Systematic review of influenza resistance to the neuraminidase inhibitors. *BMC Infect. Dis.* **11**, 134 (2011).
9. Lundgren, B. R. & Boddy, C. N. Sialic acid and N-acyl sialic acid analog production

by fermentation of metabolically and genetically engineered *Escherichia coli* †. *Org. Biomol. Chem.* 1903–1909 (2007). doi:10.1039/b703519e

10. Vimr, E. & Kalivoda, K. Diversity of microbial sialic acid metabolism. *Microbiol. Mol. Biol. Rev.* **68**, 132–153 (2004).

11. Almagro-Moreno, S. & Boyd, E. F. Insights into the evolution of sialic acid catabolism among bacteria. *BMC Evol. Biol.* **9**, 118 (2009).

12. Brigham, C. *et al.* Sialic acid (N-acetyl neuraminic acid) utilization by *Bacteroides fragilis* requires a novel N-acetyl mannosamine epimerase. *J. Bacteriol.* **191**, 3629–3638 (2009).

13. Lamari, F. N., Kuhn, R. & Karamanos, N. K. Derivatization of carbohydrates for chromatographic, electrophoretic and mass spectrometric structure analysis. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **793**, 15–36 (2003).

14. Klein, A. *et al.* New sialic acids from biological sources identified by a comprehensive and sensitive approach: liquid chromatography-electrospray ionization-mass spectrometry (LC-ESI-MS) of SIA quinoxalinones. *Glycobiology* **7**, 421–32 (1997).

15. Sato, C., Inoue, S., Matsuda, T. & Kitajima, K. Fluorescent-assisted detection of oligosialyl units in glycoconjugates. *Anal. Biochem.* **266**, 102–9 (1999).

16. Hassan, M. I., McSorley, F. R., Hotta, K. & Boddy, C. N. Inducible T7 RNA Polymerase-mediated Multigene Expression System, pMGX. *J. Vis. Exp.* 1–11 (2017). doi:10.3791/55187

17. Stanton, P. G. *et al.* Application of a sensitive HPLC-based fluorometric assay to determine the sialic acid content of human gonadotropin isoforms. *J. Biochem.*

Biophys. Methods **30**, 37–48 (1995).

5.3 Experimental Section

5.3.1 Plasmid Mini-Preps via P1, P2, & P3

1 mL of culture was centrifuged at 16,000 × g in a 1.5 ml eppendorf tube to pellet cells. Supernatant was discarded and the procedure was repeated 4 more times to collect the cells contained in 5 mL of culture. The pellet was resuspended in 250 µl of P1[13] (250 µl per 5 mL of culture). The cell suspension was mixed thoroughly until no cell clumps were visible in suspension. 250 µl of P2[14] were added (250 µl per 5 mL of culture) and mixed gently by inverting the tube 4-6 times to mix[15]. After mixing, 350 µl of P3[16] (350 µl per 5 mL of culture) were added and the tube was inverted immediately, but gently, 4-6 times to mix.

Tubes were centrifuged at 16,000 × g for 10 minutes to pellet debris and 700 µL of the supernatant were transferred to a tube containing 700 µL of cold isopropanol and put on ice for 15 min. After the incubation time the tubes were centrifuged at 16,000 × g for 30 minutes to pellet plasmid DNA and the supernatant was discarded.

500 µL of chilled 70% ethanol were added to the plasmid DNA pellet and centrifuged at 16,000 × g for 10 minutes. The supernatant was discarded, and the pellet was dried via air or speed-vac for 20 minutes. The plasmid DNA pellet was resuspended in elution buffer[17].

---

[13] P1: 6.1 g Tris, 3.7 g EDTA-2H2O pH 8.0 w/ HCl/1 liter, add 100 µg/ml RNAse A as needed, usually 10 mg RNase in 100 ml batches, store 4 degrees.
[14] P2: 8.0 g NaOH in 900 ml H2O plus 100 ml of 10 % SDS/1 liter, store R.T.
[15] Solution should become viscous and slightly clear. DO NOT let lysis proceed for more than 5 minutes.
[16] P3: 294 g KAcetate in 500 ml H2O pH to 5.5 with Acetic Acid (~110 ml), bring to 1 liter, store R.T.
[17] EB: 10 mM Tris, pH 8.5

5.3.2 Gene expression in and protein extraction from E. coli

To produce the seed culture one single colony of the transformants was used to inoculate 10 mL of LB broth supplemented with kanamycin. The culture was incubated at 37 °C and 200 rpm overnight. Next day, 400 mL of fresh LB were supplemented with 400 µL of 1000 × kanamycin solution (50 mg/mL) and inoculated with the seed culture. The bacterial culture was incubated at 37 °C with shaking until OD_{600} reaches 0.3-0.6.
When the proper optical density was reached, IPTG 1M was added to a final concentration of 0.5 mM and the culture was placed in the incubator at 30 °C with shaking.

After 12 to 18 hours of incubation, the culture was centrifuged at 5000 × g for 20 min. at 4 °C. Cells were resuspended in 50 mL of lysis buffer[18] and placed in an ice bath. The cells were lysed by sonication using 5 pulses of 20 s with 30 s intervals of cooling.
The cell suspension was centrifuged at 9000×g for 1 h and the supernatant was transferred to a clean falcon tube.
The recombinant proteins were purified by Ni affinity chromatography.

5.3.3 Sialic acid analogs production protocol

Single colonies are inoculated into 5 mL of LB supplemented with antibiotic and grow at 37 °C, 200 rpm overnight. The culture was harvested and spun down for 2 minutes at 13000 rpm. The pellet was washed with F1 media and used as inoculum. 5 mL of F1 media is supplemented with antibiotics and 0.3% of glycerol and 0.3% of N-

[18] Lysis buffer = 100 mM sodium phosphate, 300 mM NaCl, 10% (v/v) glycerol, 1 mg/mL lysozyme, 1 µg/mL pepstatin A, 1—2 µg/mL leupeptin, pH 8.0

acylglucosamine. The media is inoculated with the washed pellet and incubated at 37 °C, 200 rpm until reach OD_{600} = 0.5.

Once the absorbance was reached, production of analog is induced with 0.2 mM IPTG and incubated for 24 hours. After 24 hours, 0.3% of glycerol and 0.15% of *N*-acyl-D-glucosamine were added and the incubation is carried on for 24 hours more. After 48 hours of incubation, the culture was centrifuged, and the supernatant was used for the DMB derivatization procedure.

5.3.4 Construction of sialic acid biosynthetic plasmid pDS3

E. coli XL1-Blue Competent Cells were used for cloning, following supplier's protocols for transformations. Briefly, competent cells were thawed on ice and 0.5 – 1 µL vector was added, with incubation on ice for 30 minutes. Heat shock was performed at 42 ºC for 45 seconds, then incubation on ice for 2 minutes. 1 mL preheated LB medium was added, and solution incubated for 1 hour at 37 oC, 225 – 250 rpm. 100 µL were plated on appropriate selection plates and incubated overnight.

Source plasmids containing both nanE and neuB were selected from prior work in the group: pBRL21 contained neuB, and pBRL06 contained nanE, with both plasmids having NdeI and EcoRI sites flanking the respective genes. Both genes were excised by digestion with NdeI and EcoRI and cloned into pMGXAmpHis: a low-copy Amp^R vector under control of an inducible T7 promoter, with a hexa-histidine tag following the EcoRI site [25]. In this way, pDS1 (neuB) and pDS2 (nanE) were produced. To form a bicistronic synthetic inducible operon containing both nanE and neuB, nanE was excised from pDS2 by XbaI and AvrII digestion and inserted into pDS1 at the XbaI site to form pDS3.

Each plasmid was validated by restriction digests to validate successful insert ligation. Plating was performed with appropriate antibiotics for selection of successful ligation products. *E. coli* BL21(DE3) cells were transformed with vectors and cell lysates assayed by Western blotting to validate protein expression.

5.3.5 Sialic acid production in *E. coli* in shake flask culture

BL21, BRL02 and BRL04 cells were transformed with pDS3 vector. Seed cultures from single colonies were grown in LB with appropriate antibiotics at 37 ºC, 200 rpm overnight (18 h). Production cultures in F2 media (consisting of per litre, 12.24 g K_2HPO_4, 6.0 g KH_2PO_4, 4.0 g $(NH_4)_2SO_4$ and 175.5 mg $MgSO_4$) with kanamycin (50 mg/mL) were inoculated with 0.5% seed culture, 0.25% casitone, and 0.5% (v/v) glycerol, then grown at 37 oC, 200 rpm until an OD_{600} of 0.4 – 0.6 was achieved (2 – 4 h).

At this point (t = 0), cultures were inoculated with 0.5% of the appropriate *N*-acyl-D-glucosamine derivative, and protein expression was induced with 0.2 mM isopropyl β-D-1- thiogalactopyranoside (IPTG). Production cultures were incubated at 30 oC, 200 rpm. At t = 24 h, sample aliquots were collected, and antibiotics were again supplemented to the cultures. At t = 48 h sample aliquots were collected, and the growth experiment was stopped. Aliquots were centrifuged at 13,000 rpm for 2 min, and supernatant was removed for analysis. Triplicate cultures of all strains were used as possible depending on the quantity of substrate available.

Some variations of this procedure were attempted, splitting the feeding of 0.5% GlcNAc derivative to two feedings at t = 0 h and t = 24 h, as well as extending growth experiments

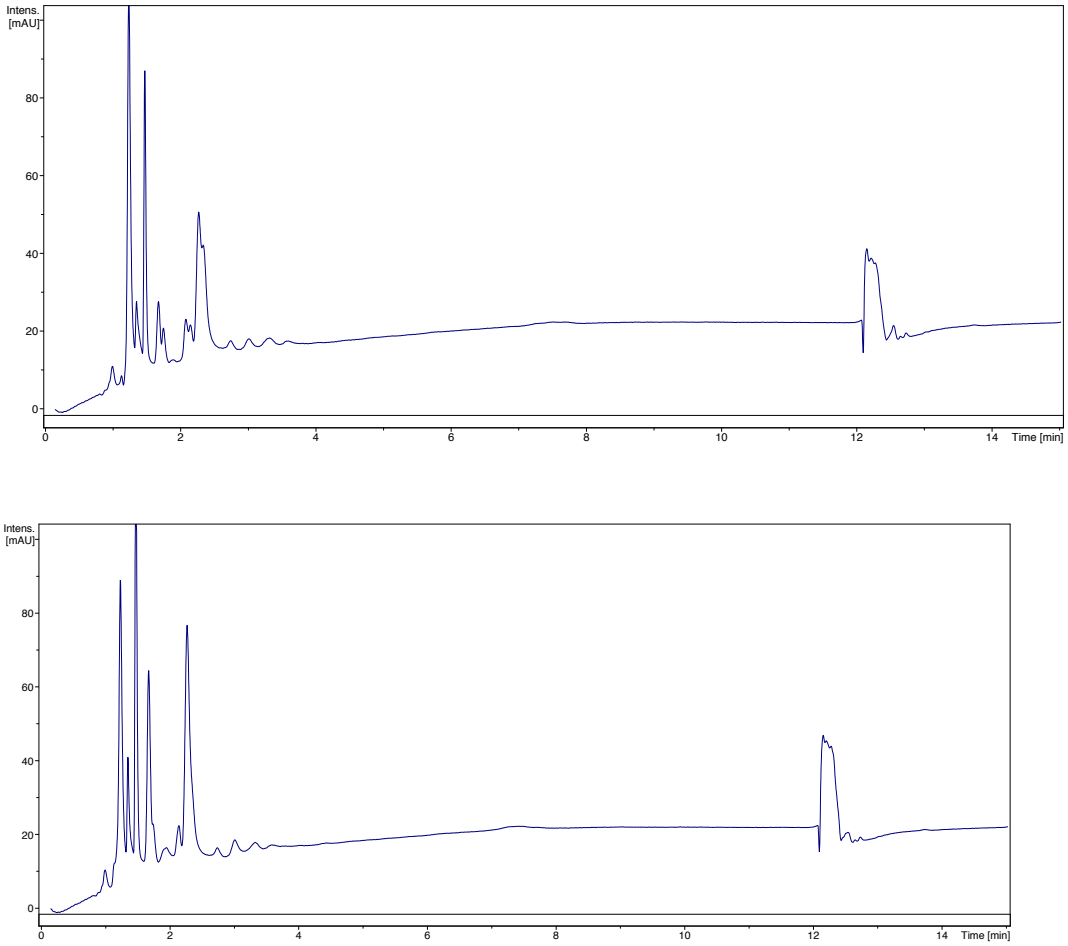
to 72 or 144 h. Aliquots were taken at t = 48 h and no significant differences in titre were found for longer experimental durations.

5.3.6 DMB derivatization protocol (adapted from Stanton *et al.*[17])

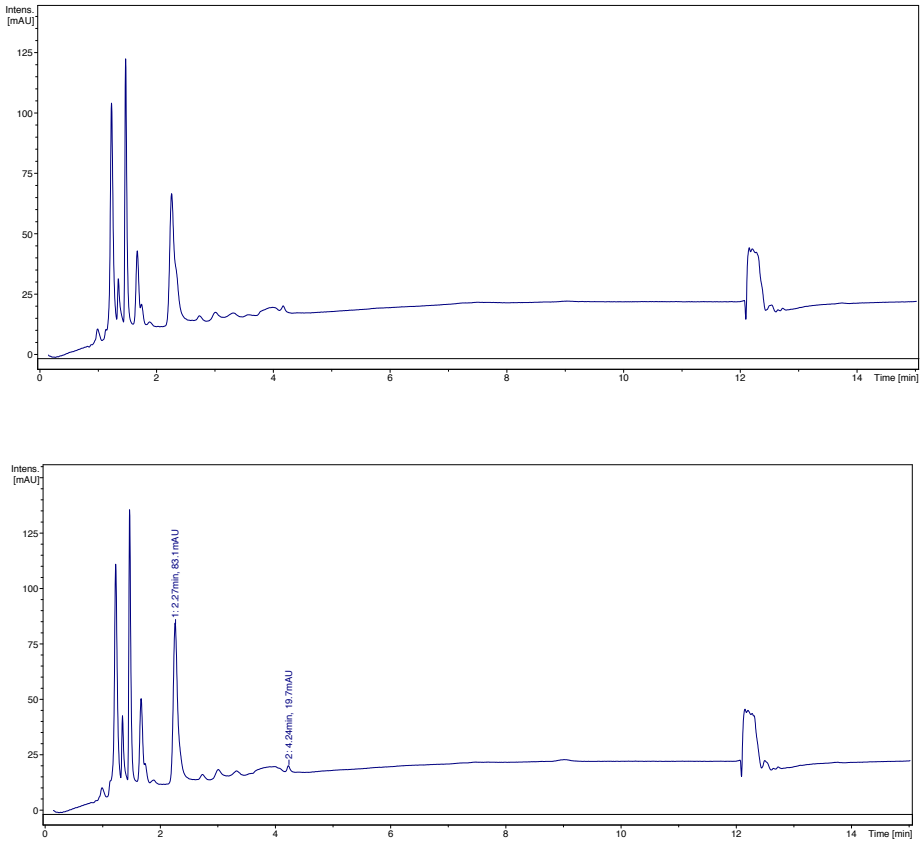
DMB solution working solution was prepared by mixing, in a 1.5 mL centrifuge tube, 858 µL of water, 56 µL acetic acid, 42 µL 2-mercaptoethanol, 1.8 mg sodium hydrosulfite and 2.4 mg DMB (4,5-methylenedioxy-1,2-phenylenediamine dihydrochloride). The tube was wrapped in aluminum foil and stored at -20 °C.

Once the samples were prepared, 20 µL of sample were mixed with 20 µL of DMB working solution in a 1.5 mL centrifuge tube. The tube was wrapped in aluminum foil and placed in a heating block at 50 °C for 2.5 hours. When the reaction time was finished, the samples were analyzed by HPLC. Derivatized samples were analyzed using a 1260 Infinity High Performance Liquid Chromatography system (Agilent Technologies) using a Prontosil C18, 5 µm, 125 mm x 4 mm column. Flow rate: 1.00 mL min^{-1}, λ= 350 nm. Mobile phase: 84% water: 9% acetonitrile: 7% methanol constant for 12 minutes, followed by a linear gradient of acetonitrile from 9% to 100% over 1 minute until completion. A standard curve was generated using known concentrations of an authentic standard of Neu5Ac from 10 – 160 mg L^{-1} (R^2 = 0.9996). The area under the curve of the peak corresponding to derivatized sialic acid was used for quantification.

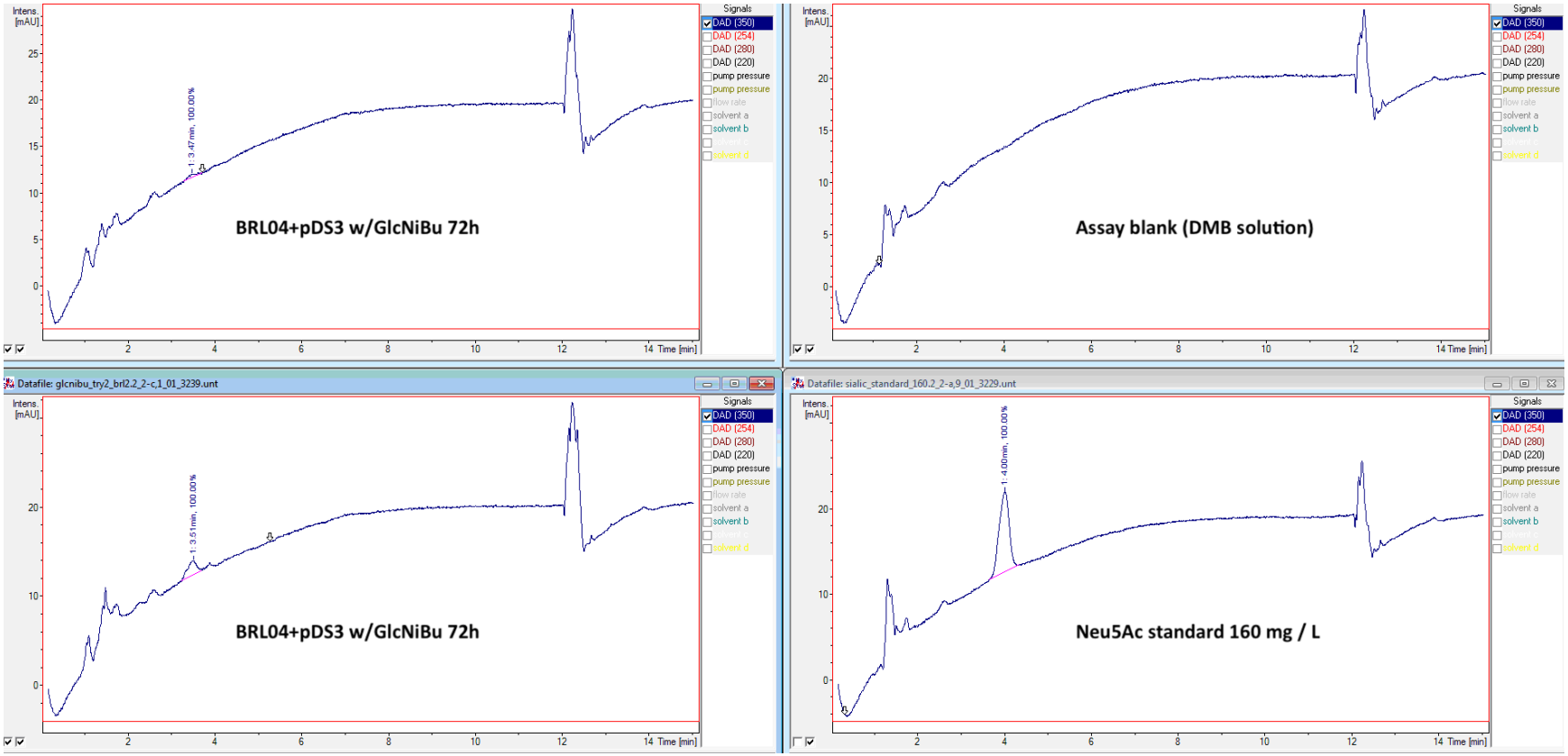
5.4 Anexx



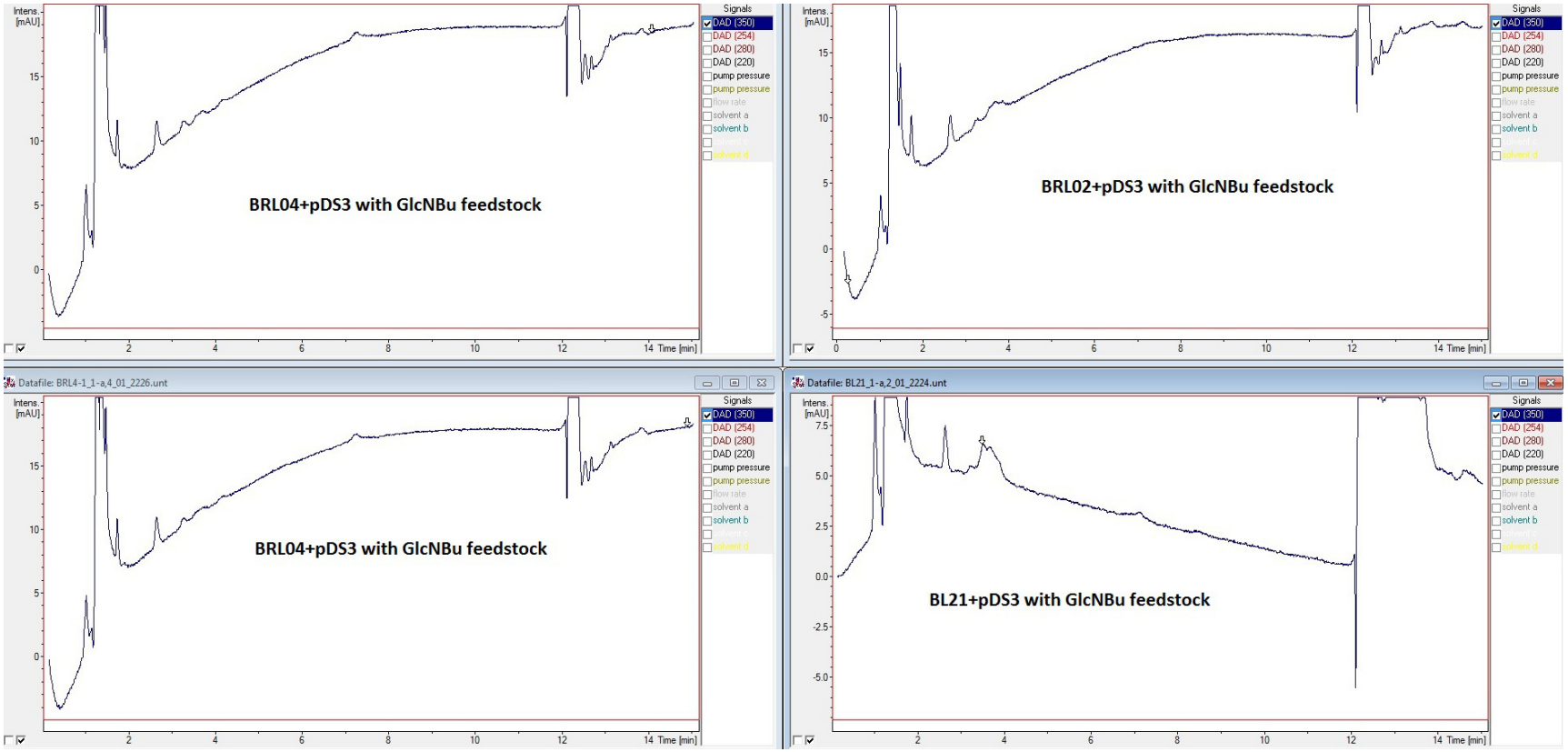
HPLC chromatogram of DMB derivatized sample from feeding experiment with GlcNPr. Top, BRL02 transformed with pBRL30; bottom, BRL02 transformed with pKH22.



HPLC chromatogram of DMB derivatized sample from feeding experiment with GlcNnBu. Top, BRL02 transformed with pBRL30; bottom, BRL02 transformed with pKH22.



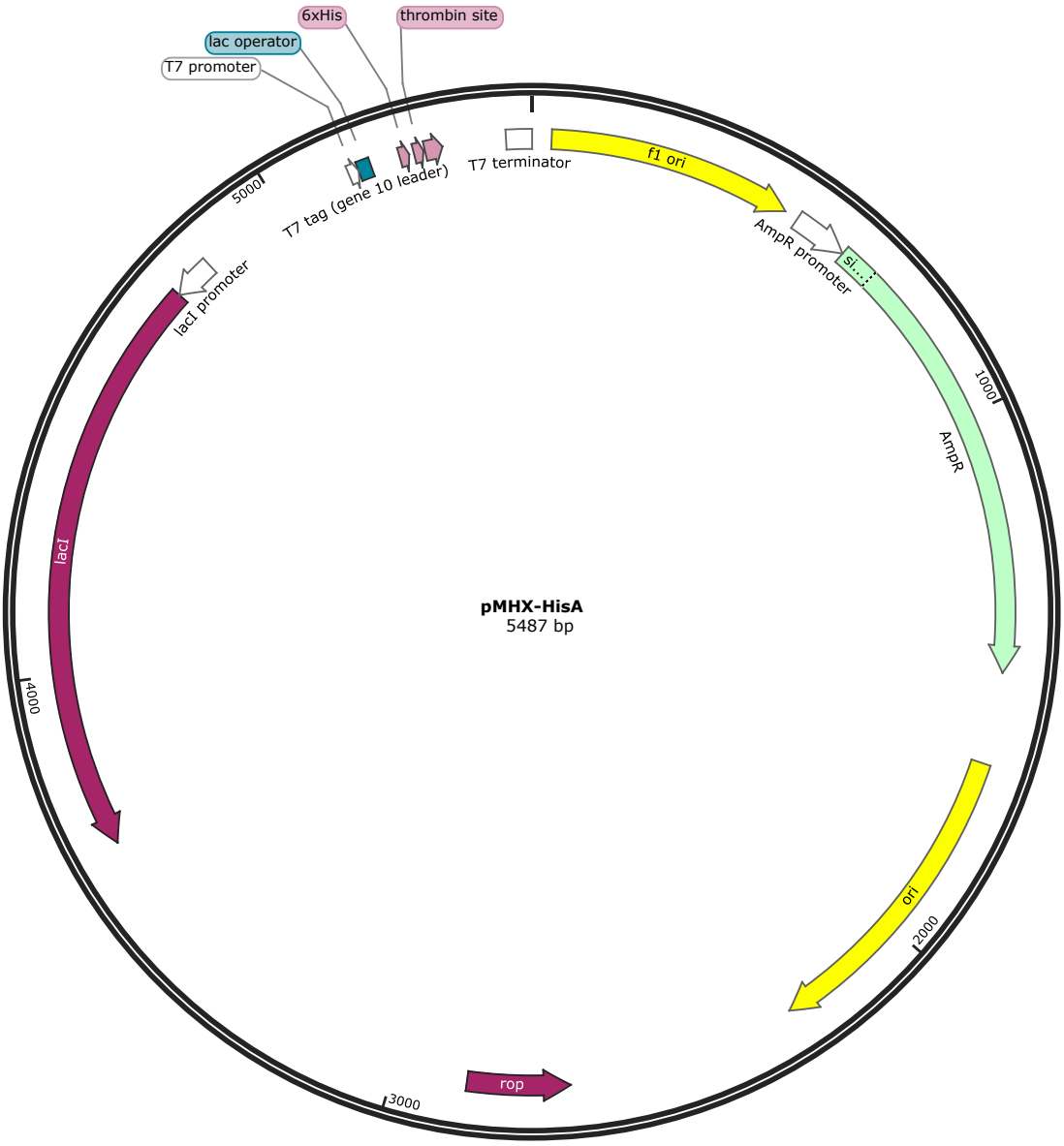
HPLC-UV chromatograms of GlcNiBu feeding experiment with BRL04, Neu5Ac standard and water as blank.



HPLC-UV chromatograms of GlcNnBu feeding experiment with BRL04 and BL21, Neu5Ac standard and water as blank.

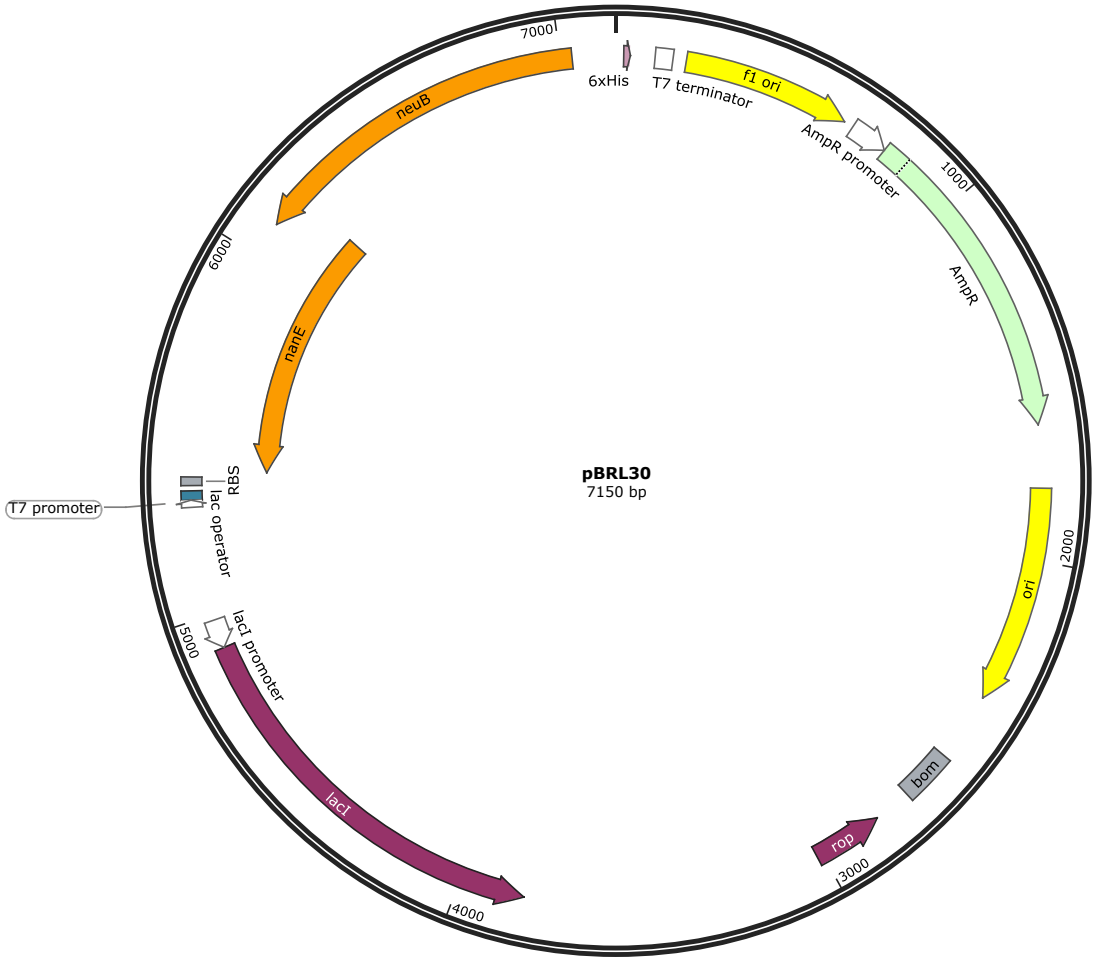
Plasmids maps

Sequence: pMHX-HisA.dna (Circular / 5487 bp)
Features: 13 total



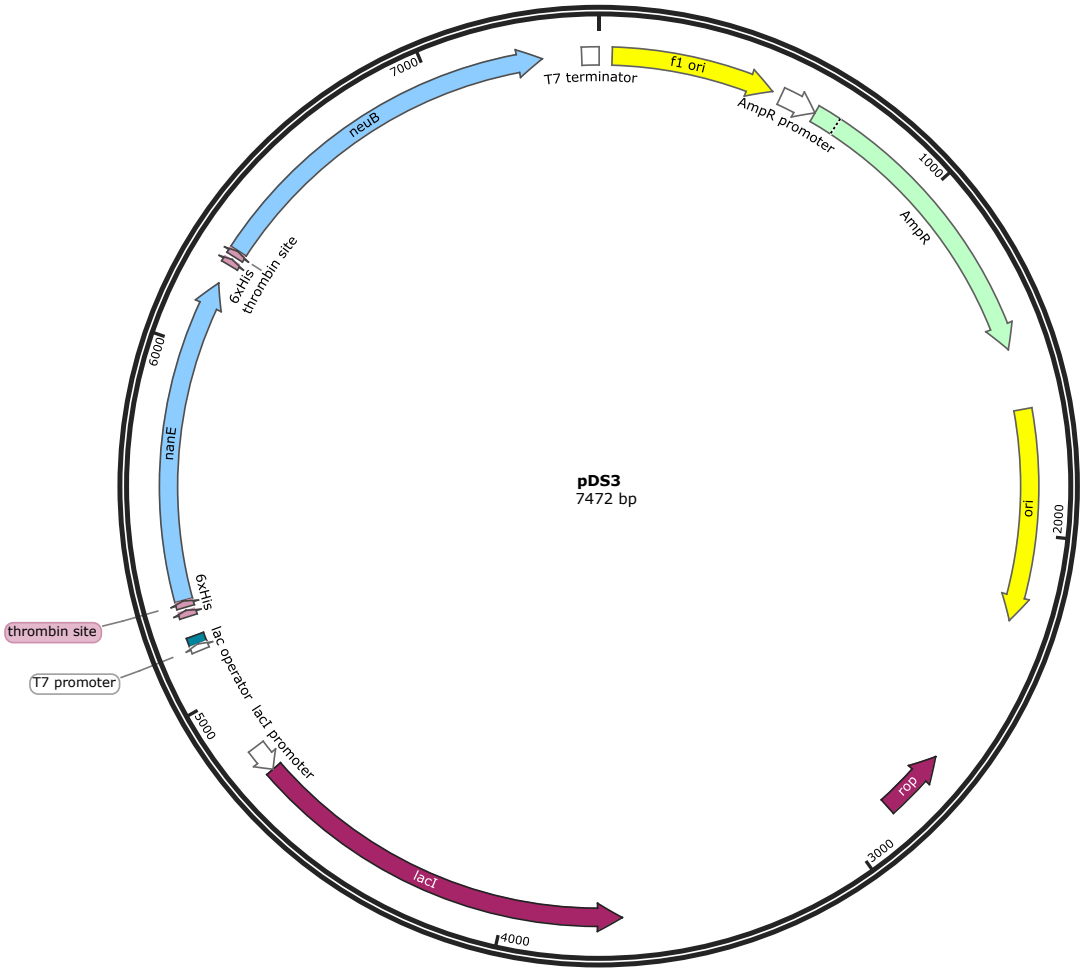
Plasmid map of pMHX-HisA

Sequence: pBRL30.dna (Circular / 7150 bp)
Features: 15 visible, 15 total



Plasmid map of pBRL30, *nanE*, *neuB* in pKH22

Sequence: pDS3.dna (Circular / 7472 bp)
Features: 16 total



Printed from SnapGene®: Jul 22, 2019 2:27 PM Page 1

Plasmid map of pDS3, *nanE, neuB* in pMHX-HisA

Chapter 6. Concluding remarks and future directions

Sialic acid plays many essential roles in mammalian and bacterial physiology and pathology. The goal in this thesis was to characterize sialic acid metabolism better and use this information to harness these enzymatic pathways for the production of non-natural sialic acid (nonulosonic acid) analogs.

The first stage of this work was to understand the substrate selectivity of NagA for non-native substrates. This work showed that NagA selectivity is related to the size and volume of the N-acyl group of the glucosamine substrate; the larger and more voluminous the acyl group, the low the activity with NagA. Three different experiments confirmed this result. The in vitro characterization demonstrated how the velocity of the reaction is affected by the size of the substituent. The in silico docking experiments rationalized the results obtained from the in vitro analysis by examining the different biding modes of the non-native substrates. Ultimately, an in vivo experiment confirmed that NagA is vital for the use of N-acyl amino sugars as the carbon source in Escherichia coli due to its role in the catabolic pathway of sialic acid and the recycling process of amino sugars.

Phosphorylation of amino sugars is critical for their biochemical processing. We found that both deacetylase (NagA) and epimerase (NanE) activity depends on the presence of the 6-phosphate group on the substrate. These findings spurred the development of a simple synthetic strategy to install a phosphate group on the C6 oxygen of hexoses. Although we found purification of the final product to be challenging due to the very high polarity, the strategy shows promise. It provides simple meth for making 6-phosphosugars of many carbohydrates. Further optimization of the purification protocol is required to complete this method.

For in vivo production of N-acyl sialic acid analogs, kinase-mediated phosphorylation of the N-acylglucosamine building block is likely not required. In E. coli, the phosphoenolpyruvate dependent phosphotransferase system (PTS) would likely be responsible for uptake of the exogenous sugar starting material concomitantly phosphorylating it. Thus, the next stage of the project was to investigate the ability of the epimerase NanE and the synthase NeuB to accept N-acyl derivatives of their native substrates.

When we tested NanE with ManNAc-6-P, the substrate was transformed into GlcNAc-6-P; however, when we tried the reverse reaction, we could not detect any product. As we knew from the literature that sugar epimerization to an axial substituent was possible, we were initially concerned that our NMR assay was not sensitive enough to detect the ManNAc-6-P product. However, careful analysis of the reaction conditions revealed that NanK was used to phosphorylate GluNAc, and due to its substrate specificity, no phosphosugar was produced. Based on the literature precedent and our results with the reverse reaction, it is expected that if we change the kinase, the conversion of GlcNAc-6-P into ManNAc-6-P will readily occur, confirming the ability of the epimerase to access our key sugar intermediate.

In the last stage of this thesis, we were able to test the last enzyme in the pathway by constructing the biosynthetic pathway for sialic acid analogs. This experiment showed that substrates with small N-acyl groups could be converted in vivo into N-acyl sialic acids. This result confirmed that NeuB could accept a range of ManNacyl derivatives. Interestingly in E coli strains possessing NagA production of N-acyl sialic acids was not observed, suggesting even the low level of background substrate hydrolysis observed for these typically poor substrates is sufficient to reduce the building block pool for sialic acid production. It

will be exciting to see if substrate with large acyl groups will be effectively converted into sialic acid and if NagA mediated hydrolysis will continue to confound the production system. This project addressed the fundamental question of could a NanE-mediated strategy to access N-acyl manosamines be useful for generating sialic acid in vivo. The answer is clearly yes, it is feasible. However, the thesis work also highlighted several questions still to be answered. It is crucial to confirm the level of promiscuity of NeuB as it will help design substrates that can be converted in their correspondent sialic acid analogs. There are two very other essential questions to be addressed; how does N-acylmanosamine gets dephosphorylated for use with NeuB? Is NanE epimerizing a phosphosugar or just the free sugar which has entered the cell via passive transport? While our preliminary data in chapter 4 suggests that the epimerization of the nonphosphorylated GlcNAc is slow, we cannot yet rule out this possibility. Answering these questions will give us a better understanding of the function of the enzymes and contribute to producing non-natural fine chemicals by a fermentation process.