Ph.D. School

University of Study of MILANO-BICOCCA

Department of *Biotechnology and Bioscience*

Ph.D. program in *Chemical, Geological and Environmental Sciences*, XXX Cycle

Curriculum in *Chemical Science*

Synthesis of Glycoderivatives as Molecular Tools in Medicinal Chemistry and Nano-Medicinal Chemistry

Ph.D. Candidate: Alice **Paiotta**

Matr. 708686

Tutor: Prof. Barbara **La Ferla**

Coordinator: Prof. Maria Luce **Frezzotti**

Academic Year 2016/2017

Non potrai mai raggiungere un reale successo a meno che tu non ami ciò che stai facendo.
(Dale Carnegie)

Non c'è abbastanza buio in tutto l'universo da spegnere la luce di una sola candela.
(Fannie Flagg)

Preface 6

PART 1. DEFINITION OF THE RESEARCH OBJECT

Chapter 1: Introduction

1.1 Introduction: overview on target 8

1.2 Aim of the Ph.D. Project 10

Chapter 2: The Cancer Disease

2.1 Cancer: "crazy" loss of properties 12

2.2 Normal cells vs tumoral cells 13

2.3 Metabolism and the *Warburg's Effect* 15

2.4 Unfolded protein response 17

2.5 *Adenocarcinoma of the Pancreatic Duct* (PDAC) 19

2.6 Conclusion 19

Chapter 3: Target Description

3.1 *Hexosamine Biosynthetic Pathway* (HBP) 20

3.2 *Alteration in cancer cells is linked with the activation of K-RAS* 22

3.3 *N*-acetylglucosaminephosphate mutase 23

3.4 Conclusion 24

PART 2. THEORETICAL FRAMEWORK AND WORKING HYPOTHESES

Chapter 4: Molecular Design

4.1 Drug Design definition 25

4.2 Drug Design on AGM1 25

4.3 Conclusion 27

Chapter 5: Computational Analysis

5.1 Introduction to Molecular Docking 28

5.2 Set up of the model for the Molecular Docking 29

5.3 Conclusion 35

PART 3. EXPERIMENTAL WORK

Chapter 6: Synthetic Strategies

6.1 Library of potential inhibitors 36

6.2 Synthetic Strategy 37

6.2.1 Synthetic Strategy 1. 38
Scheme A- *N*-acetylglucosamine-1-deoxy from *N*-acetylglucosamine 38
Scheme F - 6-sulfate derivative from 1-deoxy-*N*-acetylglucosamine 38
Scheme G - 6-sulfonate derivative from *N*-acetylglucosamine-1-deoxy 39

6.2.2 Synthetic Strategy 2. 40
Scheme C - 6-phosphoramidate derivate from *N*-acetylglucosamine 40
Scheme D - 6-sulfate derivative from *N*-acetylglucosamine 41
Scheme E - 6-sulfonate derivative from *N*-acetylglucosamine 42

6.2.3 Synthetic Strategy 3. 43
Scheme B - 2-Methyl-(3,4,6-tri-O-acetyl-1,2-dideoxy- α -D-glucopyrano)-[2,1-d]-2-oxazoline 43
Scheme H - 1-phosphoramidate derivative from *N*-acetylglucosamine 44
6.3 Conclusion 46

Chapter 7: Enzymatic test
7.1 Overview on High Performance Liquid Chromatography 47
7.2 Schematic description of HPLC instrument 48
7.3 Ion-pair chromatography 48
7.4 Aim of enzymatic test 49
7.5 HPLC: setting method 51
- Analysis UDP-GlcNAc standard and calibration line 51
- Analysis of cellular extract and enzymatic reaction 52
7.6 Results and Discussion 53
7.6.1 Cellular extract and endogenous UDP-GlcNAc 53
7.6.2 Enzymatic reaction with 1mM and 5mM substrates 54
7.6.3 Cellular extract, enzymatic reaction and quantification of UDP-GlcNAc 55
7.6.4 Enzymatic reaction with inhibitors on 10μL and 30μL of extract 55
7.7 Conclusion 59

Chapter 8: Biological tests

8.1 Introduction: potentials inhibitors and coefficient of distribution 60

8.2 Cellular test on MDA-MB-231 63

8.3 Conclusion 66

Chapter 9: Cyclodextrins and encapsulation

9.1 Introduction: traineeship program 67

9.2 Cyclodextrins properties 69

9.3 Cyclodextrins complex with drug 70

9.4 Cyclodextrins and cyclodextrins-complex characterization by NMR 71

9.5 Cyclodextrins synthesis 75

- Biotin-γ CD synthesis 76

- NBF-triazole-γ CD synthesis 79

9.6 Conclusion 82

Conclusion 83

Appendix

A.1 Technical features of Molecular Docking 86

A.2 Synthetic Techniques and characterization methods 87

A.3 Steps of synthesis 88

A.4 HPLC: setting method 135

A.5 Biological part 139

Acknowledges 142

List of Abbreviations 143

This script concerning the project carried out during my 3 years-Ph.D. In this first part a brief description of the Chapters of the thesis is done in order to get an overview of different argument treated during the Ph.D. period.

The PART 1. is divided in three chapters which describe and give a detail definition of the research object.

In Chapter 1 an introduction of the project is illustrated: this chapter introduces the goal of the work. The new approach of science, based on the exploitation of biocompatible molecules, is the synthesis of glycomimetics molecules of potential inhibitors in order to contrast cancer cells: Adenocarcinoma of the Pancreatic Duct (PDAC) is descripted as the target of studies.

The Chapter 2 concerns the description of main characteristics and functions of normal and tumoral cells. Carcinogenic cells alter their metabolism to favor their proliferation and need high quantity of glucose. The *Hexosamine Biosinthetic Pathway* (HBP), which biosynthetic starting point is the glucose, products Uridine-diphosphate *N*-acetylglucosamine (UDP-GlcNAc), the precursor for glycosylation for protein folding, which is overexpressed in this cells. Some alteration on pathway can bring to changes in *glycosylation* which are important for proliferation, motility and other functions of tumoral cells.

In the Chapter 3 there is a description of the targets of our studies: the *Hexosamine Biosynthetic Pathway* (HBP). Its role is important in particular in cancer cells in order to synthesized UDP-*N*-GlcNAc. Increas levels of glycosylation and altering expression of enzymes are the key factors in several human diseases. An illustration of *N-acetylglucosamine-phosphate mutase* (AGM1), enzyme involved in the HBP, is done: this mutase-enzyme is the focus point of our inhibition.

The PART 2. consists in two chapters with information about the theoretical framework and working hypotheses.

Molecular design is described in Chapter 4, a rational design of molecules which have similar structure of the natural substrate of enzyme AGM1 with some modification: the design of the compounds is followed by the planning of synthetic strategies to reach products. In this chapter are described the structures of *potential inhibitors*.

The Chapter 5 contains the principles of molecular docking used in order to get a virtual screening on the library of compounds which have been designed. This approach is important to have preliminary and theoretical results about the affinity and interaction between the potential inhibitors and the protein.

In the PART 3. the experimental work is described in three chapters: all the synthesis used in order to get the final potential inhibitors are described in Chapter 6; the synthetic mechanism of all the steps, the reagents used and the experimental conditions and are described.

In the Chapter 7 the validation of the analytical method of separation through the *High Performance Liquid Chromatography* is described, and the results of the enzymatic reaction on cellular extract are reported.

In the Chapter 8 are reported some preclinical evaluation of one of glycomimetics synthetized in a Triple Negative Breast Cancer (TNBC) cellular model.

Then, in Chapter 9 is described the encapsulation of one of enzymatic inhibitor in functionalized cyclodextrins for cancer cells targeting. The section include the aims of the traineeship period and the work developed in collaboration with Ph.D. Milo Malanga at *CycloLab* (Budapest), a brief description of cyclodextrins properties and the last part is about the synthesis of *cyclodextrins tools* for the encapsulation of potential inhibitors.

PART 1. DEFINITION OF THE RESEARCH OBJECT

Chapter 1
Introduction

This chapter briefly introduces the goal of the project. An initial overview explains the target of our studies i.e. a very common cancer named *Adenocarcinoma of the Pancreatic Duct* (PDAC). The new frontiers of science, based on the exploitation of biocompatible molecules, will be the starting point of our research to get new glycomimetics molecules, sugar-based structures: many biological and physiological processes involved this type of structures.
Finally, the *goal of the work* will be described.

1.1. Introduction: overview on target

In the last decades cancer is one of the main causes of death in the world ranging from young teens to old people. This is because of an increase in the average age of the population and other risk factors such as smoking and unhealthy diet: it is also linked to hereditary, environmental pollution, lifestyle and prolonged life time as well. Research on cancer mortality has declined considerably compared to the past, in particular thanks to the development of diagnostic methods that allows you to find out more cases in a timely manner and also the availability of more efficient therapies.

Every day in Italy 1,000 new cases of cancer are discovered.
The figures presented by the *Italian Association of Cancer Registries* (AIRTUM), regarding the increase in cases of cancer patients in Italy, show the scenario of a dominant social disease and point out the research improvements occurred over the time. Fortunately, cancer mortality is much decreased compared to the past, mainly due to the evolution of diagnostic methods that permit to discover early a larger number of cases beside the availability of more effective innovative therapies and a better information.

Because of the increase of new cases over the entire population and the death rate, PDAC make it one of the "big killers" in the world. Most patients do not exceed five months from the diagnosis of pancreatic cancer and, after about five years, only 5% of them survive.[1][2][3]

Our research is focused in particular on the study of PDAC: being a "silent killer", the majority of PDAC sick people are diagnosed too late - advanced cancer - , beyond the stage of a possible surgical resection. Many biological processes rely on the role of carbohydrates in regulating the functionality of different proteins and enzymes in many biosynthetic pathways.

[1] Guillaumond, F. et al., "Pancreatic tumor cell metabolism: focus on glycolysis and its connected metabolic pathways." *Archives of biochemistry and biophysics* 545 (2014): 69-73.
[2] Peto, J. et al., "Cancer epidemiology in the last century and the next decade." *Nature* 411.6835 (2001): 390-395.
[3] Doll, R. et al., "An epidemiological perspective of the biology of cancer." *Cancer Research* 38.11 Part 1 (1978): 3573-3583.

These include the *Hexosamine Biosynthetic Pathway* (HBP) essential for cell growth and involved in proliferation, survival and migration of tumor cells, in particular PDAC. The understanding of the molecular basis of HBP's role in tumor can help us to identify compounds that interfere with it: its inhibition may represent a basic research tool essential to develop a possible strategy able to arrest or kill PDAC tumor cells.[4]

Glycomimetics explore the role of carbohydrates in biological processes. Almost all cell surface and secreted proteins are modified by covalently-liked carbohydrate moieties and the glycan structures of these glycoproteins are essential mediators in processes such as protein folding, cell signaling, neuronal development, hormone activity, and the proliferation of cells and their organization into specific tissue. Many viral and bacterial infections involve recognition of host carbohydrates by microbial cell surface proteins. Tumor progression is associated with the appearance of unusual carbohydrates on the surface of cells, and the expression of these tumor associated carbohydrate antigens provide opportunities for the development of diagnostics and immune-therapies.[5]

This biological process includes carbohydrates–carbohydrates, carbohydrates–proteins, and carbohydrates–nucleic acids interactions. Carbohydrates, in the form of glycopeptides, glycolipids, glycosaminoglycans, proteoglyans, or other glyco-conjugates have long been known to participate in many biological processes. These include viral entry, signal transduction, inflammation, cell–cell interactions, bacteria–host interactions, fertility and development. Due to their frequent branching and linkage diversity, oligosaccharides have greater structural complexity than nucleic acids and proteins.

Carbohydrates are one of the main types of nutrients exploited by human body as source of energy for cells, tissues and organs. They play an important role in biological, physiological and pathological processes and they are crucial in the regulation of the function of different proteins and enzymes. Recent chemical advances, such as improved synthetic methods and enzymatic synthesis, have opened new possibilities in obtaining pure and chemically defined carbohydrates. At the same time, the field has seen growing interest in the development of carbohydrate microarrays and neoglycoconjugates to facilitate otherwise laborious biological studies.[6]

The activity carried out is focuses on the design and preparation of compounds of biological interest, particularly glycomimetics with antitumor activity.

[4] AIRC-airc.it/cancro/cos-e/statistiche-tumori-italia

[5] Carbohydrate Recognition: Biological Problems, Methods, and Applications, edited by Binghe Wang, Geert-Jan Boons (2011).

[6] Ratner, Daniel M., et al. "Tools for glycomics: mapping interactions of carbohydrates in biological systems." *ChemBioChem* 5.10 (2004): 1375-1383.

The role of sugars in human body have stimulated the design of glycomimetics, molecules with similar structures to carbohydrates, which could interfere in the biological functions of the enzymes in specific biosynthetic pathways.[7]

The Hexosamine Biosynthetic Pathway is used by normal and tumoral cells to synthesize *Uridine diphosphate-N-acetil-D-glucosamine* (UDP-*N*-GlcNAc), substrate for glycosylation for proper protein folding. Increasing levels of glycosylation and altered expression of enzymes that control this change are the key factors in several diseases such as diabetes, neurodegenerative diseases and finally cancer. Whereby an alteration of the HBP, and therefore an altered production of UDP-*N*-GlcNAc, is likely to promote the proliferation, adhesion, motility, resistance to stress, angiogenesis and tumor cell metastasis.[8]

1.2. Aim of the Ph.D. Project

Important features of tumoral cells such as proliferation, adhesion, motility, resistance to stress and metastasis are linked to glycosylations. The HBP, indeed, feeded by glucose and by glutamine is considered a cellular sensor of nutrient levels that supply cancer cells with UDP-*N*- acetylglucosamine. This one is a donor molecule for *N*- and *O*- protein glycosylation that constitute post-translational modification: its modulation may represents an important feature for cancer cells.[9] Accordingly HBP inhibition as well as its positive and negative modulation may represent a fundamental basic research tool to further define its role in cancer biology as well as a new promising therapeutic routes.

Due to the important role of the HBP in the cells, this research project is focused on the synthesis of glycomimetics which can interfere with the biological functions of enzymes in particular biosynthetic pathways, including HBP. The research key concept is based on the role that sugars play in regulating the functions of different proteins and enzymes.

The main objective of the research project is to identify and synthesize new glycomimetics as molecular tools to study the role of HBP in regulating of proliferation and survival of cancer cells. HBP uses glucose, glutamine, acetyl CoA and glucose 1,6-bisphosphate and uridine to synthesize UDP-GlcNAc, the substrate of the enzymes catalyzing the reactions of *N*- and *O*- glycosylation of the proteins.[10] Inhibition of HBP blocks the UDP-GlcNAc synthesis and consequently the glycosylation reactions. Therefore the inhibition of

[7] Hirabayashi, Jun, and Ken-ichi Kasai. "Glycomics, Coming of age!." *Trends in Glycoscience and Glycotechnology* 12.63 (2000): 1-5.
[8] Malvezzi, M., et al. "European cancer mortality predictions for the year 2014." *Annals of oncology* 25.8 (2014): 1650-1656.
9 Bacigalupa, Zachary A., Chaitali H. Bhadiadra, and Mauricio J. Reginato. "O-GlcNAcylation: key regulator of glycolytic pathways." *Journal of bioenergetics and biomembranes* (2018): 1-10.
[10] Palorini, R., et al. "Glucose starvation induces cell death in K-ras-transformed cells by interfering with the hexosamine biosynthesis pathway and activating the unfolded protein response." *Cell death & disease* 4.7 (2013): e732.

HBP, as well as its positive and negative modulation, can be the principal mean of basic research to further define its role in cancer biology.[11]

Recent studies have confirmed that both glucose and glutamine, whose adoption from cancer cells is driven by the oncogene KRAS (*V-Ki-ras2 Kirsten viral oncogen homologous sarcoma)*, are critical for the PDAC progression and to support the growth of pancreatic cancer by limiting oxygen availability. Consequently the areas of investigation for this type of cancer aims at affecting the mechanisms through which the oncogenic KRAS coordinates the metabolism to support the tumor growth.[9] The understanding of the HBP pathway and its response in PDAC, through innovative chemical tools, and the identification of modulators could allow the development of novel therapeutic strategies for PDAC tumors.

In this context a library of potential inhibitors similar to the natural substrate of the catalytic mechanism, or analogous to the product's structure but with some important modification in strategic positions have been prepared.

[11] Kim, Young Hyo, Tsuguhisa Nakayama, and Jayakar Nayak. "Glycolysis and the Hexosamine Biosynthetic Pathway as Novel Targets for Upper and Lower Airway Inflammation." *Allergy, asthma & immunology research 10.1* (2018): 6-11.

Chapter 2
The Cancer Disease

In this part characteristics and functions of both *normal and tumoral cells* are described: their different growth, cell division and proliferation capability and it will be done an accurate description about different behaviors in case of the stress of the endoplasmic reticulum due to changes in *glycosylation*.

2.1. Cancer: "crazy" loss of properties

Cancer is a genetic disease caused by the modification of genes that control several function like growth and cellular replication. Genetic modifications that cause cancer can be hereditary or can be related to a particular lifestyle or to be a result of errors during the cell division or DNA defects caused by specific environmental exposures, smoking, radiations and so on. Each type of cancer is characterized by a precise genetic combination. In general, cancer cells may have more than one single genetic modification.[12,13]

There is not a single disease called cancer but different types of diseases which have different causes and affects different organs and tissues and each one require its own specific therapy and diagnostic solution. However, there are some properties and characteristics that are common to all cancers. To use a metaphor, at one point, a cell of the organism "crazily" loses some of its properties, getting other ones and starting to multiply itself without any rules. Inside each cell there are "genes" which, in case of a "crazy" cell, don't work anymore therefore allowing the cell to survive giving rise to a tumor. Because these genes are out of order, the mechanism which controls the cells replication is compromised and a huge number of other cancerogenic cells with the same defect are generated. Healthy cells are progressively supplanted by the more exuberant cancer cells. Both a benign and malignant tumor tend to proliferate abnormally but, while a benign tumor remains confined to the organ in which it developed, the malignant tumor tends to grow and migrate in a process called metastasis, the most advanced stage of tumor progression and the real cause of cancer deaths. In this way a malignant tumor hurt the human body by compressing or invading the surrounding tissues. Its aggressiveness is tightly related to its capability to enter into the blood circulation reproducing the disease in other parts of the body, forming other cancers named metastasis.

Cancer results from an accumulation of mutations, i.e. genetic alterations that don't regulate anymore several cell's features like proliferation,-survival, membership and motility. Mutations can develop in very different times, even under the influence of external stimuli. The benign tumor can be considered the first stage of these alterations.

However, very often, this stage is skipped and you get the malignant tumor without obvious symptoms.

[12] G. Karp, *Biologia cellulare e molecolare*, Casa Editrice EdiSES, Napoli, III edizione, 2008.
[13] Chial, Heidi. "Proto-oncogenes to oncogenes to cancer." *Nature Education 1.1* (2008): 33.

The necessary and sufficient causes for tumor development are already "written" at the origin in the genes, that are hereditary. In the majority of tumors, the alterations of genes are determined by environmental causes such as prolonged exposure to carcinogens, chemical, physical or other dangerous substances. However the cigarette smoke, asbestos, certain substances developed burning coal or oil, an unbalanced diet, alcohol, ultraviolet rays from the Sun, chemicals used by employees in certain industrial processes or in agriculture, can increase the predetermined genetic "fragility" and get to cause mutations. In some cases the mutations are generated by errors in the mechanism of cell replication regardless of the external environment. Tumor needs oxygen and nutrients to grow and develop. For this reason the tumor produces substances that stimulate growth of new blood vessels – angiogenesis - that will feed the new growing tissue.[14]

Pancreatic cancer seems to be one of the most aggressive and deadly; this is due to the propensity for local invasion and metastasis, the appearance of an extensive desmoplastic stromal reaction and the tumor induced immunotollerance response. In addition a low vascular network creates a hypoxic environment preventing the proper spread of oxygen and nutrients to cancer cells. To survive and proliferate under these conditions, pancreatic tumor cells exploit specific metabolic pathways to satisfy their enormous requirement of energy and biomass. However, PDAC is a heterogeneous tumor and the reprogramming of metabolic processes involved in cancer cells to get oxygenation and nutrients is quite complex.[1]
In addition, KRAS gene mutations (also known as *V-Ki-ras2 Kirsten viral oncogen homologous sarcoma*), found in more than 90% of cases, play an essential role in the initiation, development and maintenance of the KRAS gene: KRAS alters the expression of the enzymes involved in the use of glucose by promoting the nucleotide's biosynthesis and the glycosylation of proteins that support tumor growth and favors the production of NADPH, with a decrease in the production of oxygen-reactive cytotoxic species generated during cell proliferation.[15,16,17]

2.2. Normal cells vs tumoral cells

The most important feature of a tumor cell regardless if it is inside the human body or in a culture plate, is the loss of growth control. The growth and division capacity is not drastically different between a tumor cell and most healthy cells: when healthy cells are grown in tissue culture, they develop and divide at a similar rate. However, healthy cells proliferate and tend to remain in a single cell layer.

[14] http://www.airc.it/cancro/cos-e/come-nasce-tumore/
[15] Cogoi, Susanna, et al. "MAZ-binding G4-decoy with locked nucleic acid and twisted intercalating nucleic acid modifications suppresses KRAS in pancreatic cancer cells and delays tumor growth in mice." *Nucleic acids research 41.7* (2013): 4049-4064.
[16] Ji, Zhenyu, et al. "Oncogenic KRAS activates hedgehog signaling pathway in pancreatic cancer cells." *Journal of Biological Chemistry 282.19* (2007): 14048-14055.
[17] E. Ozres "Adenocarcinoma del pancreas, conoscere le mutazioni genetiche per diagnosticare il tumore prima che sia troppo tardi"(2014).

Malignant cells, under the same conditions, continue to grow stacked on each other by forming masses: this is because they no longer respond to regulatory signals that stop the growth and division. Tumor cells continue to grow in the absence of growth stimulating signals that are required by normal cells: their cell cycle is not dependent on the interaction of growth factors and their receptors. Normal cell cultures exhibit a limited ability to divide; after a number of mitotic divisions, they are facing an aging process that make them unable to grow and multiply. The tumor cells, on the other hand are to be considered *immortal* because they continue to multiply infinitely.

A normal cell before turning into a cancer cell has to undergo numerous transformations (one or more DNA mutations). The first mutation affects cellular reproduction mechanisms and causes uncontrolled proliferation of the cell itself (initiation). But the cell is transformed into cancer cell (promotion) only after developing further mutations that give it the following characteristics:

- uncontrollable proliferation.
- escape from death due to aging and "programmed suicide": cancer cells can activate an enzyme (telomerase) that addresses aging cell damage and avoids death. This phenomenon refers to the tumor cell, which guarantees greater survival and therefore more chances to accumulate mutations that increase the malignant potential. The cancerous cell manages to escape the apoptosis program and sometimes, paradoxically, is the body itself that produce anti-apoptotic substances that help the cell survive.[18,19,20]
- stimulate blood vessels formation to provide nutrition: newly crafted vessels move through the cellular mass by providing nutrition and allowing them to grow further in size. This process is called *angiogenesis*.[21]
- ability to break away from the origin tissue.
- ability to overcome barriers (i.e. connective tissue membranes) and invade surrounding tissues, penetrate into the bloodstream and lymphatic circulation.
- proliferate in other areas of the human body in addition to the original one.

The specific characteristics of the cancerous cells, as they are not shared by healthy cells, can become the goal of targeted biological therapies, which have greater selectivity on the cancer cell than conventional chemotherapy.

[18] Weinberg, Robert A. "How cancer arises." *Scientific American 275.3* (1996): 62-71.
[19] Renehan, Andrew G., Catherine Booth, and Christopher S. Potten. "What is apoptosis, and why is it important?." BMJ: *British Medical Journal 322.7301* (2001): 1536.
[20] Sjöström, Johanna, and Jonas Bergh. "How apoptosis is regulated, and what goes wrong in cancer." BMJ: *British Medical Journal322.7301* (2001): 1538.
[21] Hart, Ian R. "Perspective: tumour spread—the problems of latency." *The Journal of pathology 187.1* (1999): 91-94.

2.3. Metabolism and the Warburg's Effect

Carcinogenic cells alter their metabolism to support rapid proliferation and cellular expansion. In particular they need glucose which is used in the oxidative phosphorylation pathway, generally exploited for aerobic glycolysis. Glucose deprivation can trigger oncogenes and these can increase – up regulate – the proteins involved in aerobic glycolysis. Proteins involved in increasing glycolysis may render tumor cells more resistant to apoptosis. Aerobic glycolysis induces an acidification of the tumor environment, favoring the development of a more aggressive and invasive phenotype.
Cellular energy metabolism is one of the processes involved in the transition from normal cell to carcinoma cell. In fact, glucose metabolism is altered in tumor cells (*Figure 2.1)*

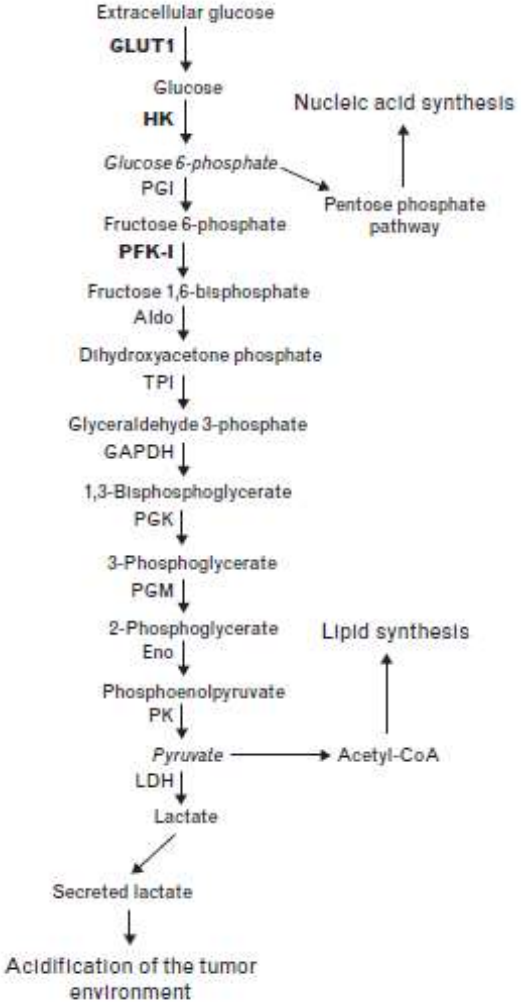


Figure 2.1: glycolysis in cancer cells.

The glucose consumption increment through the glycolytic pathway allows the production of intermediates that are essential for carcinogenic cells to rapidly proliferate. One of these intermediates, glucose 6-phosphate, is used for the synthesis of nucleic acids through the pathway of phosphate pentoses to allow rapid DNA replication. The abundant production of pyruvate stimulates lipid synthesis which is necessary for the formation of membranes in the division of carcinogenic cells. Finally, lactate secretion from cancer cells induces an acidification of the environment that promotes tumor progression and inhibition of anticancer drugs.[22] Carcinogenic cells can therefore undergo aerobic glycolysis producing large amounts of lactate regardless of oxygen availability, this is termed as *Warburg Effect*.

The role of metabolism in cancer has been a focus of extensive research. In the 1920s, Otto Warburg observed that tumor cells produce energy by undergoing high rates of glycolysis and lactate fermentation in the cytosol in the presence of oxygen, condition usually found only in the oxygen-deprived rate. In contrast, normal cells typically have lower glycolysis rates followed by pyruvate oxidation in mitochondria. This cancer-associated phenomenon was later termed as the *Warburg Effect*[23] (*Figure 2.2*): firstly it allows cells to use glucose to produce ATP. Although the ATP yield per mole of glucose consumed is low, if the glycolytic flow is high enough, the percentage of ATP produced at the cellular level by glycolysis can exceed that one produced by oxidative phosphorylation. Secondly, glucose degradation provides

[22] A. Annibaldi, C. Widmann, *Current Opinion in Clinical Nutrition and Metabolic Care* (2010), 13, 466–470
[23] Weinberg R. A., CST *Guide: pathway & protocols – Cell Signaling Technology*, INC

intermediates needed for the biosynthetic pathways used to produce sugars (ribose for nucleotides), lipids (glycerol and citrate), and through the pathway of the pentoso-phosphate the production of NADPH. It can therefore be emphasized that Warburg's effect has bioenergetic and biosynthetic benefits. The reason why the lactate production rate is significantly higher than the pyruvate one is due to the fact that in these cases glycolysis exceeds the maximum pyruvate production speed, if not, the cells should eliminate pyruvate by means of high flow. This leads the system to avoid pyruvate buildup, turning it into lactate.

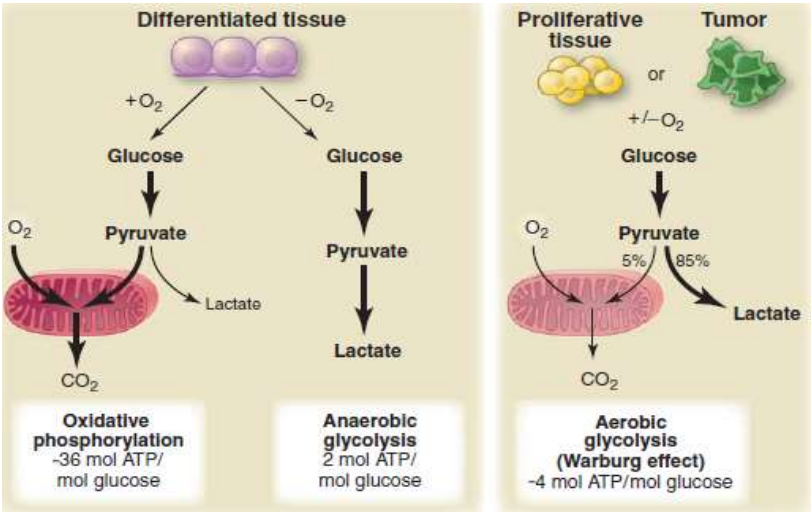


Figure 2.2: schematic representation of the difference between oxidative phosphorylation in anaerobic and aerobic glycolysis. In the presence of oxygen, the glucose is transformed into pyruvate and the latter is oxidized to CO_2 as a result of oxidative phosphorylation. Oxygen is essential in this process because is the ultimate acceptor of electrons. In limited conditions of oxygen, cells can convert pyruvate into lactate to recover glucose, but with poor production of ATP. Warburg observed that carcinogenic cells tend to convert most glucose into lactate. In the proliferation of cells about 10% of glucose is directed to the pathway of pyruvate production.

In contrast to normal differentiated cells, which rely primarily on mitochondrial oxidative phosphorylation to generate the energy needed for cellular processes, most cancer cells instead rely on the aerobic glycolysis: aerobic glycolysis is an inefficient way to generate adenosine 5′-triphosphate (ATP), however, and the advantage it confers to cancer cells has been unclear. [24]

2.4. Unfolded protein response [25]

[24] Vander Heiden, Matthew G., Lewis C. Cantley, and Craig B. Thompson. "Understanding the Warburg effect: the metabolic requirements of cell proliferation." S*cience* 324.5930 (2009): 1029-1033.
[25] Lee, Amy S. "The glucose-regulated proteins: stress induction and clinical applications." *Trends in biochemical sciences* 26.8 (2001): 504-510.

To survive in a hostile environment with limited nutrients cancer cells co-opt cellular regulatory pathways that facilitate adaptation and thereby maintain tumor growth and survival potential. The endoplasmic reticulum (ER) is uniquely positioned to sense nutritional deprivation stress and subsequently engage signaling pathways that promote adaptive strategies.

The initial growth of solid tumors is not well coordinated with the formation of new blood vessels necessary to support malignant expansion (*J. Folkman et al.(1966)* [26]). Thus, tumors rapidly outstrip the existing vasculature, resulting in a restricted microenvironment for critical nutrients (glucose and oxygen). Cellular energy and redox homeostasis is critical for aspects like proper folding, modification of transmembrane and secreted proteins within the Endoplasmic Reticulum (ER). Perturbation of this homeostasis triggers a cellular checkpoint referred to as the unfolded protein response. The capacity of the ER to respond to alterations in nutrient status makes it an effective early sensor of cellular stress associated with tumorigenesis also known as hypoxic areas, constitute a unique and dynamic feature of the tumor microenvironment that is known to contribute to cancer progression (also see other reviews about this issue).

Clinically, hypoxia has several adverse and important implications. Hypoxic tumor cells are known to be more resistant to current treatment modalities.[27]

Productive folding of secreted proteins in the endoplasmic reticulum (ER) is essential to ensure normal cell function. In order to get the proteins folded properly, ER homeostasis must be maintained. ER homeostasis is defined by the dynamic balance between the ER protein load and the ER capacity to process this load. ER homeostasis can be perturbed by pathological processes such as hypoxia, glucose deprivation, viral infections, environmental toxins, inflammatory cytokines, and mutant protein expression, as well as by physiological processes such as aging. Disruption of ER homeostasis causes accumulation of unfolded and misfolded proteins in the ER. This condition is referred to as ER stress. Cells cope with ER stress by activating the *Unfolded Protein Response* (UPR), represented in *Figure 2.3*.

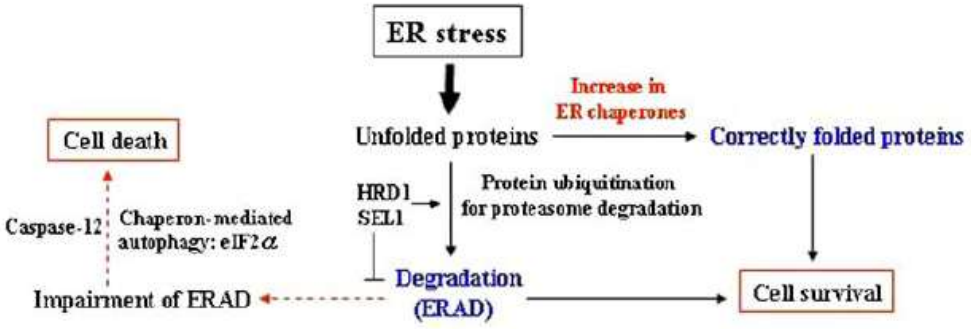


Figure 2.3: activation of the Unfolded Protein Response (UPR) due to the ER stress.

The UPR is initiated by three ER trans-membrane proteins: *Inositol Requiring 1* (IRE1), *PKR-like ER kinase* (PERK), and *Activating Transcription Factor 6* (ATF6). These three master regulators sense and interpret protein

[26]Folkman, Judah, P. E. T. E. R. Cole, and Shirley Zimmerman. "Tumor behavior in isolated perfused organs: in vitro growth and metastases of biopsy material in rabbit thyroid and canine intestinal segment." *Annals of surgery* 164.3 (1966): 491.
[27] Taylor et al., *Cancer Biology & Therapy* (2006), 5, 7

folding conditions in the ER and translate this information across the ER membrane to regulate downstream effectors.

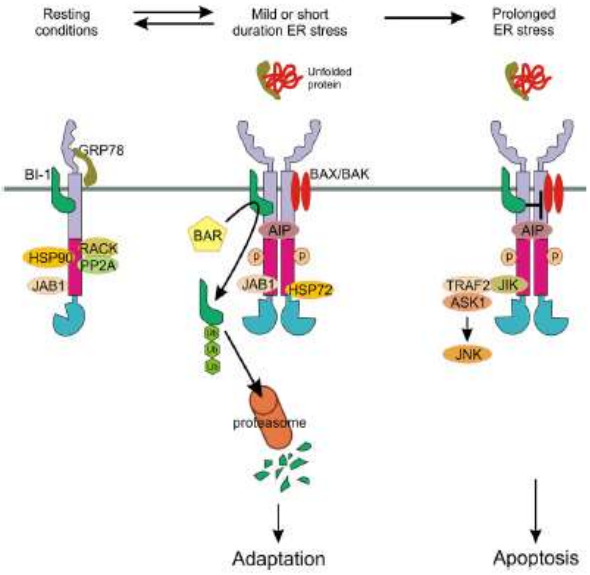


Figure 2.4: different behavior between normal and cancers cells during the ER stress.

These effectors have two distinct outputs, homeostatic and apoptotic, showed in *Figure 2.4*. Homeostatic outputs are adaptive responses to attenuate ER stress and restore ER homeostasis. These responses include the attenuation of protein translation to reduce ER workload and prevent further accumulation of unfolded proteins; upregulation of molecular chaperones and protein processing enzymes to enhance the ER folding activity, and the increase in *ER-associated degradation* (ERAD) components to promote the dregradation of unfolded proteins. When ER stress reaches the point where the cells cannot tolerate the load of unfolded proteins anymore, apoptosis sets in. Recent studies have indicated that cells suffering from insufficient blood supplies experience ER stress. The ER needs energy and oxygen for the folding process, thus nutrient deprivation (low ATP production) and hypoxia caused by insufficient blood supply leads to inefficient protein folding and ER stress in the cells, especially in cancer cells that grow and spread rapidly.[28]

2.5. *Adenocarcinoma of the Pancreatic Duct* (PDAC)

Pancreatic cancer remains a major unsolved health problem, with conventional cancer treatments having little impact on disease course. It is a highly lethal disease, which is usually diagnosed in an advanced state

[28] Gorman, Adrienne M., et al. "Stress management at the ER: regulators of ER stress-induced apoptosis." *Pharmacology & therapeutics* 134.3 (2012): 306-316.

for which there are little or no effective therapies. It has the worst prognosis of any major malignancy (3% 5-year survival).

Almost all patients who have pancreatic cancer develop metastases and die. The main risk factors are smoking, age, and some genetic disorders, although the primary causes are poorly understood. Advances in molecular biology have, however, greatly improved understanding of the pathogenesis of pancreatic cancer. Many patients have mutations of the KRAS oncogene, and various tumour-suppressor genes are also inactivated. Growth factors also play an important part. However, disease prognosis is extremely poor. Around 15–20% of patients have resectable disease, but only around 20% of these survive to 5 years. Despite pancreatic cancer's resistance to currently available treatments, new methods are being investigated.[29]

Also advances in surgical and medical therapy, little effect has been made on the mortality rate of this disease. One of the major hallmarks of pancreatic cancer is its extensive local tumor invasion and early systemic dissemination. The molecular basis for these characteristics of pancreatic cancer is incompletely understood.[30]

In certain patients with pancreatic and biliary cancer, chemotherapy may relieve tumour-related symptoms, improve quality of life and possibly prolong survival. The extent of these improvements is not completely known in spite of the extensive use of this treatment modality. Chemotherapy can add to both quantity and quality of life in advanced pancreatic and biliary cancer. The number of patients who benefit from treatment is, however, still limited.

It is necessary to look for new methods of prevention for this type of cancer, a better diagnosis capacity and with lower timing and not in advanced stages of the disease, and finally the possibility to the development of different treatments compared to chemotherapy.[31]

2.6. Conclusion

Normal and cancer cells haven’t considerable different capacity to grow and proliferate. Proliferation is linked to glycosylation and cancer cells get an overexpression of some pathway which can provide this post translational modification. Alteration of *N*- and *O*- glycosylation brings to UPR during the endoplasmic reticulum stress: we want to take advantage from the different behavior of normal and cancer cells during the ER stress in order to inhibit the proliferation of the second ones exploiting their secondary response to apoptosis.

[29] Li, Donghui, et al. "Pancreatic cancer." *The Lancet* 363.9414 (2004): 1049-1057.
[30] Li, Chenwei, et al. "Identification of pancreatic cancer stem cells." *Cancer research* 67.3 (2007): 1030-1037.
[31] Riall, Taylor S., et al. "Pancreatic cancer in the general population: improvements in survival over the last decade." *Journal of gastrointestinal surgery* 10.9 (2006): 1212-1224.

Chapter 3
Target Description

The first part of the chapter looks at the *Hexosamine Biosynthetic Pathway* (HBP), focus point of the research project, while the second part talks about *N-acetylglucosamine-phosphate mutase* (AGM1), enzyme involved in the HBP that we want to inhibit.

3.1. *Hexosamine Biosynthetic Pathway* (HBP)

One of the most used metabolic pathways by tumor cells, especially PDAC tumoral cells for their growth and proliferation, is the *Hexosamine Biosynthetic Pathway* (HBP). In order to block this metabolic pathway, we want to get an inhibition or a reduction of the enzyme catalytic action involved in the metabolic process to reduce the UDP-GlcNAc formation, a biosynthetic precursor of glycoproteins. In this way, abnormal expression of glycosylation causes production of misfolded proteins, which activate the UPR (*Chapter 2 – paragraph 2.4*).

The inhibition involves both healthy and tumorous cells: the normal cells initially activate the *primary adaptive response* followed by the activation of the proteosome, an enzyme that can "digest" the wrongly folded proteins, thus restoring initial homeostasis. If the stress of the endoplasmic reticulum is continued, healthy cells undergo to programmed cell death - apoptosis.

For the tumorous ones the HBP represents a key point for the production of glycoproteins, as a consequence, its inhibition causes an important decrease in *N*- and *O*-glycosylation, which are indispensable for cell adhesion, proliferation and motility. They give the secondary response of apoptosis (*Chapter 2 - paragraph 2.4*).

The study of HBP can help to get a better understanding on PDCA and how to defeat it. The HBP is important to synthesize essential substrates for the glycosylation processes utilizing glucose derivatives for the synthesis of the UDP-GlcNAc (*Figure 3.1*), the substrate used in the glycosylation reactions necessary for the proper folding of the proteins.

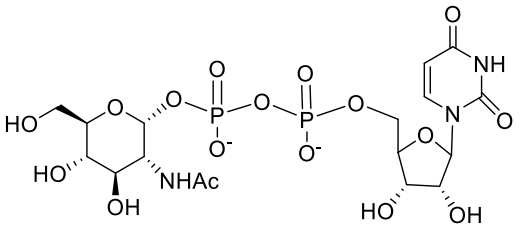


Figura 3.1: Uridine diphosphate-*N*-acetil-D-glucosamine.

An increase in glycosylation levels and alteration of the expression of enzymes that control such modulation play an important role in various pathologies such as cellular diabetes (regulation of *O*-GlcNAc involved in type II diabetes), neurodegenerative diseases and finally, in cancer. Hence, an alteration of HBP, and consequently an anomalous UDP-*N*-GlcNAc production can promote proliferation, adhesion, motility, stress resistance, angiogenesis, and ultimately metastasis of cancer cells.

The HBP pathway (*Figure 3.2*), is used to obtain the *N*-acetylglucosamine which is formed from fructose 6-phosphate through a series of enzymatic steps.

The first step is a glycolysis reaction in which fructose 6-phosphate is converted to glucosamine 6-phosphate by the GFAT (fructose 6-phosphate amidotransferase) enzyme. The glucosamine 6-phosphate is thus converted to the *N*-acetylglucosamine 6-phosphate intermediate through GAT (D-glucosamine 6-phosphate *N*-acetyltransferase). *N*-acetylglucosamine 6-phosphate is converted to *N*-acetylglucosamine 1-phosphate (GlcNAc-1-P) thanks to the AGM1 enzyme (*N*-acetylglucosamine-phosphate mutase) which belongs to the α-D-phosphoesomutase family. The final step, favored by enzyme AGX1 (UDP-GlcNAc pyrophosphorylase) with substrate UTP, gives the final product of HBP, the UDP-*N*-GlcNAc (*N*-acetyl-D-glucosamine): glycosylation reaction occurs, in fact UDP-GlcNAc can be transferred from the OGT enzyme (β-*N*-acetylglucosamine aminotransferase) to get glycoprotein. Hydrolysis made by OGA enzyme reconstitutes the unmodified protein and GlcNAc which can be recycled by the biosynthetic pathway.

The "*GlcNAc salvage HBP*" recycles the GlcNAc from the cell and the conversion into GlcNAc-6-phosphate is carried out by the GNK kinase enzyme (acetylglucosamine kinase). The process consists in the addition of a phosphate group to *N*-acetylglucosamine by the use of an ATP molecule. So the cell is able to supply *N*-acetylglucosamine 6-phosphate.[32]

[32] Vocadlo, David J., et al. "A chemical approach for identifying O-GlcNAc-modified proteins in cells." *Proceedings of the National Academy of Sciences* 100.16 (2003): 9116-9121.

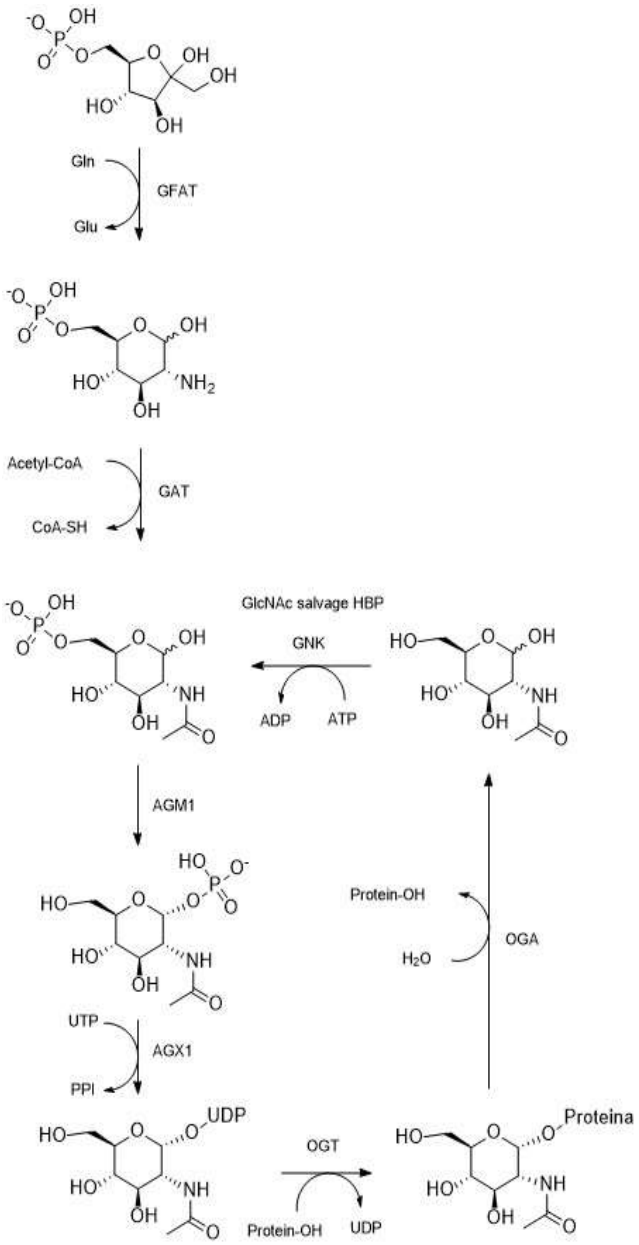


Figure 3.2: Hexosamine biosynthetic pathway (HBP).[2]

3.2. *Alteration in profile of cancer cells is linked with the activation of K-RAS*

The inhibition of the HBP causes alteration in glycosylation: this alteration in the metabolic profile of cancer cells is linked with the activation of K-RAS oncogene, important for the growth and cellular division. K-RAS promotes glycolysis by increasing the use of glutamine and decreasing mitochondrial activity. It plays a key role in tumor progression: recent studies have confirmed that glucose and glutamine are used, in particular in PDAC to support the growth of the pancreatic tumor limiting the availability of oxygen. It regulates glucose metabolites including glucose 6-phosphate and fructose 6-phosphate, which are the

precursors of HBP, and the pentose phosphate pathway. The oncogenic form of K-Ras is associated with an increase in *N*- and *O*-glycosylation. [33,34]

3.3. *N-Acetylglucosamine-phosphatemutase* (AGM1)

N-Acetylglucosamine-phosphatemutase (AGM1 – *Figure 3.3*) is an essential enzyme in the synthetic process of UDP-GlcNAc, UDP sugar that serves as a biosynthetic precursor of glycoproteins, mucopolysaccharides. AGM1 catalyzes the conversion of *N*-acetylglucosamine 6-phosphate (GlcNAc-6-P) to *N*-acetylglucosamine 1-phosphate (GlcNAc-1-P). This enzyme is a member of the D-phospho-hexomutase superfamily, which catalyzes the intramolecular phosphoryl transfer of sugar substrates. In *Figure 3.3* is shown the crystal structures of AGM1 from *Candida albicans*: the structure consists of four domains, of which three domains have essentially the same folding. The overall structure is similar to those of phosphono hexomutases; however, there are two additional strands in *domain 4*, and a circular permutation occurs in *domain 1*. The catalytic cleft is formed by four loops from each domain. The *N*-acetyl group of the substrate is recognized by *Val-370* and *Asn-389* in *domain 3*, from which the substrate specificity arises. The substrate rotates about 180° on the axis linking C-4.[35]

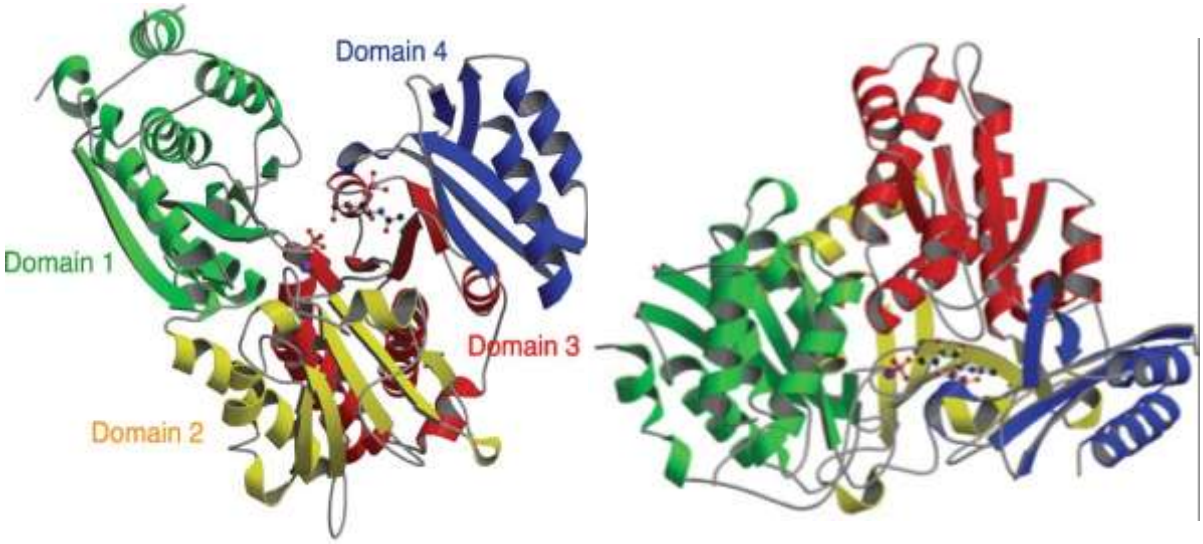


Figure 3.3: on left front view of AGM1 enzyme. On right top view of the AGM1 enzyme where both the GlcNAc-1-P molecule and the PO_4^{3-} ion are shown by the ball and stick model.

[33] Ji, Zhenyu, et al. "Oncogenic KRAS activates hedgehog signaling pathway in pancreatic cancer cells." *Journal of Biological Chemistry*282.19 (2007): 14048-14055.
[34] Ying, Haoqiang, et al. "Oncogenic Kras maintains pancreatic tumors through regulation of anabolic glucose metabolism." *Cell* 149.3 (2012): 656-670.
[35] Nishitani, Yuichi, et al. "Crystal structures of N-acetylglucosamine-phosphate mutase, a member of the α-D-phosphohexomutase superfamily, and its substrate and product complexes." *Journal of Biological Chemistry* 281.28 (2006): 19740-19747

The enzyme has an active site with four different domains:

- domain 1: active serine loop, it contains the catalytic residue phospho-serine necessary to transfer phosphate group.
- domain 2: metal-binding loop, for the coordination of Mg^{2+} ions required for the enzymatic activity.
- domain 3: sugar-binding loop, which allow the enzyme to recognize the two different orientations of the phospho-sugar located in C-1 and C-6 position in the substrate.
- domain 4: phosphate-binding loop, presents the catalytic residues that determine the orientation of the phospho-sugar substrate.

The study of the enzyme's catalytic site it necessary to understand which interactions are created between the substrate and the active site and then assess possible changes to make on glycomimetic structures. The transfer of the phosphate group from the anomeric position to C-6 is carried out by passing through a bis phosphate intermediate following the mechanism shown below:[27]

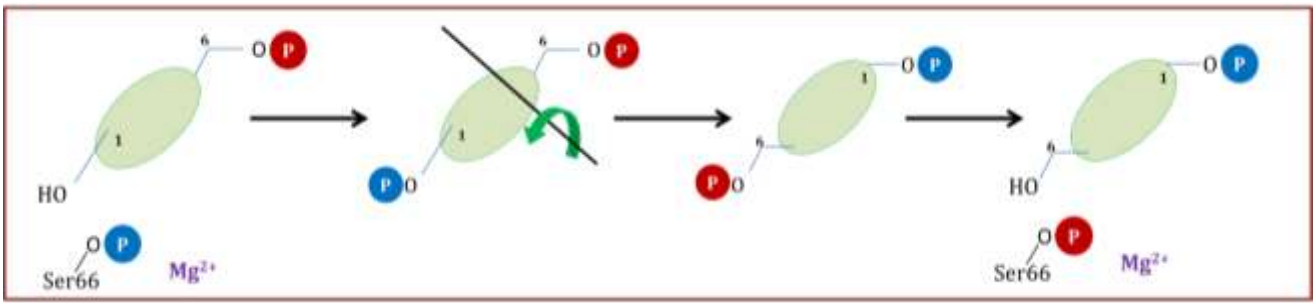


Figure 3.4: catalytic mechanism of enzyme AGM1.

The phosphate group of the serine residue Ser66 of the enzyme active site is transferred to GlcNAc-6-P substrate in the C-1 OH position and generates a bisphosphate intermediate which is rotated by 180°. The phosphate group at C-6 OH, closer to Ser66, is transferred to the aminoacidic residue in order to restore the catalytic site of the enzyme. The AGM1 enzyme activity is therefore essentially linked to the presence of the hydroxyl group in C-1 and the phosphate group in C-6: this is the reason why our project is focused on glycomimetics' synthesis with modifications in anomeric and C-6 positions.

3.4. Conclusion

The inhibition that we want to achieve on cancer cells is due to their protract ER stress for accumulation of misfolded protein. This stress is linked to modifications and alteration in glycosylation and the HBP is one of the pathway which produce glycoproteins and is used by cancer cells to proliferate: it's inhibition, due to the decrease of enzymatic activity of AGM1, could represent the way to induce their apoptosis.

PART 2. THEORETICAL FRAMEWORK AND WORKING HYPOTHESES

Chapter 4
Molecular Design

Drug Design, often referred to rational drug design, is that innovative process to discover new medications based on the knowledge of a biological target. This principle is exploited by pharmaceutical chemistry when it needs new drugs.

This chapter represents the starting point of the project describes the structures of *potential inhibitors*.

4.1. Drug Design definition

The latest drug development techniques are based on the knowledge of the target molecule of the disease, from which is possible design a drug that can counteract their activity. The starting point of all biological processes is molecular recognition, such as the formation of a link between two chemical structures on the basis of their complementarity (with a mechanism called key-lock). Reaction between an enzyme (protein acting as catalyst during a chemical reaction in the body) and an organic substrate is the focus point of this thesis project. The natural substrate is exactly complementary to the enzyme and, as it binds, triggers its activation. Through the three-dimensional knowledge of the target of the receptor – enzyme which the drug must interfere with – new molecules can be selected on the base on their similarity to the natural substrate. They are called "lead compounds" and will be the "raw material" on which is possible to work to develop a drug capable of bonding to the therapeutic target.

4.2. Drug Design on AGM1

The process of drug design in this project is based on glycomimetic structures similar to natural substrate of the AGM1 enzyme which is involved in the HBP: the potential inhibitors can be recognized by the enzyme and interfere with its enzymatic activity. The HBP is used by normal and tumoral cells to synthesize UDP-GlcNAc, substrate for glycosylation for proper protein folding. An increase in glycosylation levels and alteration of the expression of enzymes that control such modulation plays an important role in various pathologies such as cellular diabetes (regulation of *O*-GlcNAc involved in type II diabetes), neurodegenerative diseases and finally, in cancer. Hence, an alteration of HBP, and consequently an anomalous strange UDP-*N*-GlcNAc production, can promote proliferation, adhesion, motility, stress resistance, angiogenesis, and ultimately metastasis of cancer cells. [5,36]

[36] AIRC-airc.it/cancro/cos-e/statistiche-tumori-italia

Some studies did in the past on the possible inhibition of another enzyme in HBP show a strategy of inhibition, suggest by *D. J. Vocadlo et al. (2003)* [24], which involves the hijacking of HBP with glycosyltransferase inhibitors: the focus point of HBP is the last step where OGT catalyzes the *O*-GlcNAc glycosylation of a putative protein. This chemical approach may present a limit: the inhibitors which mimic the OGT structure must have a charge to mime the charge of the diphosphate group but this type of charged molecules have many problems to cross cell membranes.

The aim of the research project is to interfere with the HBP through the inhibition of *N-acetylglucosaminephosphate mutase* (AGM1), a key enzyme of the pathway, which catalyzes the conversion of *N*-acetylglucosamine-6-phosphate to *N*-acetylglucosamine-1-phosphate (*Figure 4.1*). The design of a library of potential modulators of this enzyme based on substrate/product structural similarities has been done: some analogues of *N-acetyl-D-glucosamine* (GlcNAc) with modifications on the functional groups directly involved in the catalytic process like the anomeric C-1 and C-6 positions have been designed.

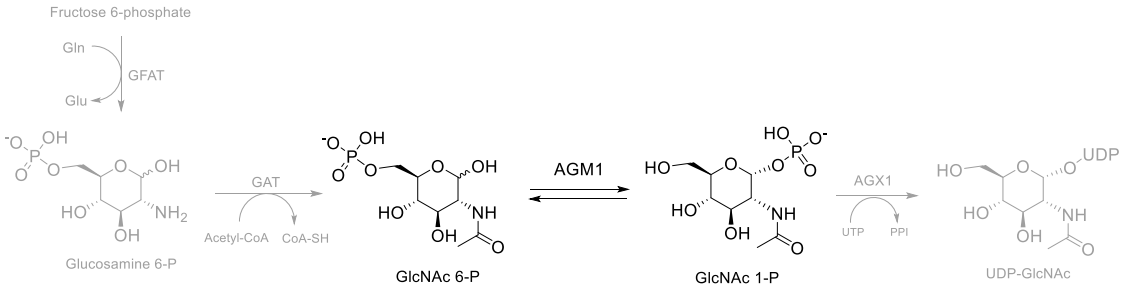


Figure 4.1: illustration of enzymatic reaction of AGM1 in HBP.

The library we want to design will include compounds which present *isosteric group* of phosphate, compounds that have very similar chemical-physical properties to the natural substrate. They can have good affinity, effectiveness and specificity to the target. In this field the compounds are also defined as bioisosters, as they have similar biological activity. The objective is to obtain good pharmacological products in terms of similar solubility and chemical stability. The isosteric groups of the phosphate used to introduce a modification on the sugar structure are shown in *Figure 4.2*.

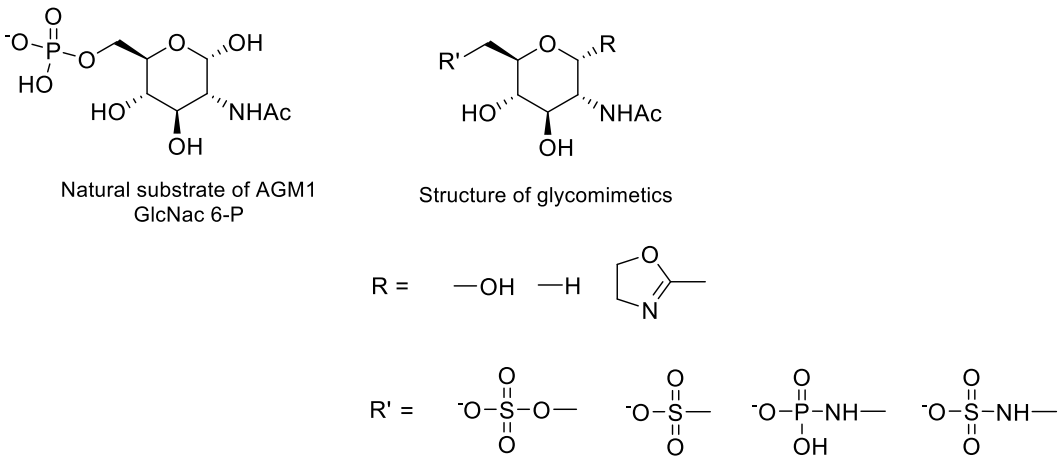


Figure 4.2: illustration of potential inhibitors with isosteric group of phosphate.

4.3. Conclusion

The first approach of the project is the design of molecules which glycomimetics structures similar to the natural substrate of the enzyme AGM1. On the other hand they contain isosteric group of phosphate in positions C-1 and C-6 in order to inhibit the catalytic reaction.

Chapter 5
Computational Analysis

In this chapter will be described computational technique to get a virtual screening on the library of compounds which have been designed as taylor-made structures: in this manner will be studied how the structural ligands inhibit the target. A computational approaches that "docks" small molecules into the structures of macromolecular targets and "score" their affinity to binding sites are widely used in hit identification and lead optimization.[37]

In the first part are described the technical parameters of the optimized model on the natural substrate of the target enzyme; while in the second part, the optimized method has been applied on the designed structures to study the interaction between the molecule and the enzyme.

5.1. Introduction to Molecular Docking

Computational methodologies have become a crucial component of many drug discovery programs, from hit identification to lead optimization and structures based virtual screening.[38,39]

The docking process involves the prediction of ligand conformation and orientation (or posing) within targeted binding site. In general, there are two aims of docking studies: accurate structural modeling and correct prediction of activity.[40]

Docking literally means mooring: this term attributed to the molecules indicates how they fit in each other. To get the docking done between two molecules, it is necessary to know their three-dimensional structure: this can be obtained experimentally by X-ray or nuclear magnetic resonance (NMR) or can be constructed theoretically (ab initio). It is also necessary to model the fundamental interactions between atomic constituents of the molecules. These interactions are well approximated by potential interactions used to study the temporal evolution of molecules.

The power of computing of the modern computers and the development of appropriate algorithms help researchers to get docking possible. In order to effectively study structural interactions between molecules (especially between proteins and ligands – *Figure 5.1*), statistical analysis is also used: the experimental structures of known proteins and ligands are used, and can be used as a template.

[37] Bajorath, Jürgen. "Integration of virtual and high-throughput screening." *Nature Reviews Drug Discovery* 1.11 (2002): 882-894.
[38] Walters, W. Patrick, Matthew T. Stahl, and Mark A. Murcko. "Virtual screening-an overview." *Drug Discovery Today* 3.4 (1998): 160-178.
[39] Langer, T., and R. D. Hoffmann. "Virtual screening an effective tool for lead structure discovery." *Current pharmaceutical design* 7.7 (2001): 509-527.
[40] Kitchen, Douglas B., et al. "Docking and scoring in virtual screening for drug discovery: methods and applications." Nature reviews Drug discovery 3.11 (2004).

For this reason, it has been necessary, in recent years, to develop a large algorithms database (Protein Data Banks - PDB) and suitable programming languages for quick database searches. Proteins interaction are extremely complex because they are flexible objects: they can absorb different conformations over time due to thermal fluctuations. This flexibility allows proteins to interact effectively between them.

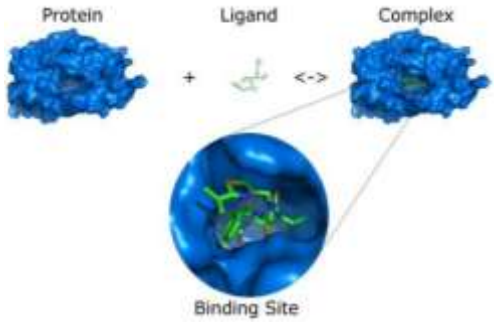


Figure 5.1: complex due to the interaction between protein and ligand.

From a computational point of view, however, it is difficult to take effectively in account the conformational layout/structure. Despite the difficulties, docking is extremely useful to study interaction between proteins and small ligands, e.g. drugs. The aim of the calculation is optimize the state in which the free energy of the protein-ligand system has the lowest value and therefore is the most stable among the configurations.

Molecular docking is a key tool in the structural molecular biology and computer-assisted drug design. The goal of ligand-protein docking is to predict the predominant way of binding between a ligand and a protein of known three-dimensional structure. Successful docking methods effectively search high-dimensional spaces and use a scoring function that correctly ranks docking candidates.[41]

5.2. Set up of the model for the Molecular Docking

The first important assessment was to understand if the pose of the substrate was right: in fact during the docking of the natural substrate isolated from the co-crystallized form we had found low energy conformations, but is prominent the comparison between both two poses. In the *Figure 5.2* are showed the poses of the natural substrate inside the protein co-crystallized (green molecule) before the preparation of the ligand and the same substrate after the preparation (grey molecule): the molecules get similar position and the RMS is 1,40 Å so the model is ready to be used for the docking of the potential inhibitors (all the experimental data described in *Appendix – A.1*).

[41] Śledź, Paweł, and Amedeo Caflisch. "Protein structure-based drug design: From docking to molecular dynamics." *Current opinion in structural biology 48* (2018): 93-102.

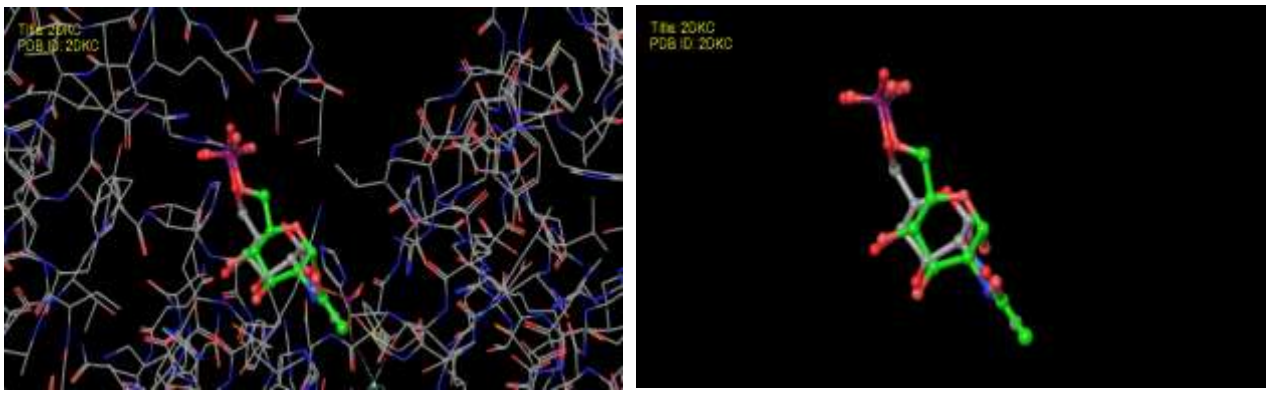


Figure 5.2: pose of the natural substrate of the enzyme and the substrate after computational analysis with and without the catalytic site.

After the grid preparation, the first part of the computational project is the evaluation of standard structures to use in order to understand the goodness of docking. The natural substrate of AGM1 has the hydroxyl group de-protected because of the best interaction with the enzyme (*Figure 5.3*). In fact it has the highest value of interaction.

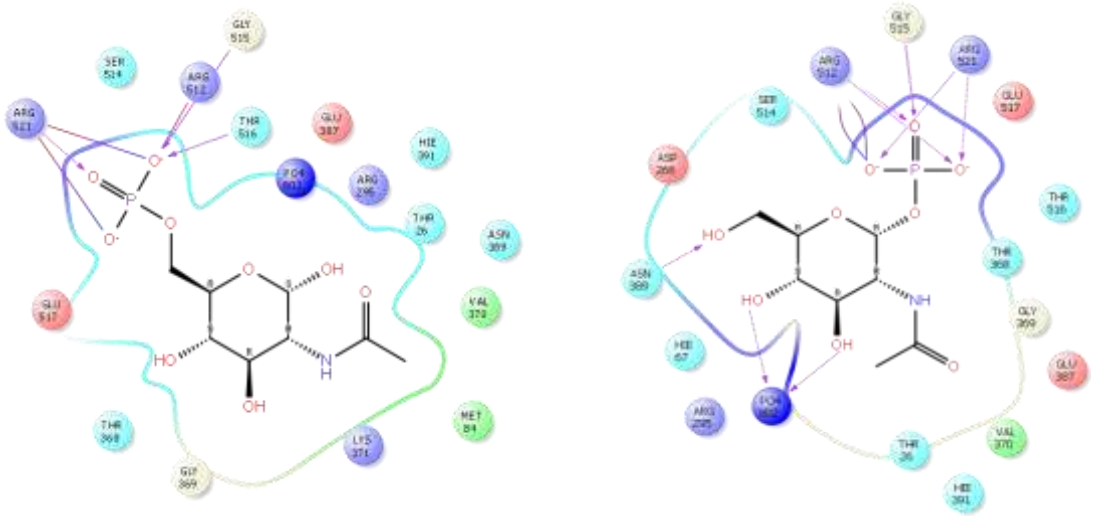


Figure 5.3: on the left the pose of the natural substrate of enzyme AGM1; on the right the product of the catalytic reaction of AGM1.

The "standard" molecules (*Figure 5.4*) used to set the model are:

- The major value of docking score belongs to natural substrate **GlcNAc-6-P** (8,013 kJ/mol), followed by the product of enzymatic reaction **GlcNAc-1-P** (7,318 kJ/mol). All the docking score of the "standard" molecules are compared to the natural substrate.
- GlcNAc-6-P peracetylated (*Figure 5.5 (A)*): it gets a value of docking score of (6,960 kJ/mol) slightly lower than the natural substrate GlcNAc-6-P, because the acetylation of the hydroxyl groups does not favor the interaction with the active site.

- GlcNAc (*Figure 5.5 (B)*, without the phosphate group, gets a decrease in the docking score (5,759 kJ/mol) due to the lowest interactions (H-bond) of the ligand with the aminoacidic residue in the active site.
- GlcNAc peracetyated (*Figure 5.5 (C)*: the missing phosphate group decreases the interaction between the molecule and the protein. Furthermore, the presence of acetyl group leads to the worst value of docking score (4,134 kJ/mol).

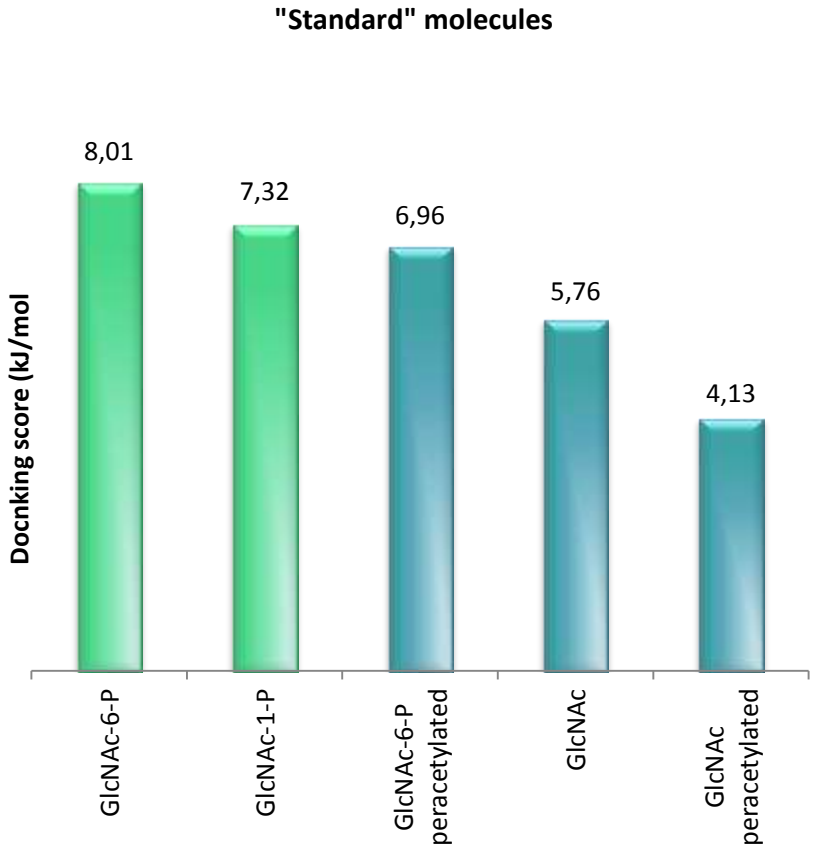


Figure 5.4: histogram of standard molecules to set up the computational method developed.

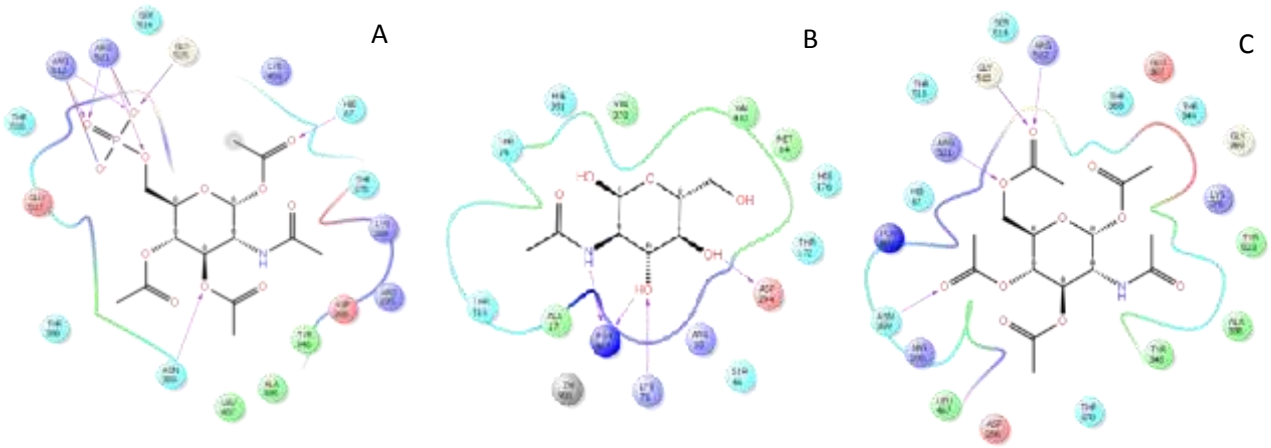


Figure 5.5: (A) pose of GlcNAc-6-P peracetylated, (B) GlcNAc and (C) GlcNAc peracetylated.

The optimized model has been used for the docking of others potential inhibitors. In *Figure 5.6* are shown the structures of the potential inhibitors in de-protected form: as it is reported in the preliminary model, the interactions with enzyme AGM1 are better if the molecule has hydroxyl groups.

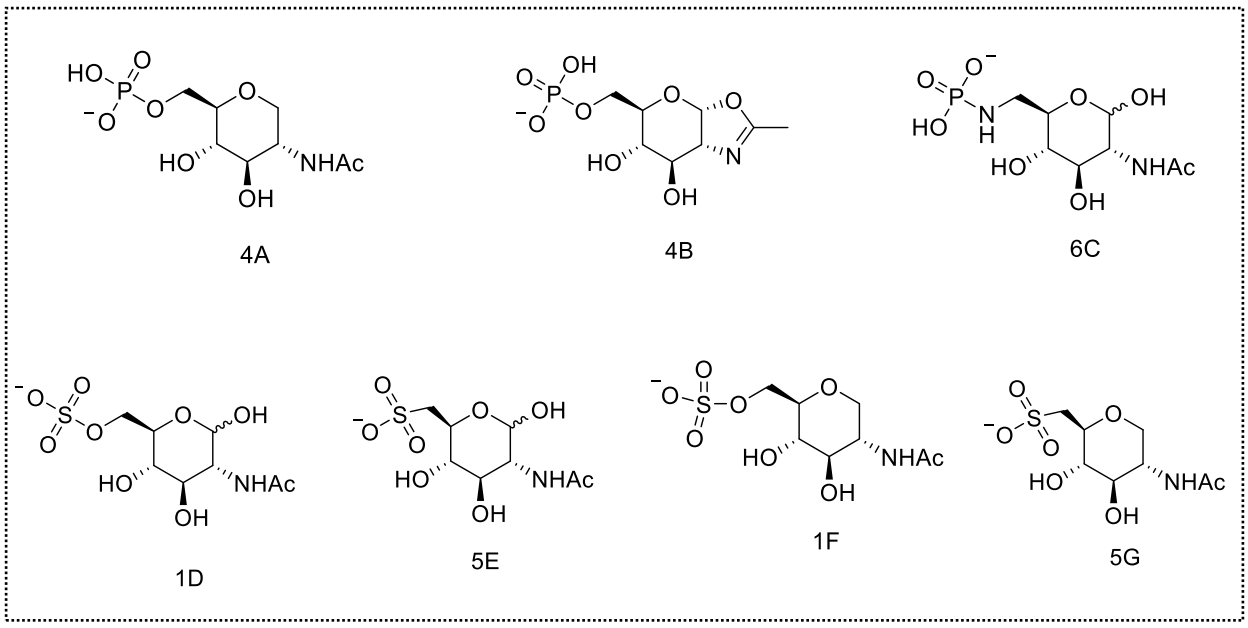


Figure 5.6: structures of potential inhibitors synthesized. N.B: the structures for docking of molecules 4A and 4B are phosphorylated because the interaction with enzyme AGM1 needs its presence.

The structures analyzed with the computational method are in two different forms: the de-protected and the acetylated one (*Figure 5.7*) in order to understand the real interaction with structures with hydroxyl groups. In the graph below is shown that the de-protected form has a better value of docking score than the acetylated one: this fact is due to the interaction between protein and inhibitors is better when the molecule has no acetylated hydroxyl group.

The evaluation and ranking of predicted ligand conformation is a crucial aspect of structure-based virtual screening. Some of the potential inhibitors docked get better value of energy than the natural substrate: it means that they have a good interaction in the active site of the enzyme.

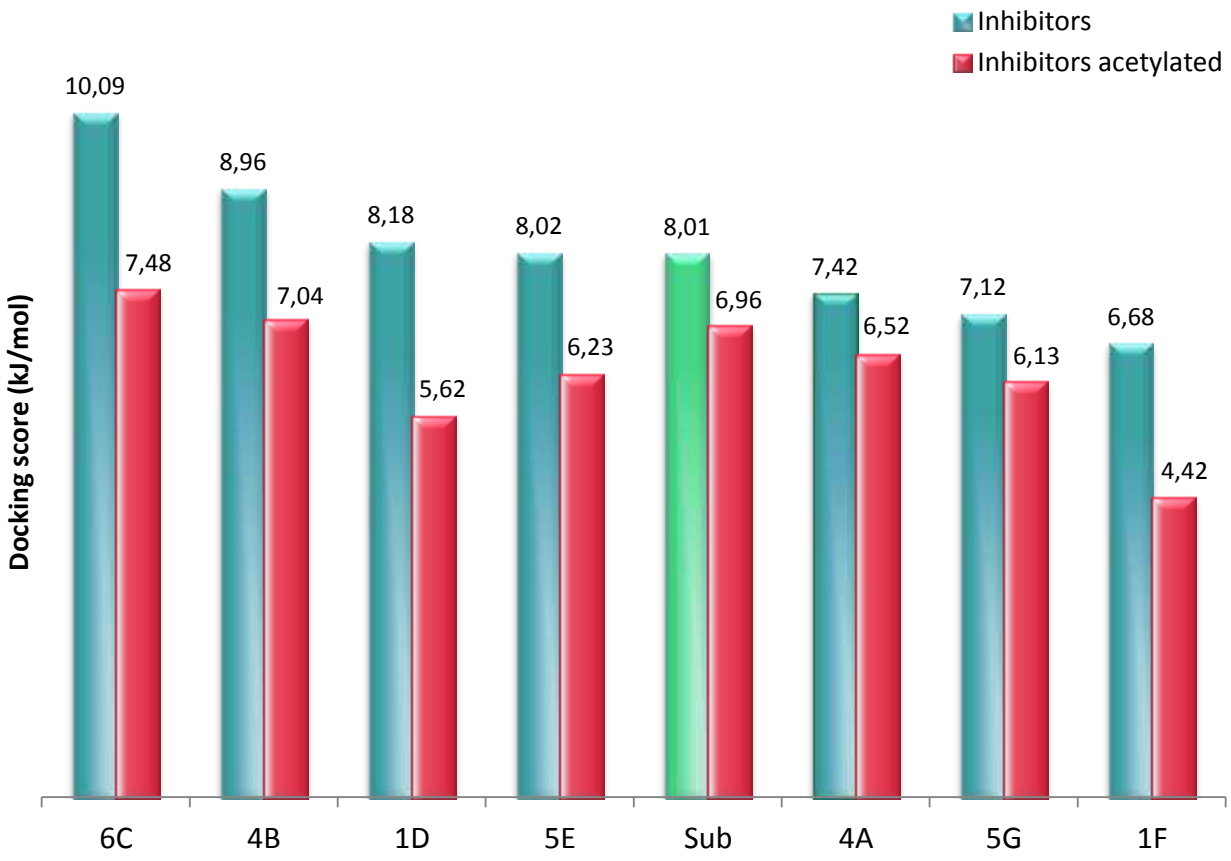
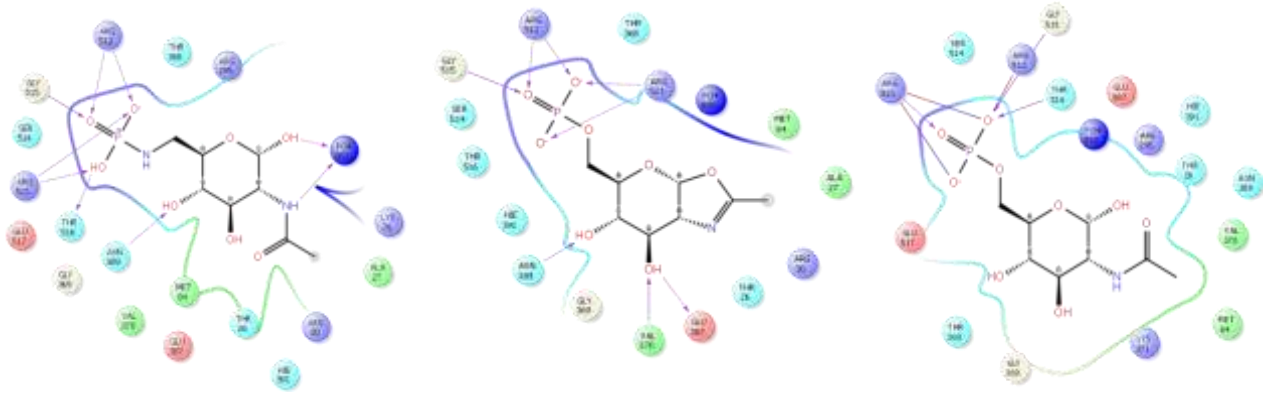


Figure 5.7: histogram of molecular docking of the potential inhibitors, in red the acetylated form, in blue the de-protected form, in green the natural substrate.

In *Figure 5.8* are displayed structures of molecules 6C and 4B which get higher interaction than the natural substrate:



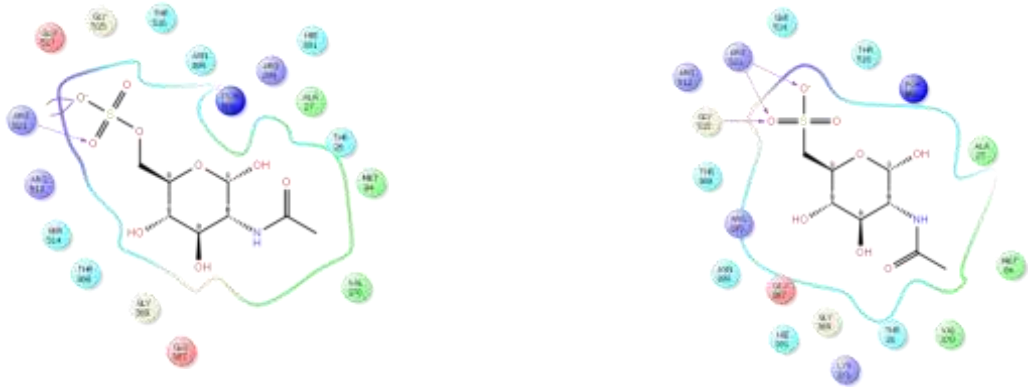
Molecule 6C (10,09 kJ/mol) Molecule 4B (8,96 kJ/mol) Natural substrate (8,01 kJ/mol)

Figure 5.8: poses of de-protected form of molecules 6C and 4B in comparison with the natural substrate.

The interactions of the molecules 6C and 4B with the binding site of the enzyme makes the complex ligand-proteins more stable, in particular molecule 6C shows interaction between hydroxyl group in position C-4 and aminoacidic residue ASN-389.

Furthermore, nitrogen atom of acetammide group and hydroxyl group in position C-1 give interactions with PO4-802. In this way this structures has more affinity with the active site of the enzyme than the natural substrate with a docking score of 10,09 kJ/mol versus 8,01 kJ/mol of the natural substrate. On the other hand, molecule 4B has interactions between hydroxyl group in position C-3 and C-4 with VAL-370 and GLU-387 and ASN-389. This interaction stabilized the complex with score 8,96 kJ/mol.

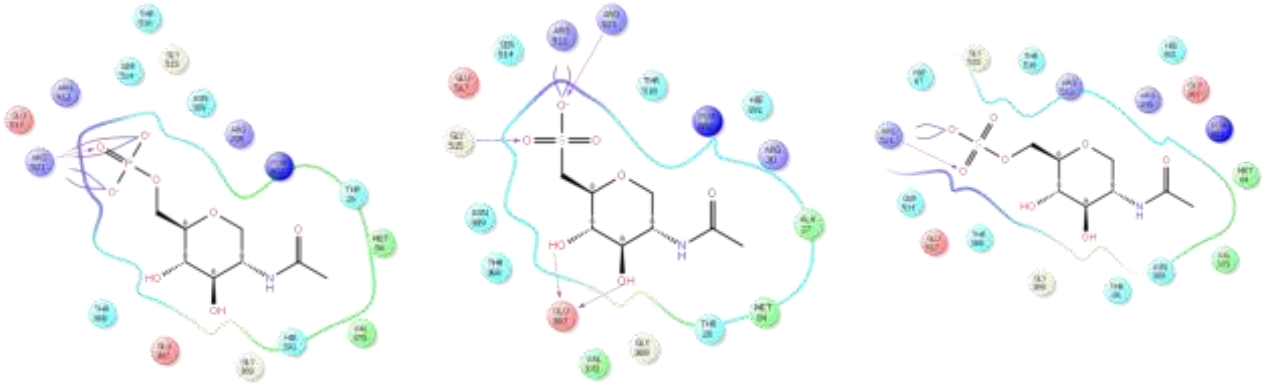
Molecules 1D and 5E have similar scores to the natural substrate because of their ligand-protein interaction is comparable with it (*Figure 5.9*):



Molecule 1D (8,18 kJ/mol) Molecule 5E (8,02 kJ/mol)

Figure 5.9: poses of de-protected form of molecules 1D and 5E.

Finally, molecules 4A, 5G and 1F, which poses are represented in *Figure 5.10*, have low value of docking scores clearly due to their minor interaction with aminoacidic residue of the active site of enzyme.



Molecule 4A (7,42 kJ/mol) Molecule 5G (7,12 kJ/mol) Molecule 1F (6,68 kJ/mol)

Figure 5.10: poses of de-protected form of molecules 4A, 5G and 1F.

5.3. Conclusion

Docking is often used to predict the binding of a small pharmacologically active molecule to its target protein in order to predict the affinity and activity of this molecule; therefore plays an important role in the rational design of the drug.

Docking calculation has been applied to a virtual screening on protein-ligands interaction in order to know the virtual structure similarity with the natural substrate of enzyme and identify potential active compounds. The interplay between docking and scoring functions is fairly complex, but it is easier to produce reliable models of bound ligands than to distinguish true ligand from "false-positive".

An interaction between a small molecule and enzyme/protein can trigger or inhibit the enzyme. Docking is commonly used in the field of drug design. The studies carried out through the Molecular Docking on the compounds of the "drawn" library, give us some information about their affinity and interaction with the active site of the enzyme. In particular structures 6C and 4B seem to get better interaction than the natural substrate, molecules 1D and 5E have docking score similar to the substrate and the other three structures get worst interactions than the natural substrate. These preliminary results are used in order to identify those structures which are most likely capable to bind to the target drug: the information obtained by theoretical calculus will be compared to the experimental results from the in vitro test in order to get a comparison between theoretical and experimental data.

Chapter 6
Synthetic Strategies

This chapter describes synthetic strategies used to obtain the *potential inhibitors*: starting from the *N*-acetylglucosamine or the glucosamine hydrochloride, the synthetic strategy consists on different steps of reaction in order to modify C-1 and C-6 positions on the sugar by introducing phosphate isosteric groups.

6.1. Library of potential inhibitors

In *Table 6.1* are showen the potential inhibitors synthetized in the acetylated form for cellular tests and in the de-protected form which will be used for the enzymatic test.

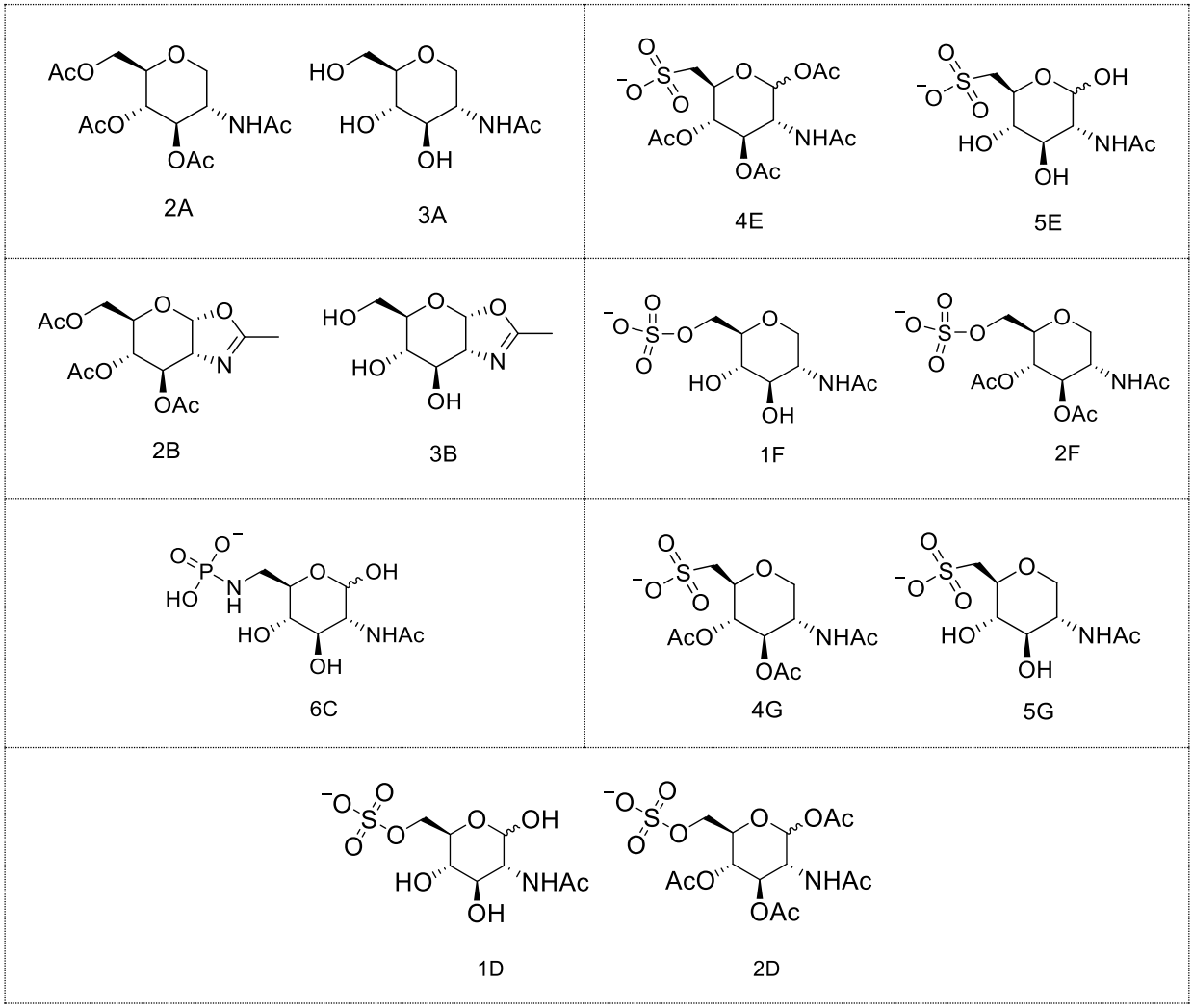


Table 6.1: structures of potential inhibitors synthetized.

6.2. Synthetic Strategy

The synthetic schemes of the potential inhibitors are described in *Figure 6.2*. The general methodology employed for the synthesis of glycomimetics of natural substrate *N*-acetylglucosamine utilizes different reactions in order to introduce a isosteric group of phosphate.

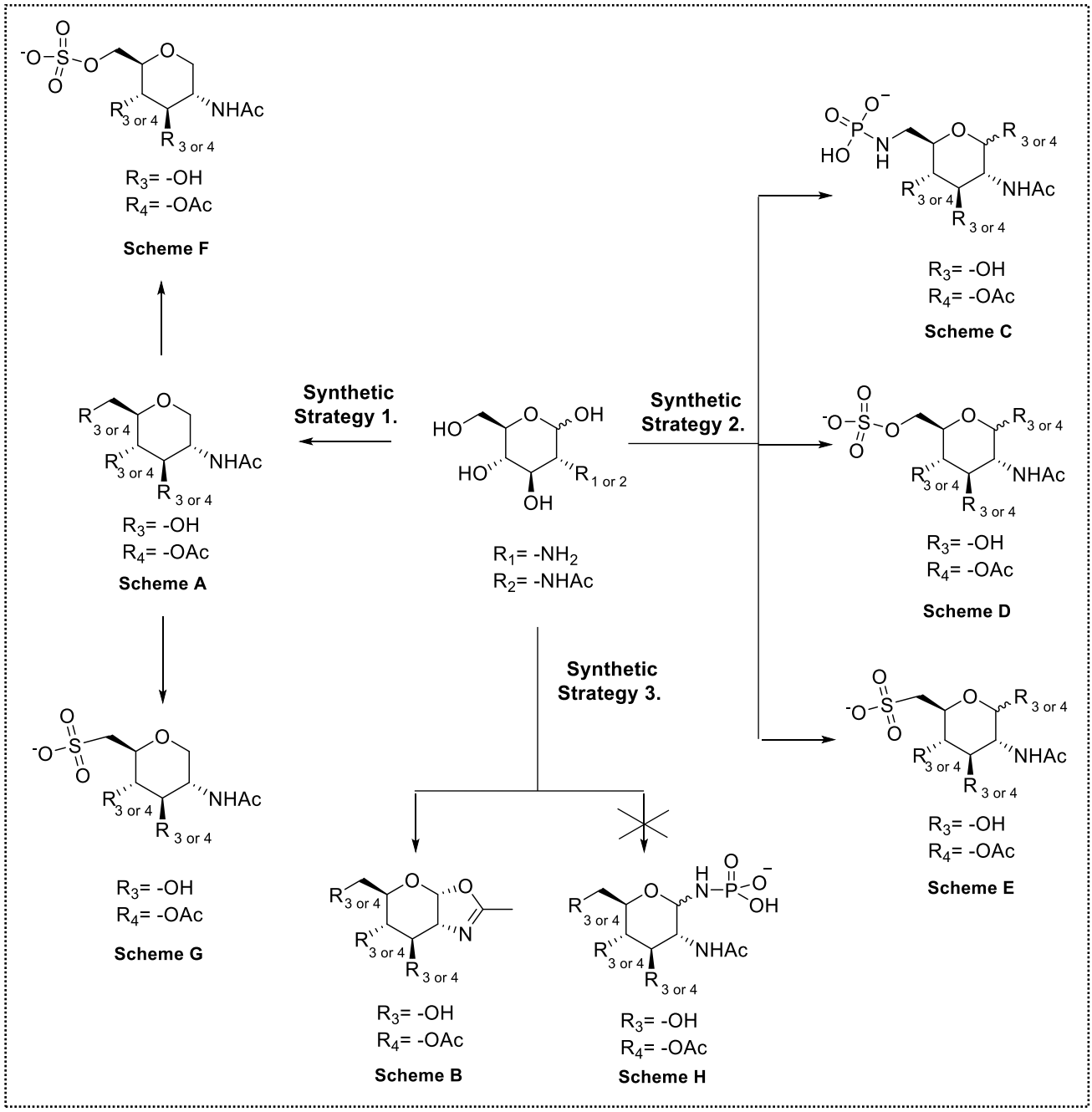
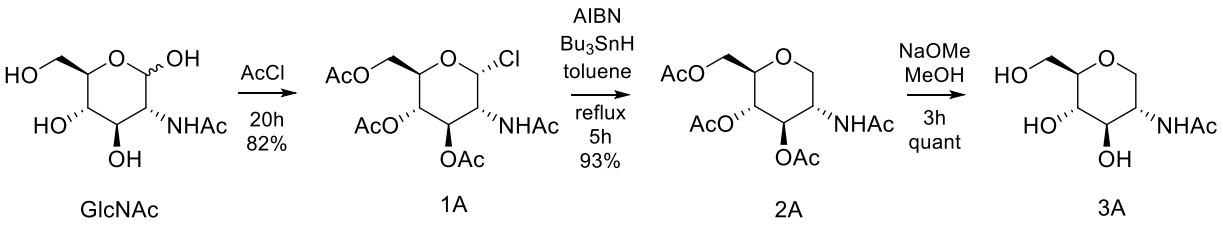


Figure 6.2: schemes of synthesis.

6.2.1. Synthetic Strategy 1.

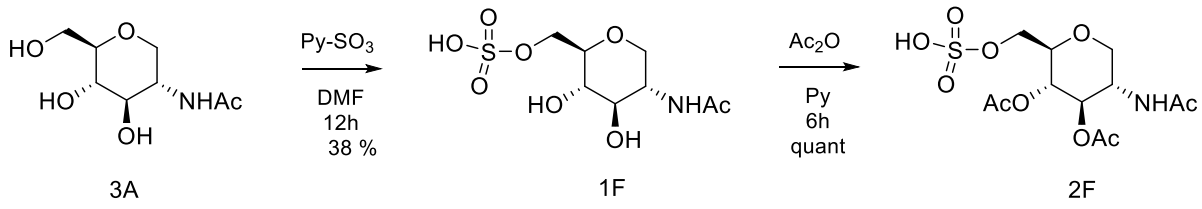
The strategy used in this synthesis brings to a first modification in position C-1, which hydroxyl group has been removed, and then a secondary modification in C-6 position: the introduction of sulphate or sulphonate group cause an alteration of structure.

Scheme A - *N*-acetylglucosamine-1-deoxy from 6-phosphoramidate



The scheme A is the starting point of modification in order to remove the hydroxyl anomeric group and obtain compound 3A[42]; the synthetic strategy starts from the commercially available *N*-acetylglucosamine. Starting material is first reacted with acetylchloride with simultaneously acetyls the hydroxyl groups and introduce the halogen atom on the C-1 (compound 1A): α-1-chloro-1-deoxy-*N*,3,4,6-tetracethylglucosamine. Subsequently, through a radical reaction with AIBN, it's possible to remove the halogen atom affording the corrisponding 1-deoxy-*N*,3,4,6-tetracethylglucosamine (2A): this compound will be used for cellular tests. Compound 2A is then deacetylated affording the 1-deoxy-*N*-acetylglucosamine (3A) which is used for the synthesis of the other 1-deoxy derivatives. Starting from compound 3A is possible to obtain compound 1F (scheme F) and 5G (scheme G).

Scheme F - 6-sulfate derivative from 1-deoxy *N*-acetylglucosamine

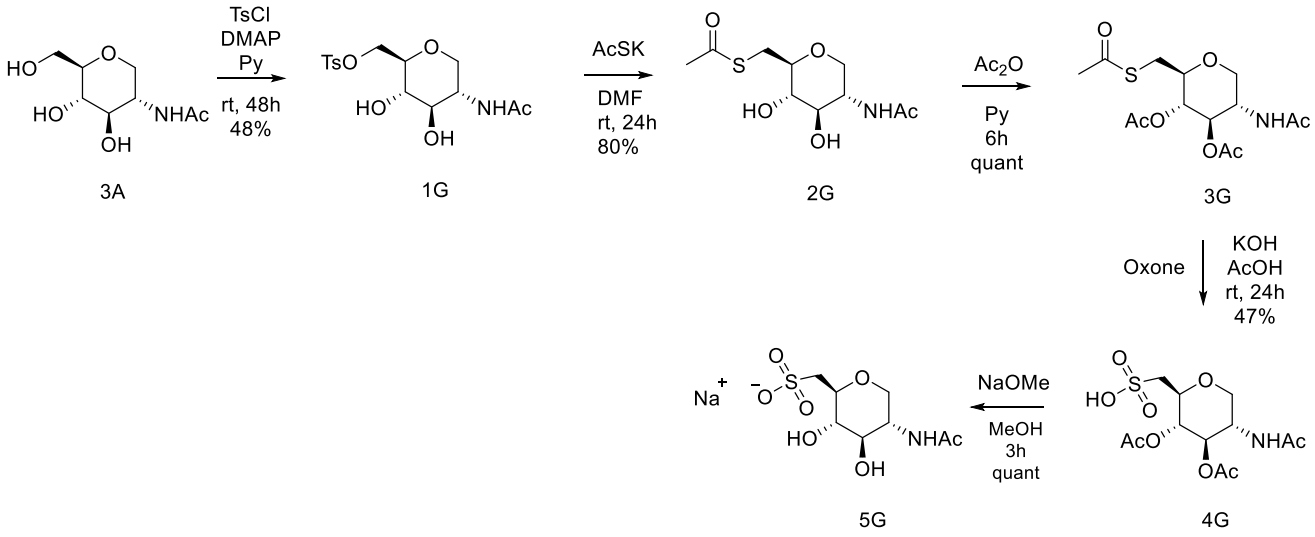


In order to synthetize compounds 1F and 2F we have optimized the synthesis starting from 1-deoxy *N*-acetylglucosamine (compound 3A) previously synthesized. The first reaction with sulphur trioxide pyridine complex introduces the sulphate group[43,44,45] in C-6 O-H and leads to the final compound for the enzymatic test. After acetylation compound 2F is formed and it's ready for the "in vivo" test.

[42] M. Chandler, M. J. Bamford, R. Conroy, B. Lamont, B. Patel, B., Patel, V. K. Williamson, C., *Journal of the Chemical Society, Perkin Transactions* (1995), 9: 1181-1188

[43] Wong, Chi Huey, et al. "Regeneration of NAD (P) H using glucose-6-sulfate and glucose-6-phosphate dehydrogenase." *The Journal of Organic Chemistry* 46.23 (1981): 4676-4679.

Scheme G - 6-sulfonate derivative from *N*-acetylglucosamine-1-deoxy



The synthesis of the sulphonate at C-6 position involves the first protection of the C-6 hydroxyl group by the tosyl starting from compound 3A.[46] After the substitution carried out by the potassium thioacetate and the next acetylation of the sugar, we obtain the compound 3G. Treatment with oxone salt introduces the sulphonate group on the molecule[47] through the oxidation of the thioacetate and its purification gives the final compound 4G for the "in vivo" test.

The reaction with sodium methoxide will generate the final compound 5G for the enzymatic test.

[44] Guiseley, Kenneth B., and Paul M. Ruoff. "Monosaccharide Sulfates. I. Glucose 6-Sulfate. Preparations, Characterization of the Crystalline Potassium Salt, and Kinetic Studies1." *The Journal of Organic Chemistry* 26.4 (1961): 1248-1254.

[45] Horrobin, Tina, Chuong Hao Tran, and David Crout. "Esterase-catalysed regioselective 6-deacylation of hexopyranose per-acetates, acid-catalysed rearrangement to the 4-deprotected products and conversions of these into hexose 4-and 6-sulfates." *Journal of the Chemical Society, Perkin Transactions* 1 6 (1998): 1069-1080.

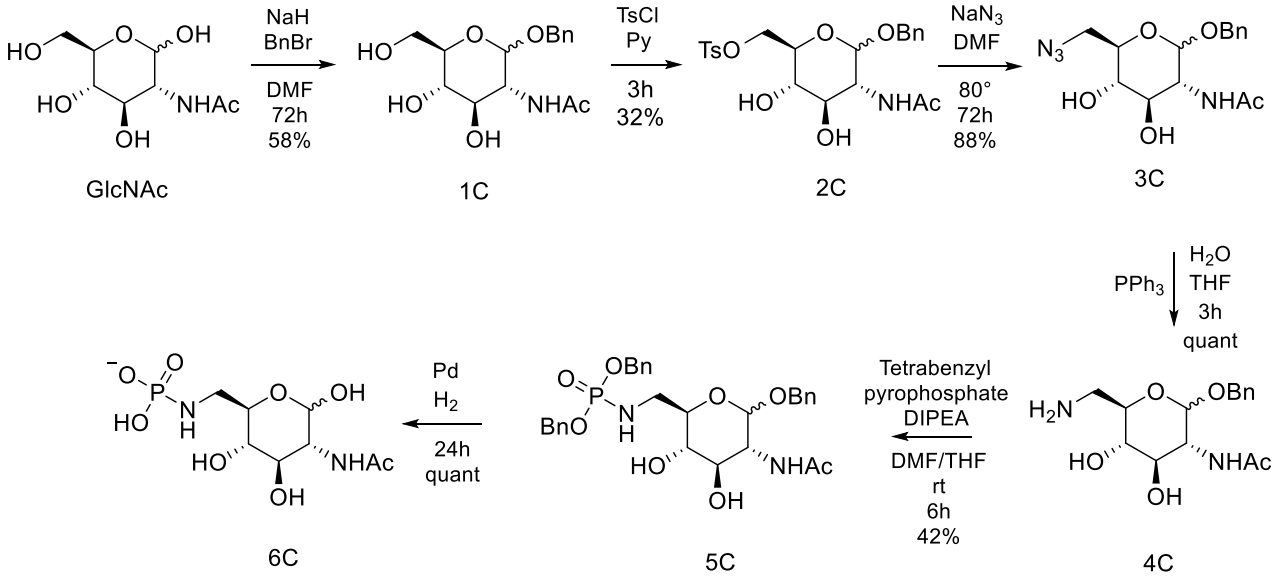
[46] Kato, Takayuki, Andrea Vasella, and David Crich. "Stereospecific synthesis of methyl 2-amino-2-deoxy-(6S)-deuterio-α, β-d-glucopyranoside and methyl 2, 6-diamino-2, 6-dideoxy-(6R)-deuterio-α, β-d-glucopyranoside: Side chain conformations of the 2-amino-2-deoxy and 2, 6-diamino-2, 6-dideoxyglucopyranosides." *Carbohydrate research 448* (2017): 10-17.

[47] Kværnø, Lisbet, et al. "Synthesis and in vitro evaluation of inhibitors of intestinal cholesterol absorption." *Journal of medicinal* chemistry, 48.19 (2005): 6035-6053.

6.2.2. Synthetic Strategy 2.

The synthetic strategy 2. is used to get modification only in position C-6 introducing phosphoramidate, sulphate or sulphonate group.

Scheme C - 6-phosphoramidate derivate from *N*-acetylglucosamine

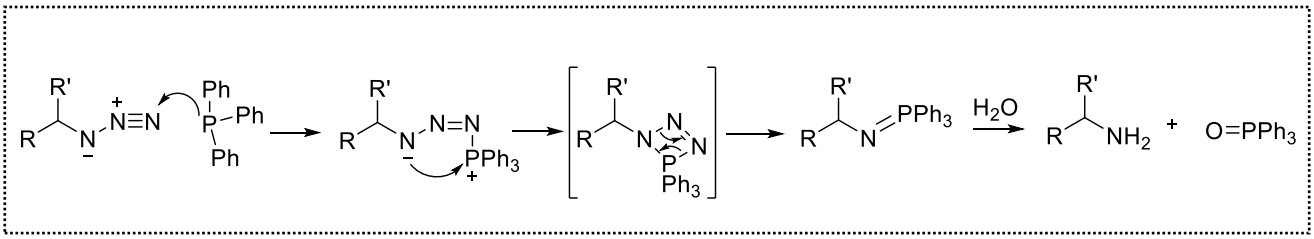


The synthesis of the phosphoramidate in C-6 position involves different reaction steps starting from *N*-acetyl-D-glucosamine. The first reaction is a regioselective protection of the anomeric hydroxyl with a benzyl group.[48] This is achieved using sodium hydride and benzylbromide because next reaction could work on OH in position C-1 or C-2 so the protection of anomeric position directs the next tosilation only in position C-6. The selective protection of the C-6 OH with tosyl generates to compound 2C.[49] After a substitution reaction with sodium azide and the reduction of the azido group with triphenylphosphine, the C-6 amino compound 4C is obtained.

Reaction with the tetrabenzylpirophosphate[50] lead to the protected phosphoramidate (compound 5C) which is deprotected through a catalytic hydrogenation reaction (Pd/H_2) in order to remove the benzyl group on the dibenzylphospharamidate in C-6 and on C-1 positions. The compound 6C is ready for the enzymatic test.

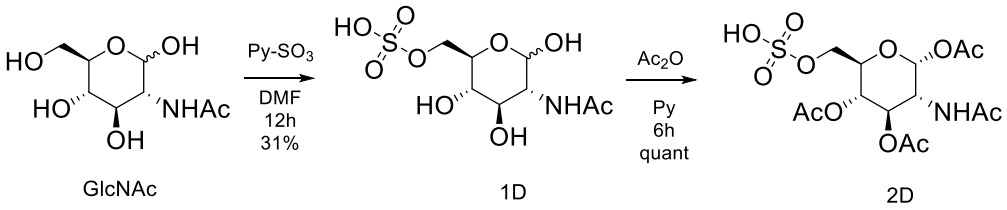
[48] Adam R. Yeager and Nathaniel S. Finney, *J. Org. Chem.* (2005), Vol. 70, No. 4, 1270 - 1275
[49] Thomson, Robin, and Mark von Itzstein. "Synthesis of 4-substituted-2-acetamido-2, 4-dideoxy-mannopyranoses using 1, 6-anhydro sugar chemistry." *Carbohydrate research* 274 (1995): 29-44.
[50] Knapp, Spencer, et al. "A family of mycothiol analogues." *Journal of carbohydrate chemistry* 24.2 (2005): 103-130.

The azido group is then reduced through a *Staudinger reaction*, in the presence of triphenylphosphine, (*Mechanism 6.1*). The mechanism of the reaction is focused on the formation of a iminophosphorane through a nucleophilic addition of phosphine on end nitrogen atom of azide, follow by nitrogen expulsion. In the second step, the intermediate is then hydrolyzed to the corresponding amine and phosphine oxide is released.



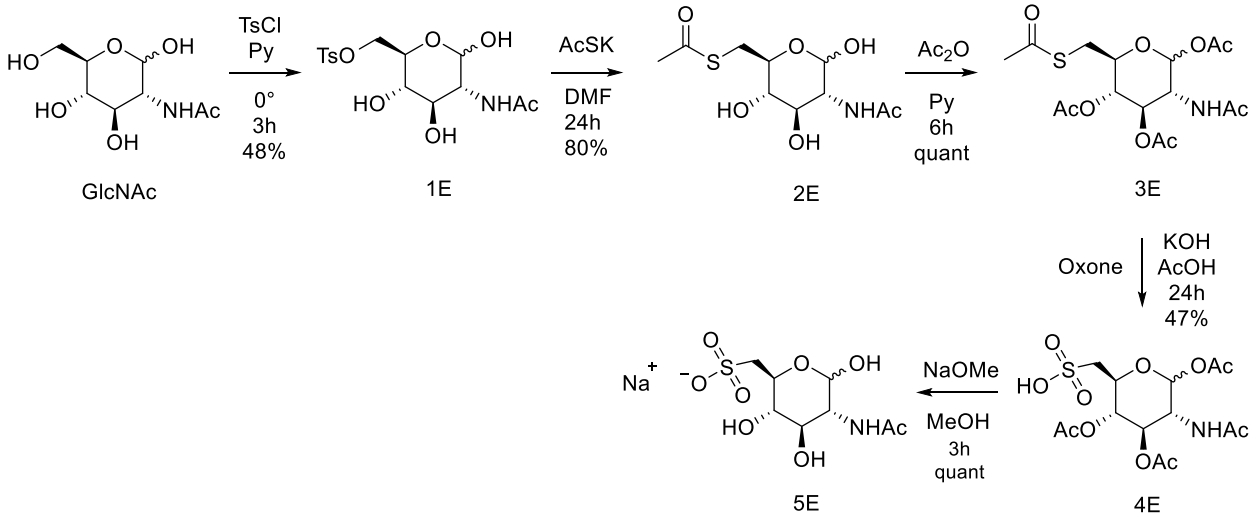
Mechanism 6.1: the mechanism of Staudinger's reaction.

Scheme D - 6-sulfate derivative from *N*-acetylglucosamine



In order to synthetize compounds 1D and 2D we have optimized the synthesis starting from the *N*-acetylglucosamine: the first reaction with sulfur-trioxide complex and pyridine introduces the sulphate group [43,44,45] in C-6 position affording directly compound 1D. After the acetylation reaction, compound 2D is ready to be tested in vivo.

Scheme E - 6-sulfonate derivative from *N*-acetylglucosamine



The synthesis of the sulphonate in C-6 position involves the first regioselective protection of the C-6 hydroxyl group by the tosyl.[51] The tosylate is then trated with potassium thioacetate affording the thioester which is then acetylated affording compound 3E. Oxone salt oxidized the thioester to the corresponding sulphonate[47] through the oxidation of the thioacetate and its purification gives the final compound 5E for the "in vivo" test.

Deacetylation with sodium methoxide affords the final compound 5E for the enzymatic test.

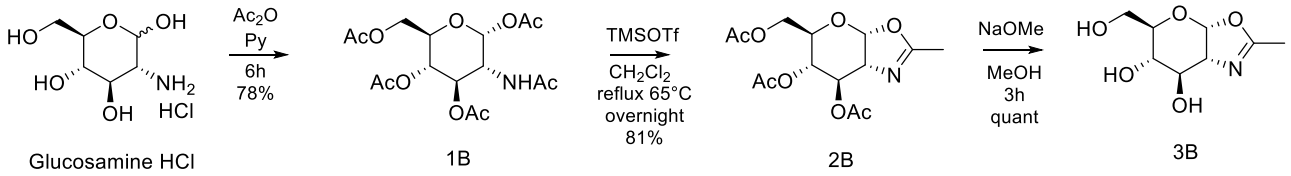
[51] Kabalka, George W., et al. "The tosylation of alcohols." *The Journal of Organic Chemistry 51.12* (1986): 2386-2388.

6.2.3. Synthetic Strategy 3.

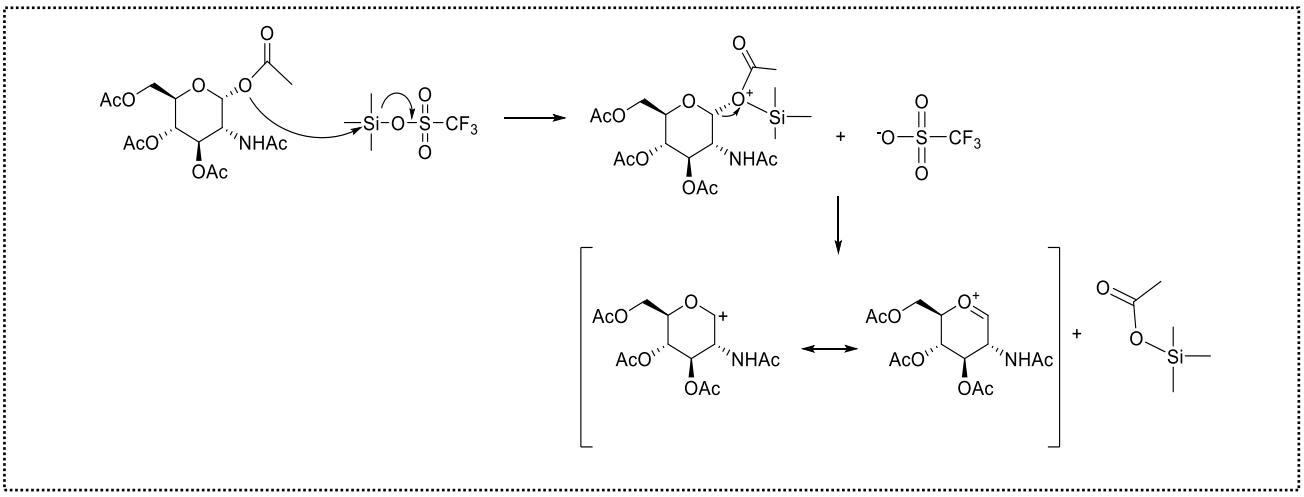
The strategy 3. affords modification in position C-1 through the introduction of the phosphoramidate group or the oxazoline ring.

Scheme B - 2-Methyl-(3,4,6-tri-O-acetyl-1,2-dideoxy-α-D-glucopyrano)-[2,1-d]-2-oxazoline[52]

Scheme B describes the protection with acetylation[53] and a reaction with Lewis acid, followed by de-protection.



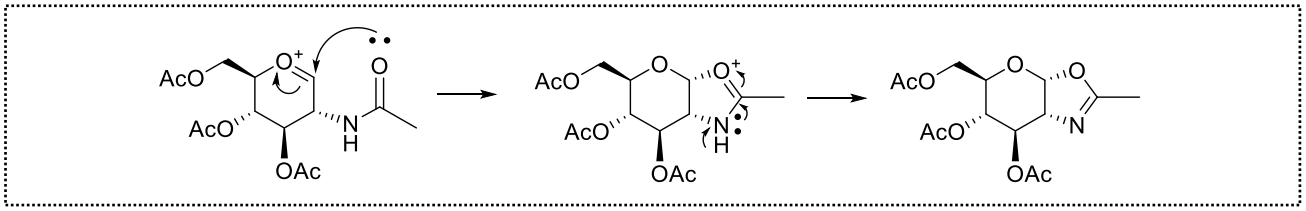
The closure of the heterocyclic ring occurs with the intramolecular attack of the carbonyl oxygen of the acetamide group on the oxonium ion, generated from the release of the anomeric acetate promoted by the Lewis acid (*Mechanism 6.2*):



Mechanism 6.2: the mechanism of formation of oxonium ion.

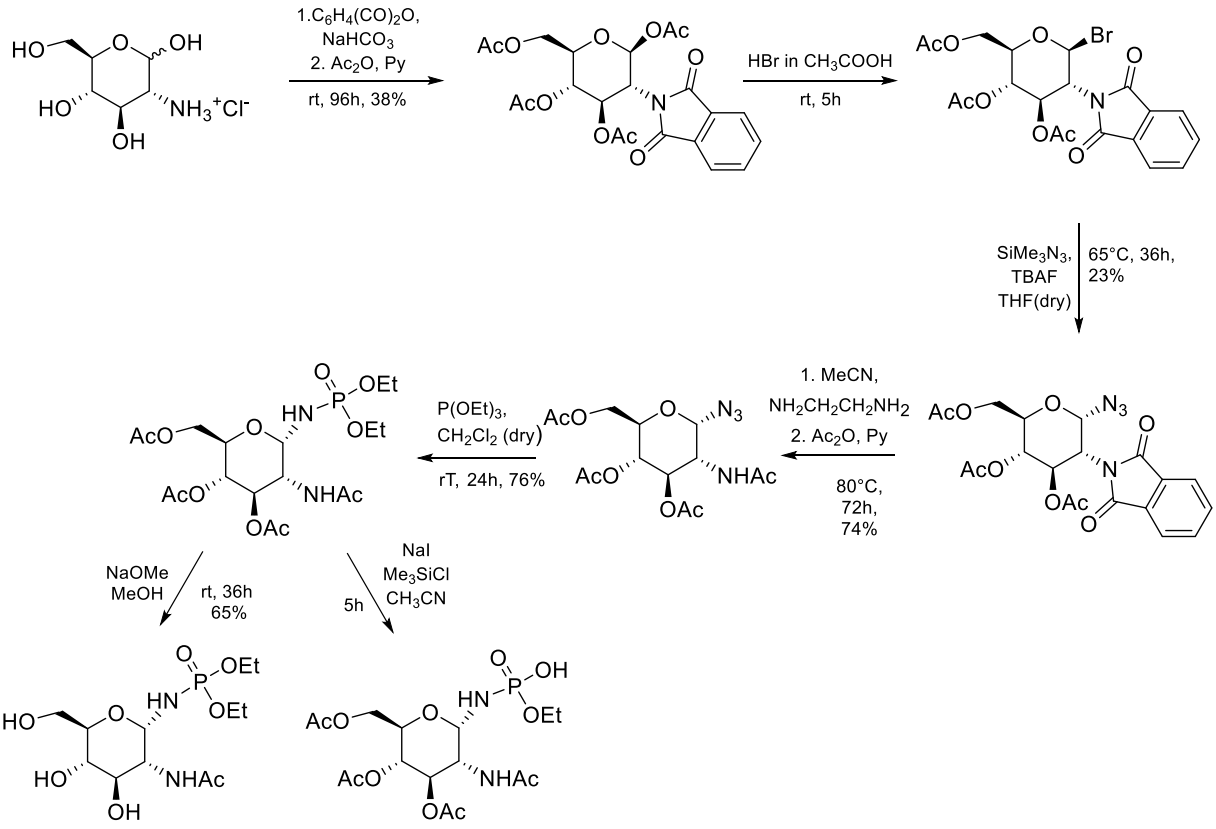
[52] R. Haddoub, N. Laurent, M. M. Meloni and S. L. Flitsch, S. L., *Synlett* (2009), 20, 3328-3332.
[53] B. Bhatt, R. Böhm, P. S. Kerry, J. C. Dyason, R. J. M. Russell, R. J. Thomson, and M. von Itzstein, *J. Med. Chem.* (2012), 55, 8963−8968

The resonance stabilized oxonium ion reacts with carbonyl group of the acetamide through an intramolecular reaction to give the glycofused five-member ring of the methyloxazoline (*Mechanism 6.3*)



Mechanism 6.3: the mechanism of reaction of closure on heterocyclic ring.

Scheme H - 1-phosphoramidate derivative from *N*-acetylglucosamine

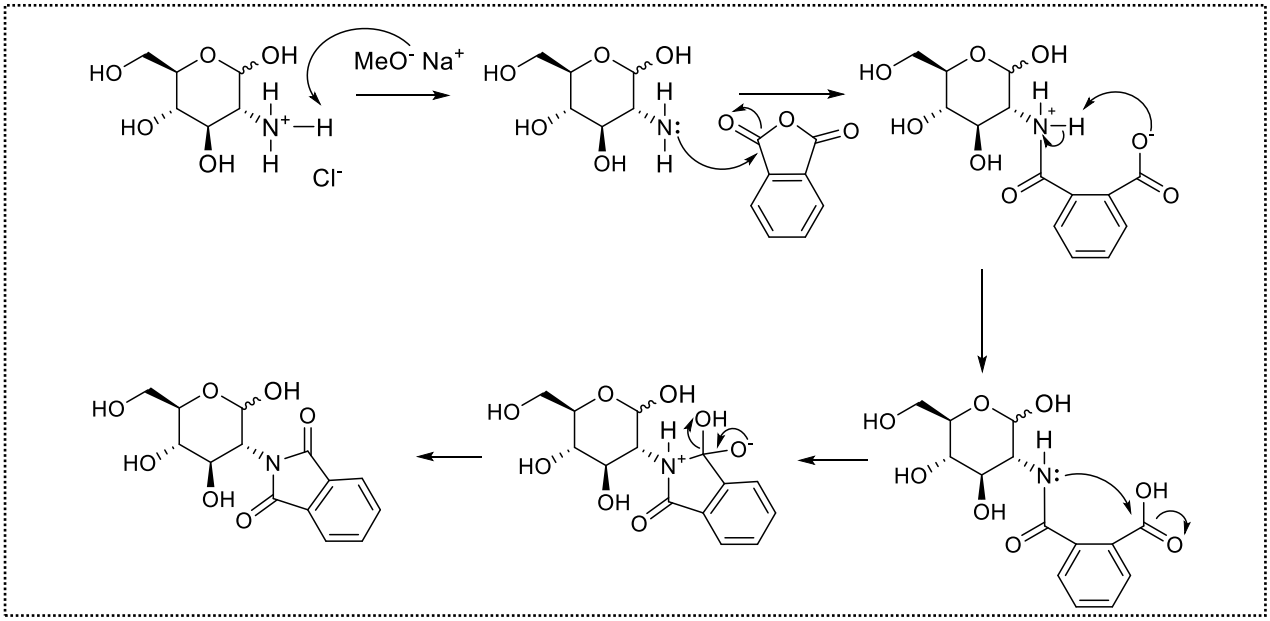


The scheme H has been used to introduce a modification in position C-1 with the introduction of phosphoramidate using a synthetic strategy already setted up by *Bianchi et al. (2006)*.[54] The synthesis starts with a first protection of nitrogen in position 2 of the sugar with the phtalimmide followed by acetylation of the hydroxyl groups. The phatalimmido group has the characteristic to orientate the anomeric acetate in β.

[54] Bianchi, Aldo, and Anna Bernardi. "Traceless Staudinger ligation of glycosyl azides with triaryl phosphines: Stereoselective synthesis of glycosyl amides." *The Journal of organic chemistry* 71.12 (2006): 4565-4577.

Reaction with hydrobromic acid generates the corresponding bromo glycoside with the bromine oriented in β. Subsequently the reactionwith trimethylsylil azide allows the introduction of the azido group in the anomeric position α oriented. Finally, the de-protection of the C-2 amino group and the subsequent reaction with trethylphosphite generates the C-1 diethylphosphoramidate product.[55] The difficulty to hydrolyzed both two esters groups, is still to be faced and hasn't allowed us to obtain the final deprotected product..

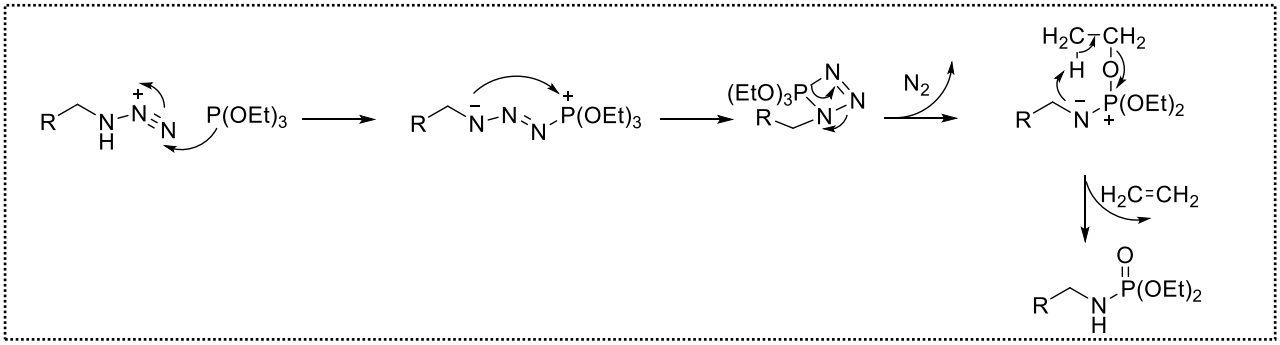
The first step of synthesis is a SN_2 reaction with formation of glucosamine with elimination of NaCl and MeOH as elimination products. Then the reaction with phthalic anhydride allows the protection of the amine formed an open-cycle intermediate. The following deprotonation, in presence of MeONa, of the amide with formation of the carboxylic acid and the coupling of the electronic doublet on nitrogen and carbonyl with closure of the ring formed the product of phthalimide (*Mechanism 6.4*).



Mechanism 6.4: the mechanism of formation of phtalimide.

Interesting is the mechanism of formation of diethylphosphoramidate by *Michaelis-Arbuzov's reaction* (*Mechanism 6.5*). The mechanism involves the nucleophilic attack by the the phosphorus atom of the triethylphosphite which gives a phosphonium intermediate. Nitrogen with negative charge, through a nucleophilic attack, allows the formation of a four-atoms cycle and the P-N bond is cleaved with the elimination of molecular nitrogen. By using a concerted mechanism, diethylphosphoramidate and ethylene are formed.

[55] Kannan, Thanukrishnan, et al. "Synthesis of glycosyl phosphoramidates: novel isosteric analogues of glycosyl phosphates." *Bioorganic & medicinal chemistry letters* 11.18 (2001): 2433-2435.



Mechanism 6.5: the mechanism of Michaelis–Arbuzov's reaction.

6.3. Conclusion

Through different synthetic strategies, the glycomimetics have been modified:

- removal of hydroxyl group on anomeric position and then the introduction of polar group such as sulphate or sulphonate groups;
- synthesis which brings to a modification on anomeric carbon to product the potential inhibitor as oxazoline derivative;
- introduction in position C-6 on sugar *N-acetyl-D-glucosamine* of polar group (sulphate, sulphonate, phosphoramidate)

In this way, the new compound synthetized can mimic the polar character of phosphate group in the natural substrate with the introduction of isosteric groups are similar to the natural substrate of enzyme, *N-acetyl-D-glucosamine 6-phosphate*, but deceive it blocking its activity.

All the compound synthetized have been characterized by NMR spectra (^{1}H and ^{13}C and 2D spectra) and *m/z* (*Appendix A.3*).

Chapter 7
Enzymatic test

This chapter will describes the screening of the potential inhibitors synthetized in order to understand their interaction with AGM1 enzyme in the cellular extract. The developed method use *High Performance Liquid Chromatography* and gives us preliminary results about inhibition capacity of our glycomimetics.

7.1. Overview on *High Performance Liquid Chromatography*

High Performance Liquid Chromatography (HPLC) is an analytical technique used to separate, identify, and quantify each component in a complex mixture. In this case the stationary phase must have different chemical properties depending on the type of separation needed. It's an improved form of column chromatography: instead of a solvent with a flux driven by the gravity, HPLC is subjected to high pressures up to 400 atm which allows to separate very complex mixtures in a short time. The sample to be separated is mixed with the mobile phase and then transported into the column in contact with the stationary phase. The molecule, depending on its chemical structure, will show different affinity for the mobile phase and for the stationary one. The different interaction molecule-stationary phase and molecule-mobile phase leads to different retention time, that is the time needed by the molecule to get out from the column.

The focus point of HPLC system is that the column and the stationary phase must be packed together inside the system. Normally the columns have a length varying from 10 to 30 cm, although lately shorter columns around 3 cm are produced. The pressurized liquid solvent containing the sample mixture is injected into the HPLC and separated through a column filled with an adsorbent solid material.

Chromatography can be described as a mass transfer process involving adsorption phenomenon. HPLC relies on pumps to pass a pressurized liquid and a sample mixture through a column filled with adsorbent, leading to the separation of the sample components. The active component of the column, the adsorbent, is typically a granular material made of solid particles (e.g. silica, polymers, etc.) of 2–50 μm. The components of the sample mixture are separated from each other due to their different retention with the adsorbent particles. The pressurized liquid or "mobile phase" is typically a mixture of solvents: its composition plays the major role in the separation process by influencing the interactions taking place between sample components and adsorbent. These interactions are hydrophobic (dispersive), dipole–dipole and ionic interactions.

7.2. Schematic description of HPLC instrument

The HPLC instrument used for the analysis includes:

- a degasser in order to remove small bubbles which can be transported into the column
- the sampler, with manual loop injection, brings the sample mixture (20µL) into the mobile phase stream which carries it into the column.
- the pumps set the desired flow and composition of the mobile phase through the column.
- the column: reverse phase LiChroCART® 250-4 LiChrospher® 100 RP-18 (5 µm)
- the UV/Vis detector generates a signal proportional to the amount of sample component emerging from the column, hence allowing for quantitative analysis of the sample components. The analyte studied, (UDP-GlcNAc metabolite), is detected by UV/Vis at 254 nm
- the digital microprocessor and user software control the HPLC instrument and provide data analysis.

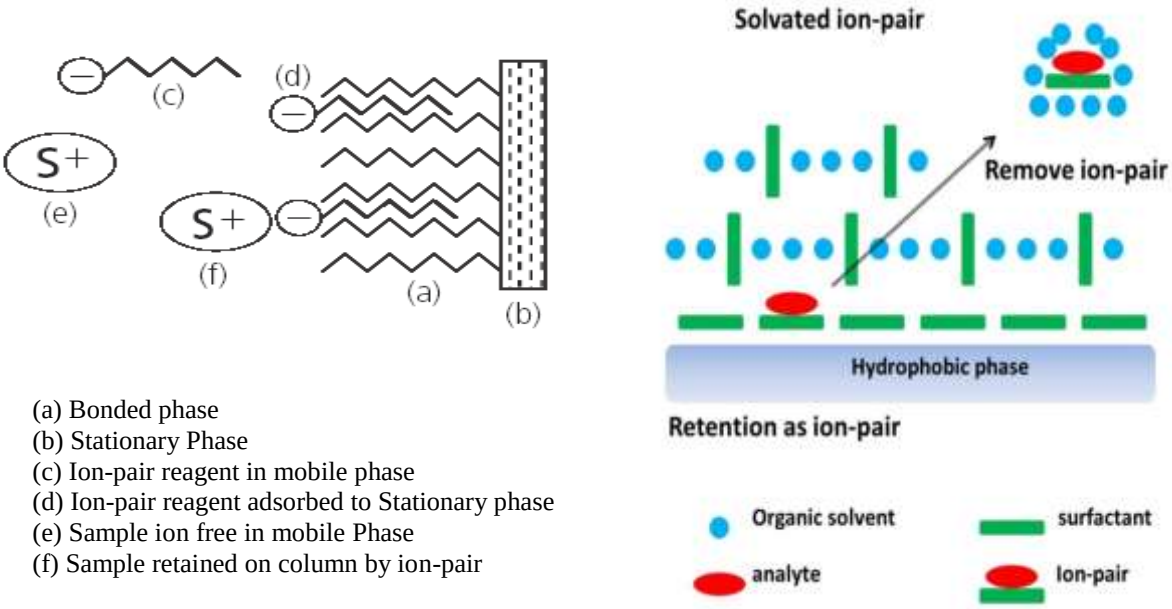
7.3. Ion-pair chromatography

The method used for the analysis was the *ion-pair chromatography* (IPC) in order to affect retention and selectivity of ionic compounds. This strategy is used when a sample containing ionic components tends to be very hydrophilic, so reversed-phase retention can be problematic, and when other changes in reverse-phase chromatography conditions fail to achieve acceptable resolution.

The IPC was developed by Dr. Gordon Schill in 1973[56] to reduce peak tailing for basic solutes, ability to increase retention of weakly retained ionized acids and provide additional options to control selectivity in the separation of ionic samples.

In reversed phase chromatography, ionic compounds are not usually retained by hydrophobic stationary phase. By adding an ion-pair reagent with a ionic head and a hydrophobic tail into the mobile phase, the hydrophobic tail of the reagent is retained by the stationary phase. The sample's ions exchange with the counter ion of the ion-pair reagent retained by the stationary phase, thus resulting in greater retention of the sample.

[56] Cecchi, Teresa. "Ion pairing chromatography." *Critical Reviews in Analytical Chemistry* 38.3 (2008): 161-213.



(a) Bonded phase
(b) Stationary Phase
(c) Ion-pair reagent in mobile phase
(d) Ion-pair reagent adsorbed to Stationary phase
(e) Sample ion free in mobile Phase
(f) Sample retained on column by ion-pair

Figure 7.1: on left description of component IPC and (on right) its mechanism.

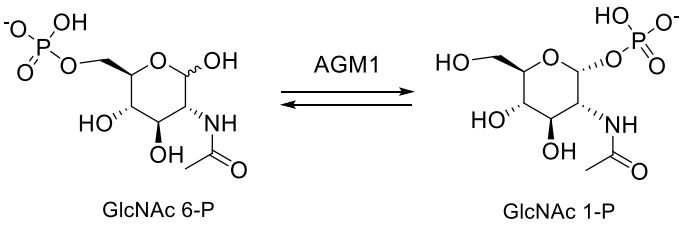
Separation is achieved by virtue of the different constants of formation of the ion pair and of their different degrees of adsorption. In this way ion exchange groups are formed on the surface of the stationary phase (*Figure 7.1*). Although IPC has been traditionally performed in isocratic elucidation mode, some recent examples of gradient mobile phases demonstrate that this strategy is valuable as it leads to high peak capacity, causes rapid separations by improving retention time, avoiding extreme pH conditions and is more selective for ionic compounds that tend to be very hydrophilic. In fact, a reverse phase chromatography consisting solely of a hydrophobic stationary phase fails to retain these compounds. The pH of the eluent is a key factor to control the charge status of the analyte and selectivity can be manipulated essentially by the concentration of IPC and by the percentage of organic additive in the eluent (e.g. ACN). The preferred mobile phase for IPC is an eluent rich in water; the purpose of adding organic additives is to increase the solubility of the analyte, improve the wetting of RP packing and regulate the retention and the selectivity of the analyte.

7.4. Aim of enzymatic test

The enzymatic test developed is important in order to have preliminary “in vitro” results about the interaction and affinity between potential inhibitors and the enzyme.

The focus point of this enzymatic test is the production of the Uridine-diphosphate-*N*-acetylglucosamine (UDP-GlcNAc), the precursor of glycosylation: the inhibition of the HBP pathway leads to a decrease of the production of this metabolite modifying the level of glycosylation; it means that the formation of misfolded protein are accumulated in the ER which activates the UPR (*Chapter 2 - Paragraph 2.4*).

Our first interest was to analyze the potential enzymatic inhibitors by studying the enzymatic reaction using the commercially available purified enzyme. In the literature no enzymatic test based on the enzyme of interest AGM1 is described. Moreover, considering that the enzymatic reaction is reversible, and that in the HBP pathway the product of the enzymatic reaction GlcNAc 1-P is not accumulated but converted to the final product UDP-GlcNAc, we speculated that working with the purified enzyme and quantifying the product GlcNAc 1-P could be problematic.



On the contrary, the quantification of UDP-GlcNAc produced using cellular enzymatic extracts is reported in the literature, and methods based on UV/Vis spectrometry are described.

Taking in consideration what just reported, we developed an enzymatic assay based on the use of an enzymatic extract to perform the reaction and on a separation/quantification methodology based on a HPLC instrument attached to a UV/Vis detector.

The test developed will quantify the UDP-GlcNAc because it is the final product of the coupled enzymatic reaction (*Figure 7.2*): the potential inhibitors are designed on the natural substrate of AGM1, so we will expect the inhibition of this enzyme followed by a decrease of production of GlcNAc 1-P, substrate of the next enzyme AGX1. The activity of AGX1 will be minor because of the minor quantity of substrate. At the end a minor quantity of UDP-GlcNAc will be produced, so we will identify the production of the metabolite using HPLC with UV/Vis detector and quantify the metabolite through the area of the peak in HPLC from the enzymatic reaction.

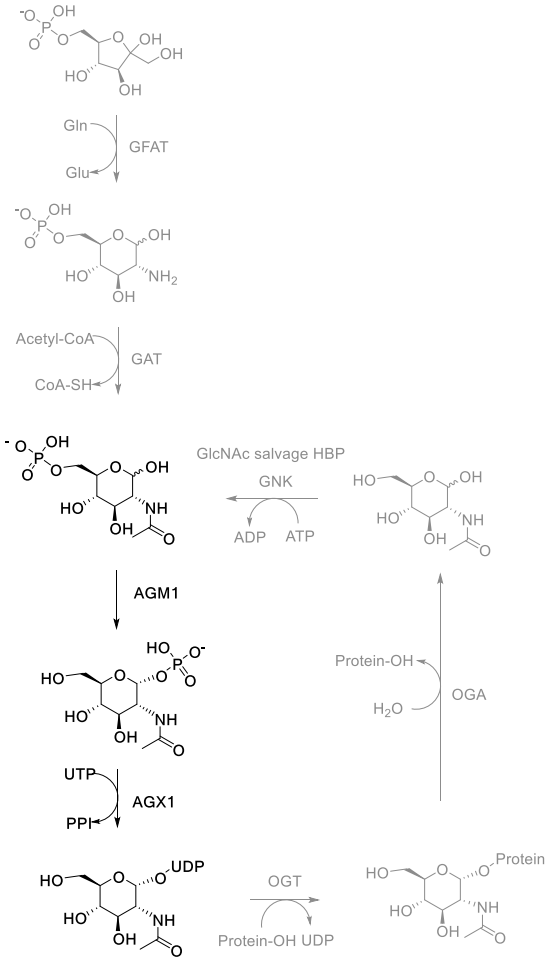


Figure 7.2: description of the coupled enzymatic reactions in HPB in order to get the UDP-GlcNAc.

7.5. HPLC: setting method

The experimental procedure used for both for the step-up of the method and the analysis of biological metabolite in the cellular extract, consists in different parts: 1) the achievement of UDP-GlcNAc calibration line in order to get the quantification of the metabolite; 2) the preparation of enzymatic/cellular extract; 3) the tuning the enzymatic reaction; 4) the set-up of the sample treatment prior HPLC analysis.

- **Analysis UDP-GlcNAc standard and calibration line**

In literature, *Nakajima et al. (2010)* [57], have developed a conventional HPLC method for simultaneous determination of nucleotide sugars (*Figure 7.3*). A mixture of nucleotide sugars (CMP-NeuAc, UDP-Gal, UDP-Glc, UDP-GalNAc, UDPGlcNAc, GDP-Man, GDP-Fuc and UDP-GlcUA) and relevant nucleotides were perfectly separated in an optimized ion-pair reversed-phase mode. The method was newly developed in order to determine the nucleotide sugars in cellular extracts from 1×10^6 cells in a single run. This method had been used to characterize nucleotide sugar levels in breast and pancreatic cancer cell lines. The new HPLC method allowed the authors to separate not only each nucleotide sugar involved in *N*- and *O*-glycan synthesis but also a series of nucleotides, contrary to most of the previously reported methods.

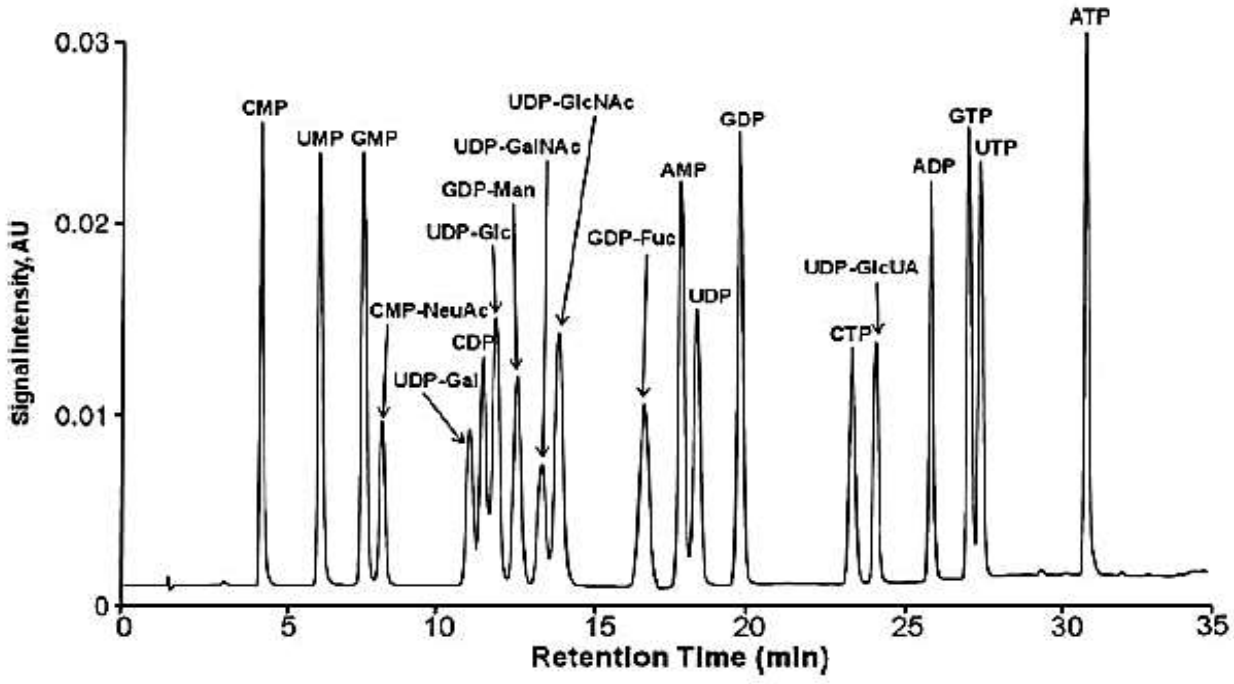


Figure 7.3: separation of a series of nucleotide sugars by optimized ion-pair reversed-phase HPLC; compounds were detected by their absorbance at 254 nm.

Starting from the paper of *Nakajima et al. (2010)* [46], first of all, we optimized the elution gradient (for details *Appendix – A.4*) in order to get the best retention times for our analyte UDP-GlcNAc which was

[57] Nakajima, Kazuki, et al. "Simultaneous determination of nucleotide sugars with ion-pair reversed-phase HPLC." *Glycobiology 20.7* (2010): 865-871.

visualized at 254 nm. The retention time was acquired determined the chromatogram's analysis of standard solution UDP-GlcNAc 1.5 mM (*Figure 7.4*).

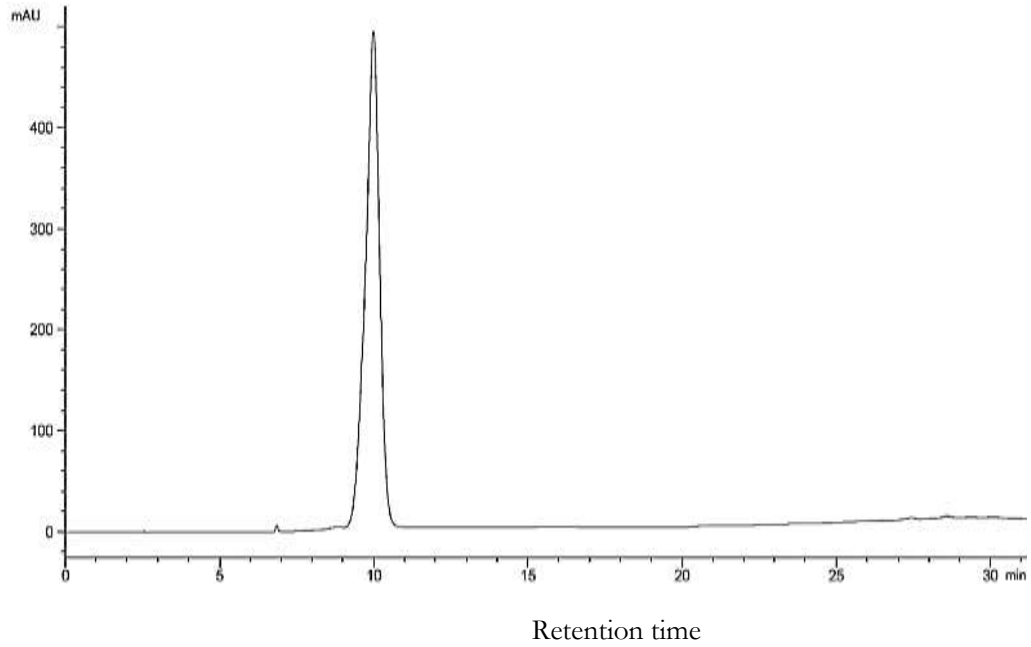


Figure 7.4: Chromatogram of UDP-GlcNAc stock solution (1.5 mM): retention time 10 minutes.

Subsequently a calibration line was generated using different concentration of analyte obtained by progressive diluition of the stock solution. (*Appendix – A.4*).

The model built on calibration line (*Appendix – A.4*) is important to predict the values of a dependent variable (quantitative), in our case the concentration of UDP-GlcNAc, starting from the values of one or more independent variables, the area of peaks. From the calibration line we get the R^2 (0,9987), a statistical measure of how close the data are to the fitted regression line, and the limit of detectability (0,006 mM) of instrument and limit of quantification (0,019 mM).

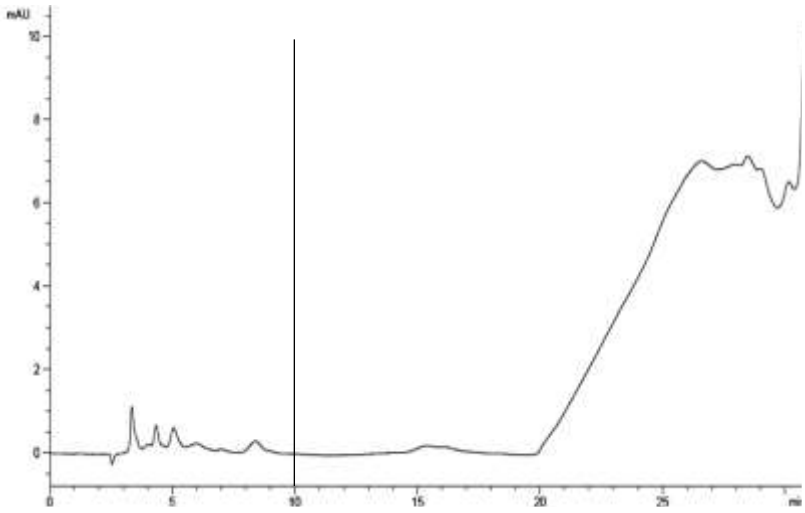
- **Analysis of cellular extract and enzymatic reaction**

The extract was used to carry out preliminary studies with the aim to find the best quantity of extract to be used in order to get a measurable quantity of UDP-GlcNAc and to verify the presence of endogenous UDP-GlcNAc in the extract.

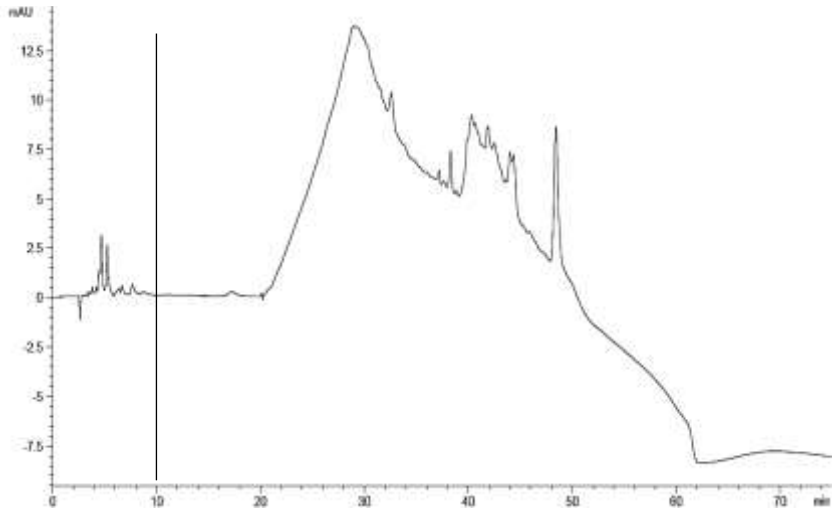
7.6. Results and Discussion

7.6.1. Cellular extract and endogenous UDP-GlcNAc

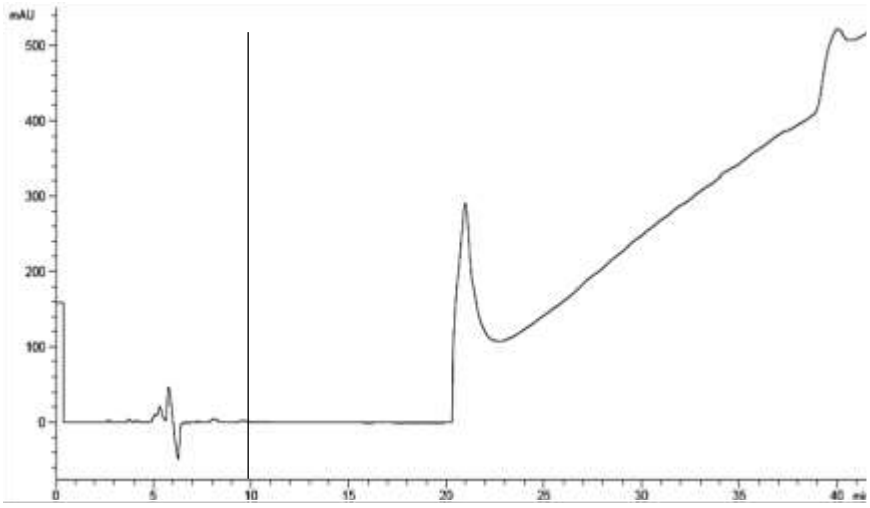
In *Figure* below are reported the chromatogram's analysis of 10µL, 30µL and 100µL of the extract:



10 µL of extract without enzimatic reaction



30 µL of extract without enzimatic reaction



100 µL of extract without enzimatic reaction

The different aliquots of extract don't present any quantifiable amount of endogenous UDP-GlcNAc, so, during the analysis, all the concentration of metabolite obtained is considered to be producted by the enzymatic reaction.

7.6.2. Enzymatic reaction with 1mM and 5 mM substrates

The first test was carried out by adding an appropriate volume of substrates to produce a quantifiable quantity of UDP-GlcNAc after the enzymatic reaction (*Appendix – A.4*).
The blank was prepared with lysate extract (30 µL) and concentration of 1mM or 5mM of substrates.

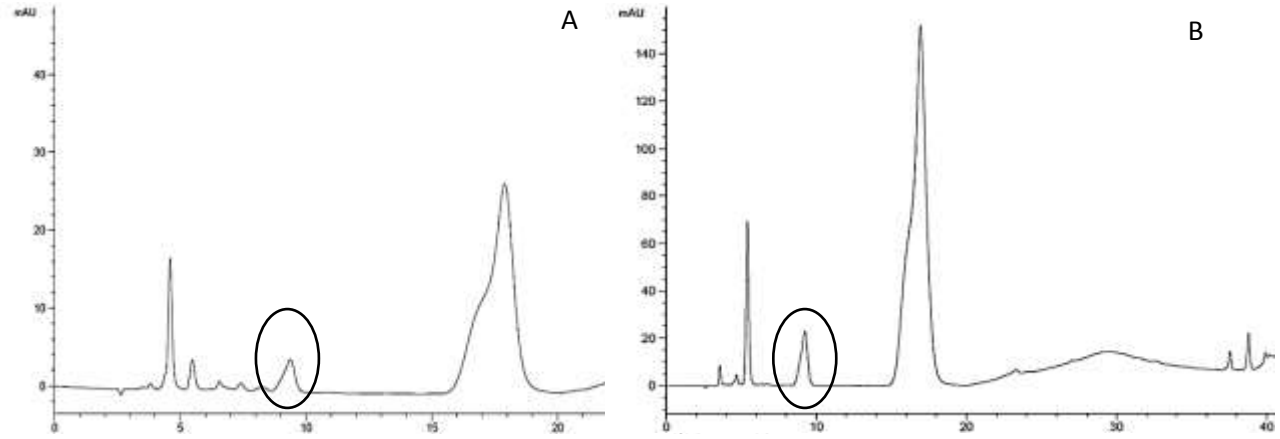


Figure 7.5: A)30uL cellular extract, 1mM substrate; B) 30uL cellular extract, 5mM substrate.

The chromatogram in *Figure 7.5* shows the two analysis performed: with 1mM substrate the effective concentration of UDP-GlcNAc produced by the enzyme reaction is below LOQ (0.0197 mM) calculated on the line of calibration (A). While using 5mM substrate the amount of UDP-GlcNAc producted is 0.100 mM, quantifiable with the method (B).

Therefore all the analysis have be carried out using concentration of substrate of 5mM.

7.6.3. Cellular extract, enzymatic reaction and quantification of UDP-GlcNAc

Next, the enzymatic reaction with 5mM of substrate inside the reaction was done on different volume of extract in order to identify the best quantity for the analysis.

The quantity of UDP-GlcNAc products by the enzymatic reaction in different volume of cellular extract is comparable and greater than the LOQ (0,019 mM – red line), so the analysis will be done with 10µL of extract at first.

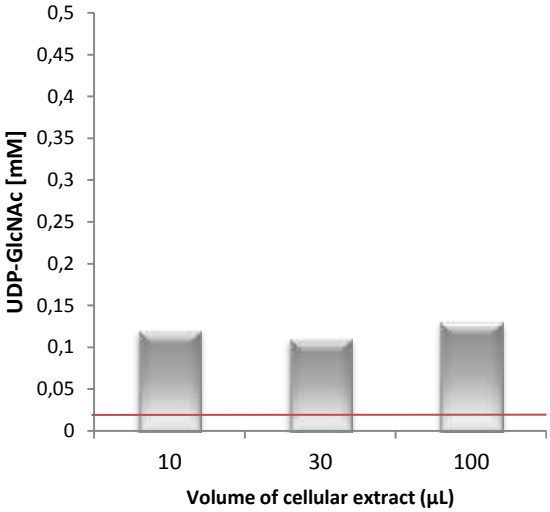


Figura 7.6: Quantity of UDP-GlcNAc products by the enzymatic reaction (5mM substrates) in 10, 30 and 100µL of extract.

7.6.4. Enzymatic reaction with inhibitors on 10µL and 30µL of extract

Next, the analysis carried out have been done thought the preparations of sample with 10µL of extract and enzymatic reaction with 5mM of substrate and 1mM of test compound (potential inhibitors). The enzymatic reaction in 10µL of extract (blank) producted 0,119 mM of UDP-GlcNAc and 30µL 0,110 mM of UDP-GlcNAc (*Figure 7.7*).

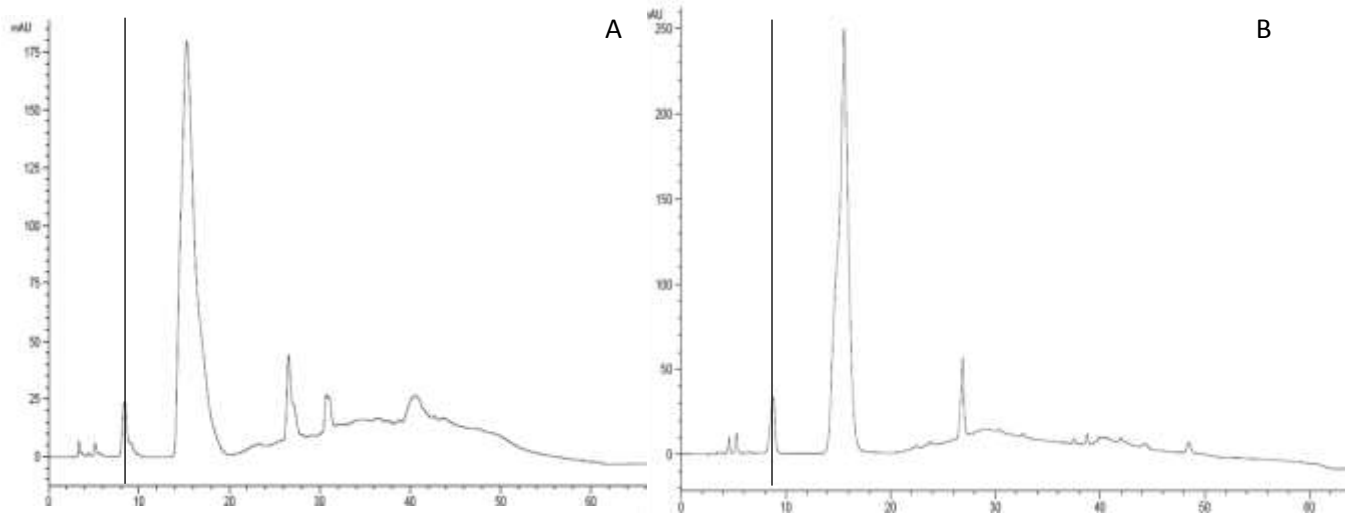
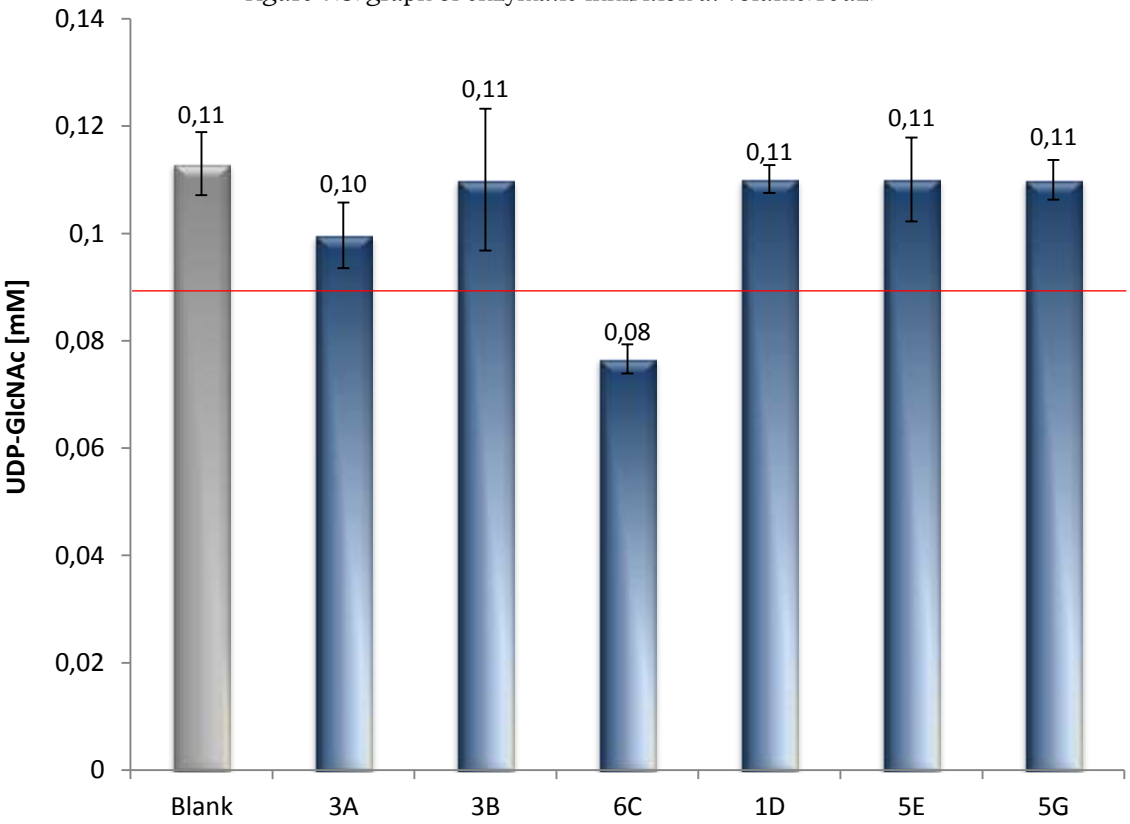


Figure 7.7 – Chromatogram of blank (A) 10 µL and (B)30 µL.

In graph below (*Figure 7.8*) are shown the mean of variation of UDP-GlcNAc quantity producted during the reaction in presence of inhibitor (1mM) and Dev. Std. in comparison with the blank in 10µL extract. Decrease of peak area of UDP-GlcNAc compared to the area of the blank, means that the inhibitor has a capacity to inhibit the activity of enzyme, so the production of UDP-GlcNAc is lower.

Figure 7.8: graph of enzymatic inhibition at volume:10uL.

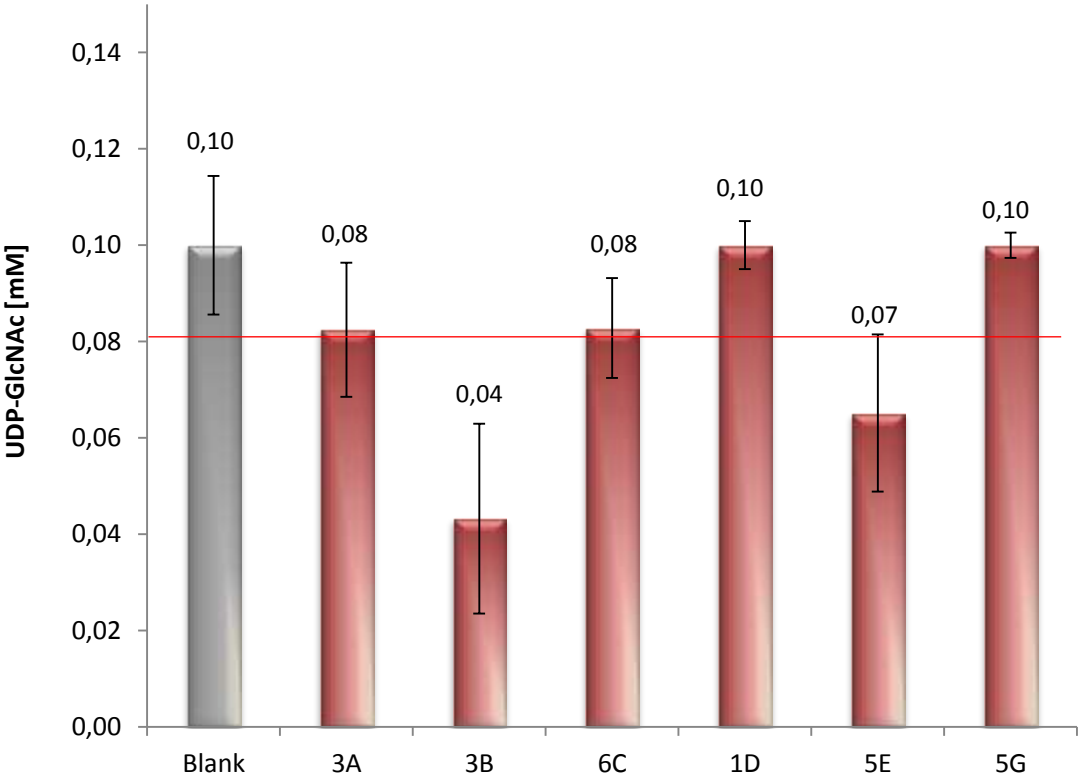


As it's shown in the histogram, from analysis of chromatogram of inhibitors tested, the variation of the quantity of UDP-GlcNAc produced in the presence of the tested compound compared to the blank reaction (without potential inhibitors) isn't quantifiable because in the range of LOQ (red line). Only in case of compound 6C there is a little decrease in the production of UDP-GlcNAc.

Secondary, analysis have been carried out using 30µL of extract in order to identify diffences between the production of UDP-GlcNAc in different volumes of extract (*Figure 7.9*).

As it's showed in histogram, from chromatogram analysis compounds (1mM) 3A, 6C, 1D and 5G the variation of quantity of UDP-GlcNAc produced isn't quantifiable because in the range of LOQ (red line). In case of compounds 3B there is a substantial decrease of UDP-GlcNAc production, while compound 5E has high standard deviation, no evaluation could be done.

Figure 7.9: graph of enzymatic inhibition at volume:30uL.



In *Figure 7.10* histogram shows the comparison of the inhibition capacities of different molecules: the potential inhibitor 1F seems to influence the activity of enzyme in 30µL of cellular extract. The molecule 6C has more effect at 10µL than 30µL of extract than molecule 3B has an effect in 30µL of extract.

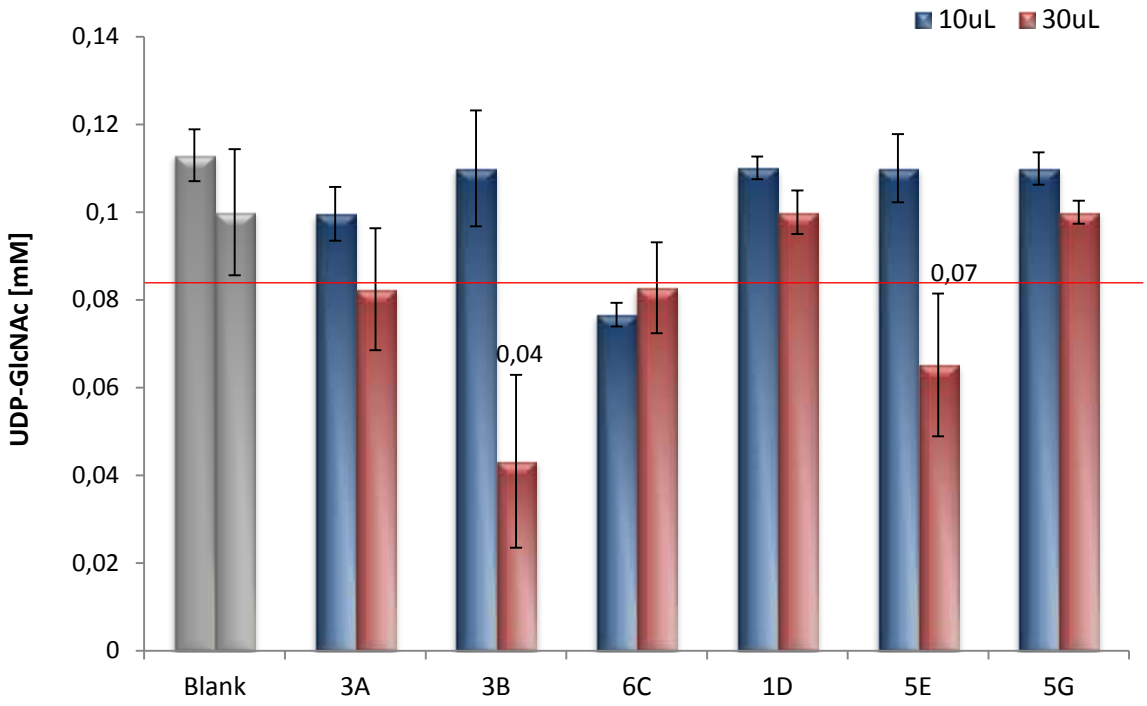


Figure 7.10: graph of enzymatic inhibition at volume both volume 10 and 30uL.

These preliminary studies on cellular extracts show that, in case of compounds 3B, the increasing concentration of proteins influences the quantity of UDP-GlcNAc product, maybe because there is a major presence of the enzyme of interest inside 30µL of extract with respect to 10µL: the concentration in proteins of the lisate corresponds to 4,99 µg/mL.

Noteworthy is the case of compound 5E, for which a significant inhibition is observed with 30µL of lisate, probably due to the increasing enzyme quantity: this event could be link to the major quantity of enzyme present.

The aim of the experimental work was to acquire data and to compare them with the theoretical ones. In *Figure 7.11*, the graph shows in green the calculated docking score of the potential inhibitors and in red the capacity of inhibition at 30µL. As you can see compound 6C has a good docking score but it hasn't capacity of inhibition is high.

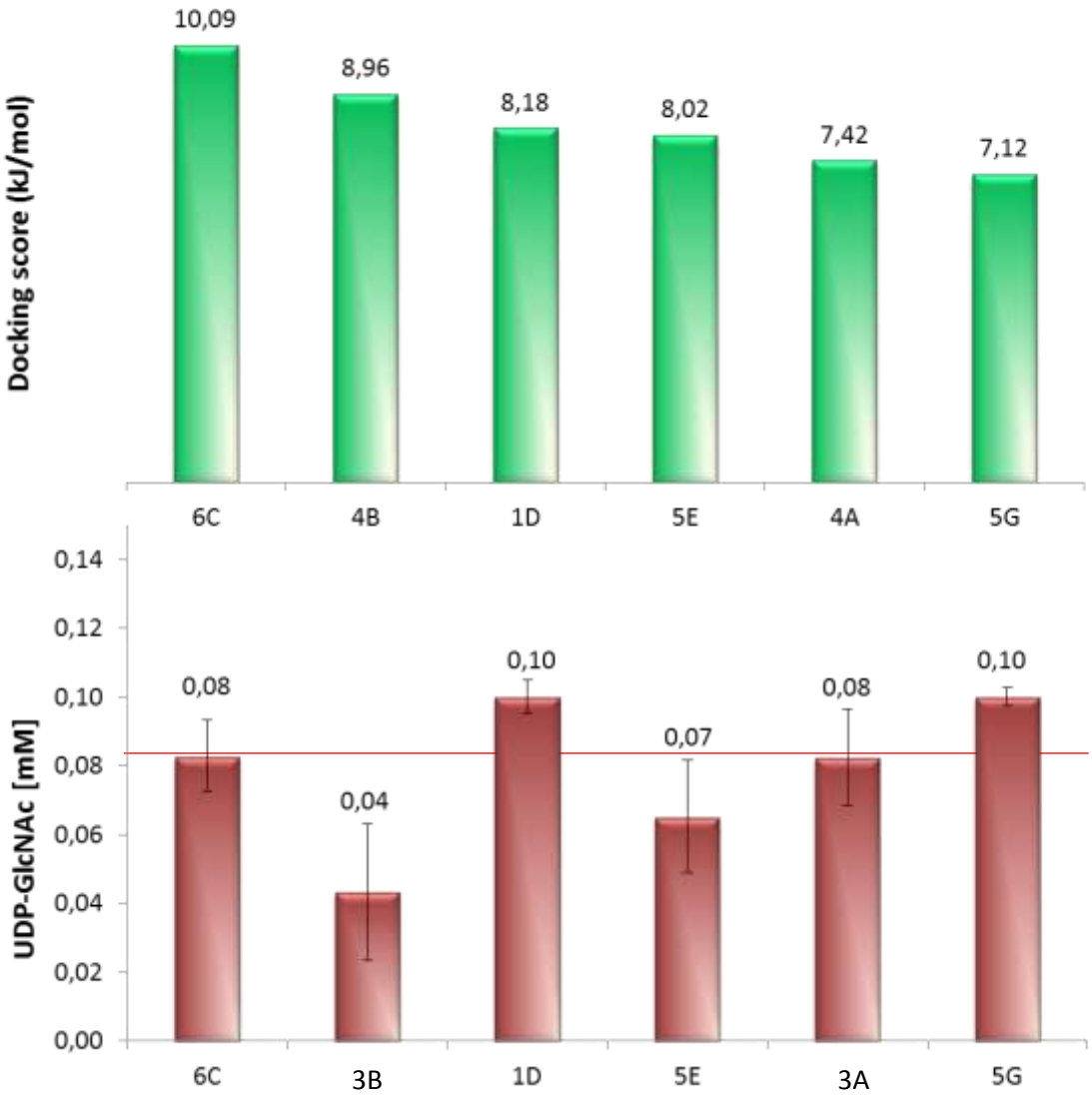
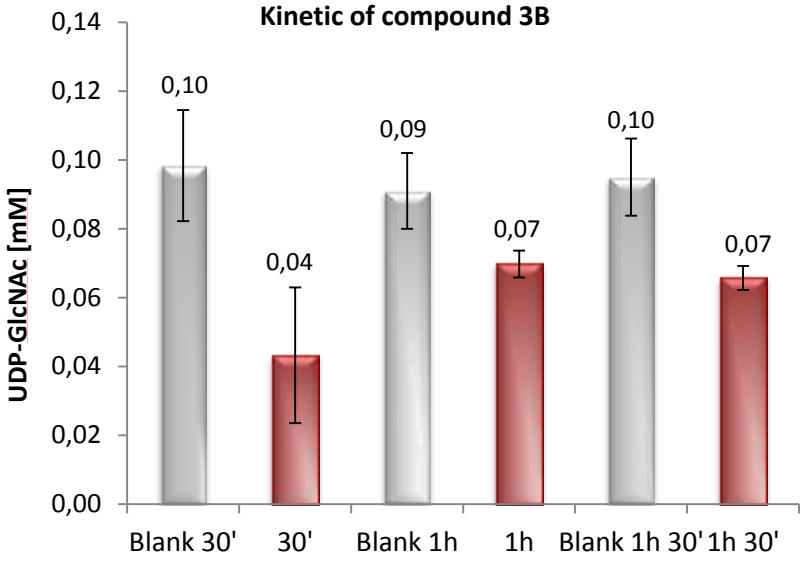


Figure 7.11: comparison between the calculated docking values (green) and experimental enzymatic inhibition at volume 30uL (in red).

The docking score of molecule 3B underlines good affinity with the enzyme and its capacity of inhibition is good too. The compound 5E has discrepancy between the value of theoretical affinity and its experimental inhibitor capacity.



Starting from this evidences and keeping in mind the minor polarity of the acetylated form of compound 3B (compound 2B) in order to test it on cells, the cinetic of enzymatic reaction has been done.

Time	Mean	Dev. St.
30'	0,04	0,02
1h	0,07	0,01
1h e 30'	0,07	0,01

From these experiments it emergs that after 30' it is not possible to observe differences between the production of UDP-GlcNAc in the presence or absence of compound 3B.

7.7. Conclusion

The validation of new analytical method of separation through the *High Performance Liquid Chromatography* to quantify UDP-GlcNAc is important in order to check all the potential inhibitors of the library, testing their capacity to inhibit the activity of enzyme AGM1 in cellular extracts.

Preliminary analysis on different volume of extract give us the indication of the best quantity of extract to use for the assay; also the concentration of substrate (1mM or 5mM) to be used in the reaction to get a quantitiafiable amount of UDP-GlcNAc produced.

Then the inhibitors (1mM) tested in the enzymatic reaction give results of production decreasing of UDP-GlcNAc after the enzymatic reaction.

The comparison with the computational data and results of experimental in vitro text, give us a description of the best interaction between enzyme and molecules: compound 3B, according to the docking score, has a better affinity than the natural substrate and, this data is confirmed by the experimental enzymatic assay.

Chapter 8
Biological test

In this part are reported some preclinical evaluation of one of glycomimetics synthetized in a Triple Negative Breast Cancer (TNBC) cell model. We will show that this novel inhibitor reducing *N*- and *O*-glycosylation protein levels leads to a decrease in cell proliferation, survival and migration.

8.1. Introduction: potentials inhibitors and *coefficient of distribution*

In order to identify a potential selective inhibitor of AGM1, we designed a series of compounds that would resemble the structure of the substrate (*Chapter 4*), but presenting at the same time, a modification in one of the key positions involved in the catalytic mechanism. In fact, during the catalytic reaction the enzyme formally transfers the phosphate monoester group from the C-6 hydroxyl to the anomeric OH: the two hydroxyl groups at position C-1 OH and C-6 OH are involved in the enzymatic reaction (*Chapter 3 – Paragraph 3.3*).

In *Chapter 5– Paragraph 5.2* are described the values of interaction between the potentials inhibitors, which structures (*Figure 8.1*) get de-protected hydroxyl group on sugar structure and, in case of molecules 3A and 3B, get the phosphate group in order to be similar to the natural substrate of enzyme AGM1, so they will be recognized from enzyme because of comparable structures but the introduction of isosteric group in the important position of C-1 and C-6 blocks its activity. With the values of docking score we can have a preliminary idea about the affinity between inhibitors and protein. The structures of the potential inhibitors possess chemical-physical characteristics that make their passage through cell membrane very harsh: they are very polar and some of them possess negative charges (phosphonates, sulphonates, sulphates).

In *Figure 8.1* are showed the structures of potential inhibitors in the acetylated form because of their lipophilic character of the cellular membrane, so they have been synthesis also in the acetylated form without phosphate group in position C-6 because of their best entry through the cell membranes during the test on cells: molecules 2A and 2B have an apolar character that can favor their passage across the membrane through a diffusion mechanism. The other molecules (7C, 2D, 4E, 2F, 4G) are very polar because possess negative charges in order to increase the uptake in cells.

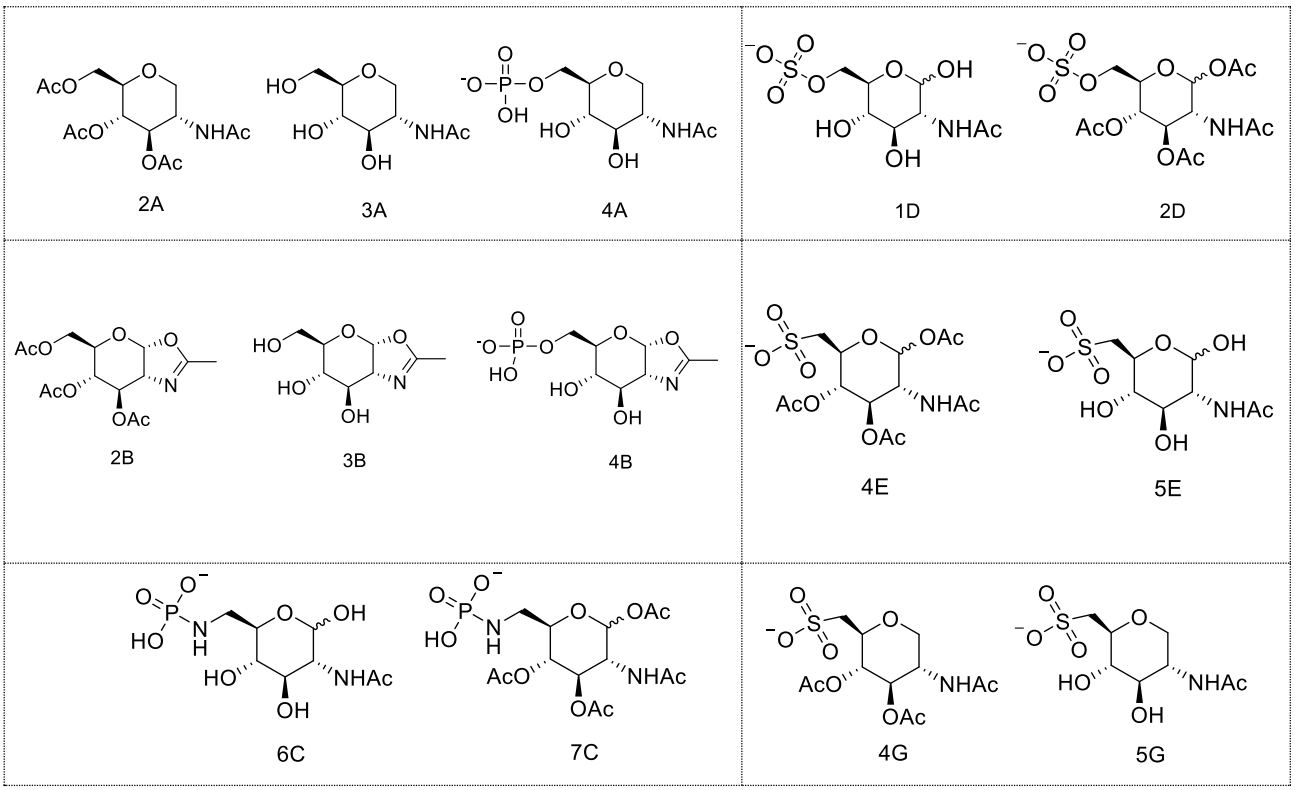


Figure 8.1: acetylated and de-protected structures of potentials inhibitors.

With the proposal to identify the potential inhibitors to be tested on cells with the best features to cross cellular membranes, I have calculated the value of **C Log*P*** through software *ChemBioDrawUltra 14.0*.

In literature there are evidence about the hydrolysis of esters group by cellular esterases *(M. Gloster et al. 2011).* [58]

Secondary, when the potential inhibitors enter in the pathway of hexosamine, will be phosphorylated directly in cell.

C Log*P* or "*coefficient of distribution*", in chemistry and in pharmaceutical sciences, represents the ratio of the concentrations of a compound within the two phases of a mixture of two equilibrium immiscible liquids. The importance of this coefficient is the indication of the hydrophobic or hydrophobic level of a chemical: through using as the mixture formed by the octanol (hydrophobic compound) and water, coefficient P_{OW}, in practical terms, is commonly indicated on a logarithmic scale in base 10, or as log. The following mathematical expression is thus obtained:

$$\log P_{ow} = \mathbf{log} \frac{[\text{solute}]\ \text{octanol}}{[\text{solute}]\text{water}}$$

[58] Gloster, Tracey M., et al. "Hijacking a biosynthetic pathway yields a glycosyltransferase inhibitor within cells." *Nature chemical biology*7.3 (2011): 174-181.

The P_{OW} log values are typically negative for hydrophilic substances, while they are positive and growing as the hydrophobic character increases. The value is a widely used parameter to assess compound solubility and cell permeability.

In graph *Figure 8.2* are reported values of the **docking score** and **C Log*P*** of the potential inhibitors. As expected, C Log*P* of polar molecules is very negative because of the presence of negative charges on the structures (6C, 1D, 5E, 1F, 5G) so they can't across cellular membrane even if they are acetylated. While molecules 2B and 2A get positive values because they don't present any charges on their structures. In particular, molecule 2B gets the highest values of C Log*P*.

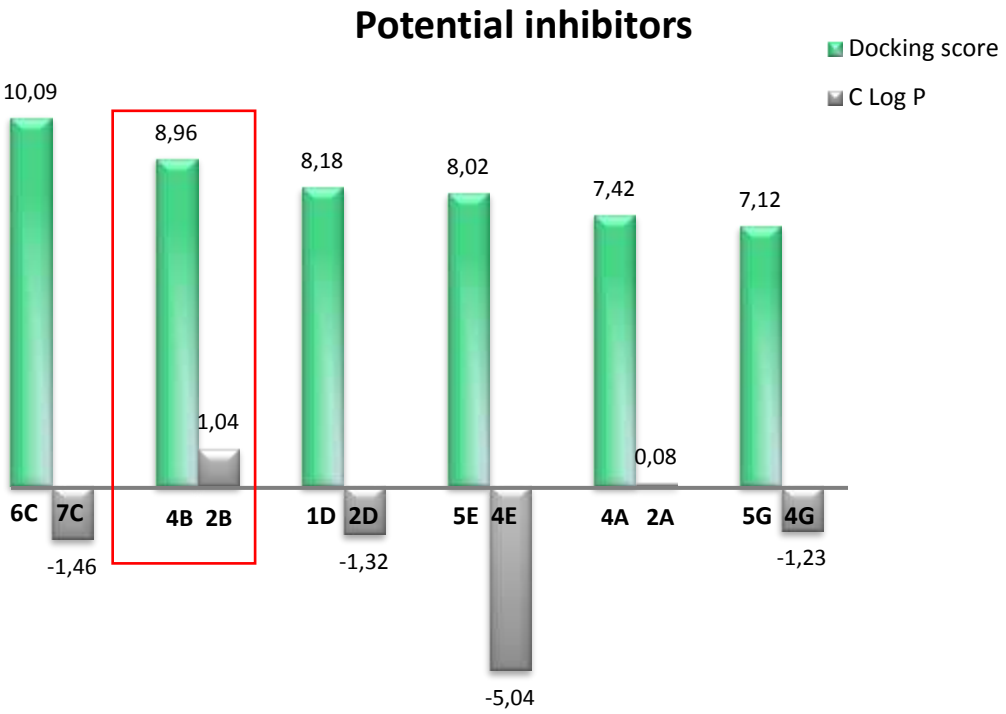
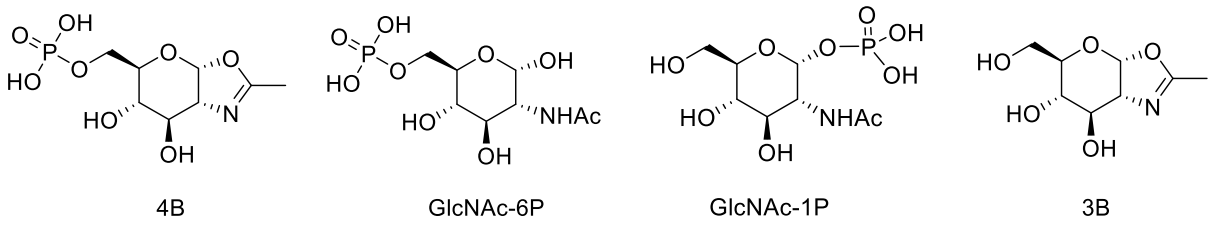


Figure 8.2: docking score of de-protected form of potentials inhibitors in green. Molecules 4B and 4A are phosphorylated in position C-6. In grey are reported the C Log*P*.

Molecule 2B gets good value of docking score, so good affinity and interactions with enzyme AGM1. Furthermore its positive value of C Log*P* is an evidence of its apolar character, so its capacity to across cellular membrane. So it will be test on cells.

8.2. Cellular test on MDA-MB-231

To assess the potential efficacy of molecule 2B as inhibitor of AGM1 enzyme, docking calculations were carried out (PDB code: 2dkd – crystal structure of AGM1 in complex with its substrate). The structures of the best poses obtained by docking calculations, the hydrogen bonds and Van der Walls interactions with the enzyme active site and the docking scores are reported in *Figure 8.3*. Since the docking score referred to the interaction of 2B in the active site of the enzyme is higher (8.96 kJ/mol) than the score referred to the natural substrate GlcNAc-6P (8.01 kJ/mol) and the product GlcNAc-1P (7.32 kJ/mol), the potential inhibitory effect of the designed molecule is strongly supported.



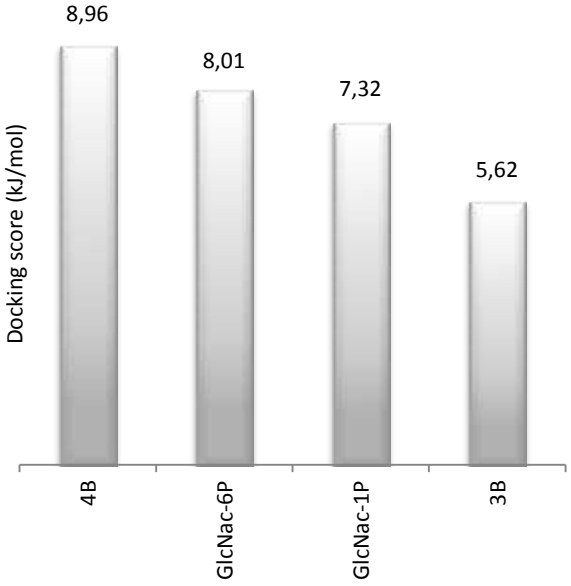


Figure 8.3: comparison between docking score of molecule 3B, natural substrate and product if enzymatic reaction.

Notably, docking calculation of the unphosphorylated 2B (2B) indicated a significant decrease of its docking score value, suggesting that the phosphorylation of the C-6 hydroxyl, allowing the formation of salt bridges interactions and hydrogen bonds with the amino acid residues present in active site, is crucial for the binding (*Figure 8.4*).

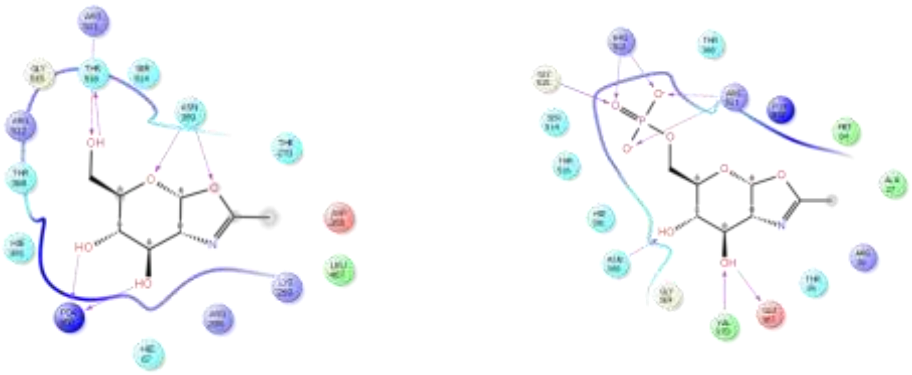


Figure 8.4: on left molecule 3B unphosphorylated, on right 4B.

In order to increase 3B cell permeability, we designed the precursor without the phosphate ester and acetylated at the hydroxyl group in position C-3,4,6 (compound 2B), assuming a subsequent intracellular hydrolysis of the acetate esters and a phosphorylation in order to become the 4B, and calculated its C Log*P*. The C Log*P* value of this compound, as expected, was strongly enhanced as compared to the de-acetylated form (*Figure 8.5*) suggesting that increased lipophilicity could greatly improve the membrane permeability.

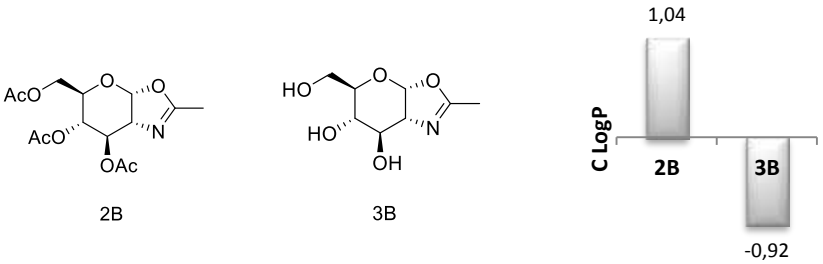


Figure 8.5: on the left the structures of molecules 2B and 3B; on right the C Log*P* of both two.

Preliminary studies were carried out from lab of prof.re Ferdinando Chiaradonna – University of Milano-Bicocca. In particular, 2B-dependent cell death is associated with activation of the Unfolded Protein Response (UPR).

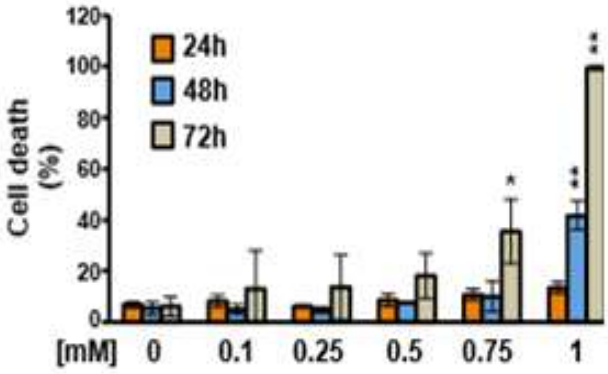


Figure 8.6: 2B treatment affects proliferation and viability of MDA-MB-231 cells by interfering with AGM1 enzyme (doses of 2B time-dependent).

As reported in literature, alteration of glycosylation level is associated to several disease, among which cancer (Ozcan et al. 2010) (Lau et al. 2007). Given that, we decided to analyzed the relevance of the HBP, the metabolic pathway regulating cellular level of glycosylation, in a triple negative breast cancer cell line MDA-MB-231, carrying an oncogenic form of K-RAS. This cell line has been chosen because their metabolism is well characterized, and in particular the role of glucose metabolism in their ability to proliferate and survive is know. When MDA-MB-231 cells were treated for 72h at different doses (from 100µM to 1mM), the

proliferation, in concentration and time-dependent manner was markedly reduced as well as the cell viability, since after 72h at high doses (750μM and 1mM) nearly 40% and 100% of the cells appeared dead (*Figure 8.6*).

Previous reports indicated that change in membrane protein GlcNAc alters cell adhesion and migration (*Gu and Taniguchi, 2008*)[59] as well as cell transformation has been correlated with a shift to larger complex-type *N*-linked oligosaccharides (*Bolscher et al., 1986*; *Dennis et al., 1987*).[60,61,62] Therefore we reasoned that 2B, interfering with HBP, could hamper cell attachment and migration as well. Morphological analysis of MDA-MB-231 cells, treated with 2B for 48h caused marked morphological alterations and cell detachment (*Figure 8.7*). In fact, in 1mM 2B almost 36% of the cells appeared rounded and detached of which only around 60% were dead, suggesting that 2B interfered with cell attachment.

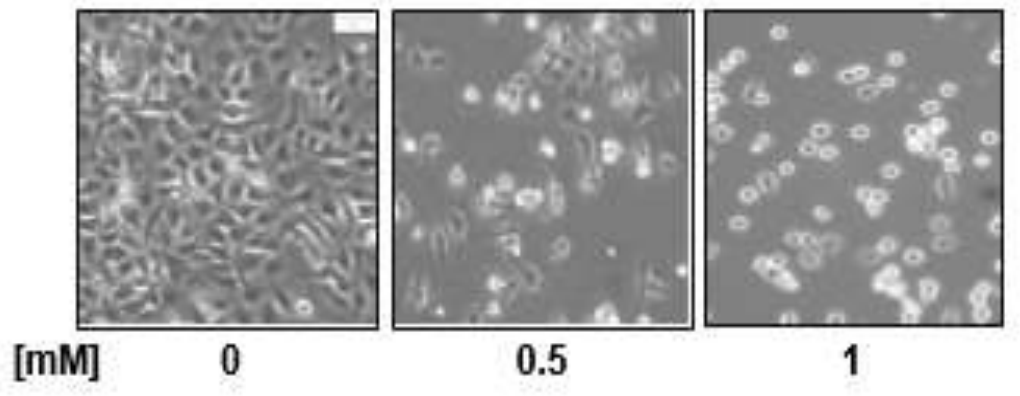


Figure 8.7: morphological alterations in cell attachment and migration. 2B specifically induces cell detachment and reduces MDA-MB-231 cell ability to adhere and migrate. MDA-MB-231 cells appearance upon 48h treatment with 2B (4X magnification, 50μm scale).

[59] Zhao, Yan-Yang, et al. "Functional roles of N-glycans in cell signaling and cell adhesion in cancer." *Cancer science* 99.7 (2008): 1304-1310.
[60] Collard, John G., et al. "Cell surface sialic acid and the invasive and metastatic potential of T-cell hybridomas." *Cancer Research* 46.7 (1986): 3521-3527.
[61] Bolscher, Jan GM, et al. "Effect of cancer-related and drug-induced alterations in surface carbohydrates on the invasive capacity of mouse and rat cells." *Cancer research* 46.8 (1986): 4080-4086.
[62] Storme, Guy A., et al. "Effect of lipid derivatives on invasion in vitro and on surface glycoproteins of three rodent cell types." *Lipids* 22.11 (1987): 847-850.

Therefore, in order to measure the effective intracellular content of 2B, MDA-MB-231 cells, treated for 24h with 1mM 2B, were analyzed by using LC/MS. In particular the acetylated (2B), the de acetylated (3B) and the de-acetylated and phosphorylated compounds were measured. All three forms were detected in cell lysate, indicating that penetration, de-acetylation of 2B and its phosphorylation occurred. Total concentration of the three compounds (0.925+/-0.106µM) indicated that 2B, despite the higher amount added to the culture medium, has an effect at nanomolar concentration, since the intracellular active compound had a concentration of 106nM +/-50nM as measured by semiquantitative tandem mass spectrometry (*Figure 8.8*).

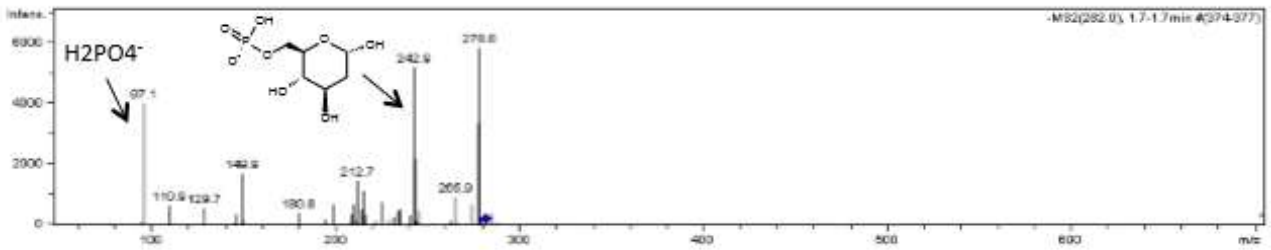


Figure 8.8: semiquantitative tandem mass spectrometry.

8.3. Conclusion

Preliminary studies have been carried out on compound 2B which reducing *N*- and *O*-glycosylation protein levels leads to a dramatic decrease in cell proliferation, survival and migration. It suppresses cancer growth in MDA-MB-231 at different doses (from 100µM to 1mM): the proliferation, in concentration and time-dependent manner was markedly reduced as well as the cell viability.

Chapter 9
Cyclodextrins and encapsulation

This section described the encapsulation of enzymatic inhibitor, compound 2D, in functionalized cyclodextrins for cancer cells targeting. In the first part, the aims of the traineeship period and the work developed at *CycloLab* are explained. The following paragraphs include a brief description of cyclodextrins properties and the last part is about the synthesis of *cyclodextrins tools* for the encapsulation of potential inhibitors.

9.1. Introduction: traineeship program

The *Hexosamine Biosynthesic Pathway* (HBP) is essential for proliferation, survival and migration ability of cancer cells. Its inhibition as well as its positive and negative modulation, may represent a fundamental basic research tool to further define its role in cancer biology and new potential therapeutic routes. In this context inhibitors of *N-acetylglucosaminephosphate mutase* (AGM1), a key enzyme of HBP, have been prepared. These compounds possess physico-chemical characteristics that make their passage through cell membrane very difficult: they are very polar and some of them possess negative charges (phosphonates, solfonates, carboxylates). The aim of the project was to synthesize and exploit ad-hoc functionalized *cyclodextrins* (CDs) as Trojan Horse to deliver the enzymatic inhibitors inside cancer cells.

In order to identify the interactions between ligand (potential inhibitors synthesized) and protein (AGM1), studies of docking have been carried out. In *Figure 9.1*, histogram of molecular docking for the potential inhibitors described the interactions of the potential inhibitors in the active site of the enzyme: docked score with better energy value than the natural substrate means that the ligands have a good affinity for the enzyme (*Chapter 4 – Paragraph 5.2*). The compounds 6C, 4B and 1D, in the de-protected form, have been docked and the derivatives get a high value of interaction compared to the natural substrate.

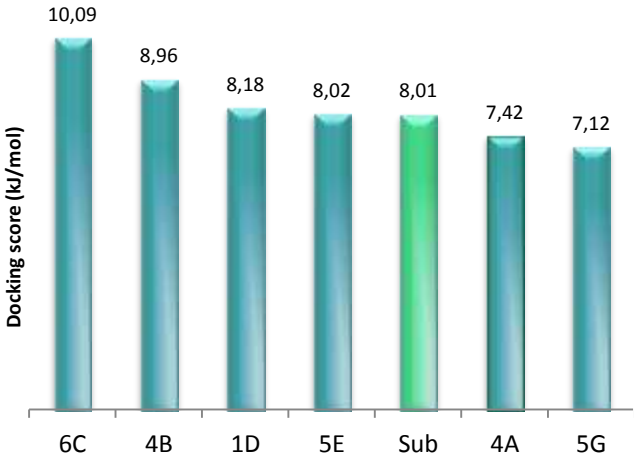
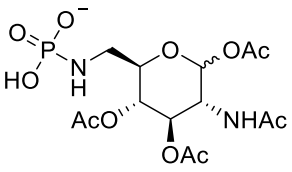


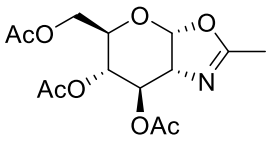
Figure 9.1: molecular docking for the potential inhibitors.

In order to prepare suitable amount of the potential inhibitors, several synthetic strategies were explored; some of the synthetized compounds are shown in Figure xx. In particular compound 7C, the acetylated form of compound 6C, has a polar character due to the phosphoramidate group in position C-6, and it could be a good candidate for encapsulation in the cyclodextrine, but it has been prepared after the period of *Erasmus Traineeship* at *CycloLab*.

Compound 2B is the acetylated version of compound 4B, without phosphate group in position C-6, and the *N*-acetyl moiety has been replaced with an oxazoline ring: its C Log*P* (1,04) value is an evidence of its apolar character, so it can across cellular membrane without any carrier.

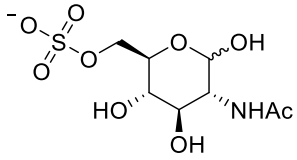


7C

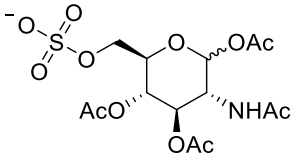


2B

Compound 2D is the acetylated form of compound 1D and has a sulphate group on C-6. The presence of acetyl groups improves the interactions with the cyclodextrins: it can be considered a "pro-drug", compound that, after administration, is metabolized (i.e. converted into the body) into a pharmacologically active drug.[63]



1D



2D

Compound 2D presents a sulfonic groups which is well tolerated biologically. A trial of their high biological tolerance, for example, is the highly safe use of beta cyclodextrin sulfobutil (Dexolve or Captisol)[64] as excipients for endovenous preparations.

Globally, compound 2D has acetylated groups that make the glucid skeleton more suitable for complexation with cyclodextrin, the sulfonic group is biocompatible and can further stabilize the complex with ion-dipole or ion-induced dipole interactions.[65,66]

[63] Miles Hacker, William S. Messer II, Kenneth A. Bachmann *Farmacologia: principi e pratica.* Academic Press, 19 giugno 2009. pp. 216-217.
[64] SBE-β-CD:Sulfobutylether-β-Cyclodextrin – CAS Number 182410-00-0
[65] Cram, Donald J., and Jane M. Cram. "Host-Guest Chemistry." *Science*, vol. 183, no. 4127, 1974, pp. 803–809
[66] Rekharsky, Mikhail V., and Yoshihisa Inoue. "Complexation thermodynamics of cyclodextrins." *Chemical reviews* 98.5 (1998): 1875-1918.

9.2. Cyclodextrins properties

Current treatments of cancer are based on drugs whose main drawback is their non-selective uptake by both normal and tumor cells. On the other hand, our potential inhibitors have shown a selective uptake thanks to the different response at UPR (*Chapter 2 – Paragraph 2.4*). The high demand for the new drug delivery strategies to combat major diseases in discriminatory ways has stimulated our research to ward the development of tumor cells selective molecules: ad-hoc modified cyclodextrins can be used as drug carriers targeting PDAC cells.

Also known as cycloamylose, clomaltic or Schardinger's dextrin, they originate from an intramolecular transglycosylation reaction by cyclodextrin-glucanotransferase (CGTase) enzyme degradation: CDs are cyclic, non-reducing oligosaccharides built up from six (α), seven (β) or eight (γ) glucopyranose units. The three major CDs are crystalline, homogeneous, non-hygroscopic, water soluble substances which are torus-like macrocyles (*Figure 9.2*) [67]

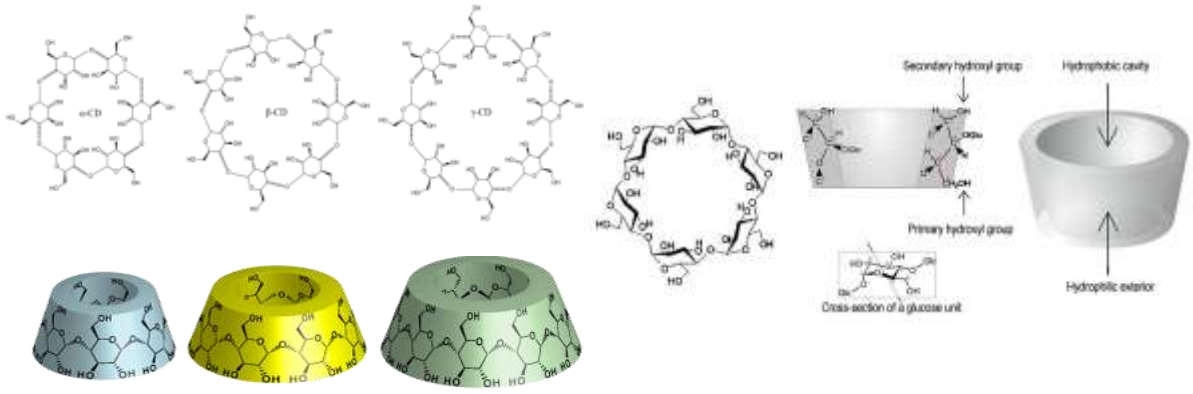


Figure 9.2: on left structures of α, β and γ cyclodextrines. On right torus-like macrocyles.

Because of their hydrophobic cavity, they can form complexes with a large variety of compounds, thus they can enhance the solubility and bioavailability of such compounds. The complexing ability of these sugars makes them valuable and versatile tools for enhancing the solubility and bioavailability of a multitude of guest-molecules. It is not surprising that CDs are well known and widely used in the pharmaceutical field.

Cyclodextrins are typically constituted by 6-8 glucopyranoside units, can be topologically represented as toroids with the larger and smaller openings of the toroid exposing to the solvent secondary and primary hydroxyl groups, respectively. Because of this arrangement, the interior of the toroid is not hydrophobic, but

[67] Szejtli, Jozsef. "Introduction and general overview of cyclodextrin chemistry." *Chemical reviews* 98.5 (1998): 1743-1754.

considerably less hydrophilic than the aqueous environment, thus enabling to host of hydrophobic molecules. The exterior part is sufficiently hydrophilic to impart cyclodextrins water solubility.

9.3. Cyclodextrins complex with drug

Suitable methods for selecting the best cyclodextrins-guest molecule match were ad-hoc investigated: NMR spectroscopy experiments, molecular dynamics simulations, and theoretical chemistry calculations provide insight into the structural and energetic determinants of the distinct binding of cyclodextrins and the migration order reversal of their respective inclusion complexes in capillary electrophoresis. Among the tested ones, *Capillary Electrophoresis* (CE) was the most effective. It is fast, cheap and one of the best method for screening charged analytes. CE is relatively recent instrumental separation technique, which has become increasingly important over last decade. The intrinsic efficiency of CE as a separation technique, together with the developments in detection, are two key aspects for its progressive implementation as a routine separation technique. The analytes separate as they migrate according to their electrophoretic mobility, and are detected near the outlet end of the capillary. The output of the detector is sent to a data output and handling device and the data is then displayed as an electropherogram, which reports detector response as a function of time. Separated chemical compounds appear as peaks with different migration times in an electropherogram.[68]

Several cyclodextrin derivatives were screen by capillary electrophoresis in order to choose the most favorable interaction between the compound 2D and the cyclodextrins: in case of ionic guest a wide variety of CDs can be used for screening; while for non-ionic compounds, ionic CDs are injected as sample solution in the run buffer.[69,70]

[68] Marina, Maria Luisa, Angel Ríos, and Miguel Valcárcel, eds. *Analysis and detection by capillary electrophoresis*. Vol. 45. Newnes, 2005.
[69] Plätzer, Manuela, Maria A. Schwarz, and Reinhard HH Neubert. "Determination of formation constants of cyclodextrin inclusion complexes using affinity capillary electrophoresis." *Journal of Microcolumn Separations* 11.3 (1999): 215-222.
[70] Loftsson, T., and Már Másson. "The effects of water-soluble polymers on cyclodextrins and cyclodextrin solubilization of drugs." *Journal of Drug Delivery Science and Technology* 14.1 (2004): 35-43.

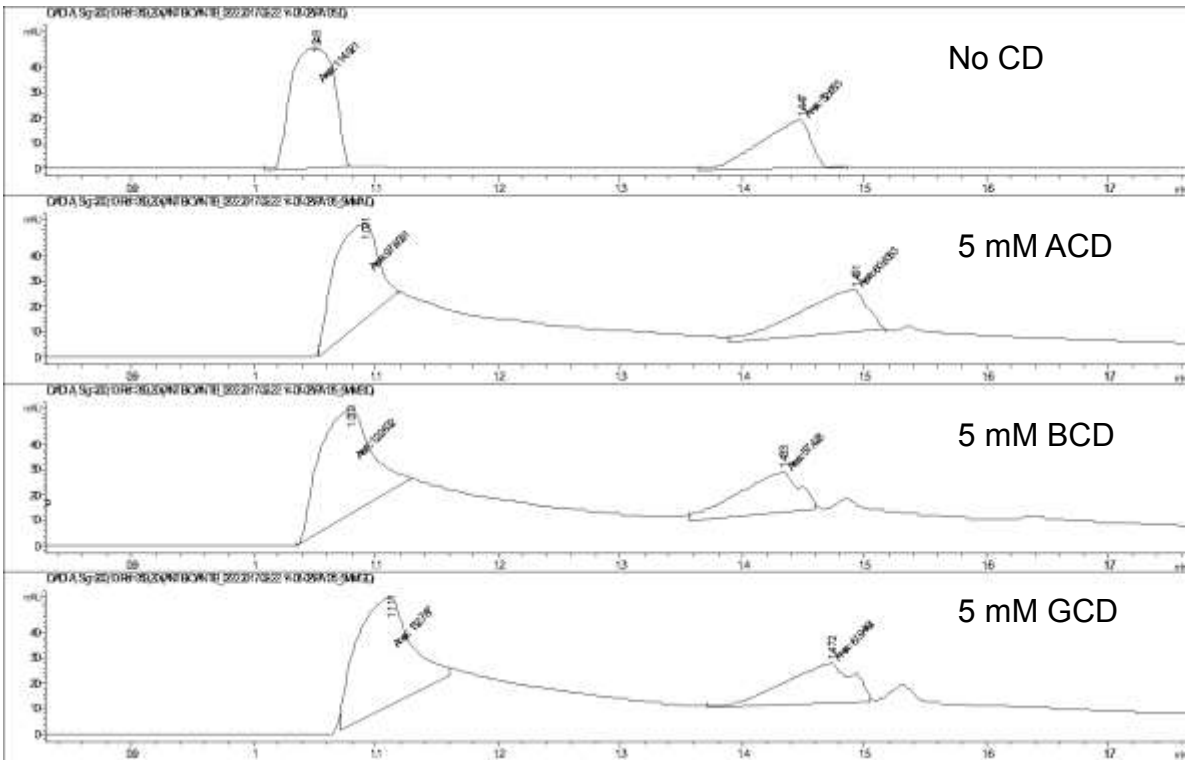


Figure 9.3: complex CD-compound.

In *Figure 9.3* is shown the effect of different CDs on the migration time of the guest-analyte (2D) added at the same concentration to the buffer. This indirect method detection allows the quantification of apparent binding constant (K values) and permits to quantify strength of complexation. The higher the K value (apparent stability constant, 1/mol unit), the stronger the interaction. These values of the constants are calculated from a series of experiments each using different CD concentrations with the analyte: the electro-osmotic flow (EOF) is a reference marker point (peak at 14/15 min), and the other peak is for the analyte. When no CD is added, the benchmark is obtained. When CDs is added, if the analyte interacts with the host molecule, then changes in the mobility are detected. The closer the EOF and analyte peaks get, the stronger the complexation is. The best CDs to complex compound 2D is γCD.

9.4. Cyclodextrins and cyclodextrins-complexes characterization by NMR

The cyclodextrin derivatives were characterized by mono and two dimensional NMR spectroscopy while the complex formation was confirmed by ^{1}H-NMR analysis.

In order to analyze and evaluate the ^{1}H spectrum of the host-guest complex and to confirm the interactions between the two entities by NMR, it is necessary to fully (or partially) assign the NMR frequencies of the host and guest species alone. In *Figure 9.4* the ^{1}H-NMR spectrum of native γ-CD is shown The spectrum is relatively simple, with sharp peaks and almost well-defined multiplicities, it is a typical spectrum of a symmetric, single isomer CD. Three set of frequencies can be identified:

i) the anomeric region, at around 5.2 ppm, with a sharp doublet, representing the eight anomeric protons (H1). This is the starting point for all the NMR-based CD assignment.

ii) the region between 4.1 and 3.8 ppm that comprises the frequencies of the H3, H5 and H6 protons.

iii) the set of frequencies between 3.8 and 3.6 ppm, these resonances can be assigned to the protons H2 and H4.

A full assignment of the γCD cannot be achieved without the aid of 2D NMR techniques. At this regard, Distortionless Enhancement by Polarization Transfer-edited Heteronuclear Single Quantum Coherence (DEPT-ed HSQC) spectra are essential tools for a clear an unambiguous evaluation. This technique allows a fast discrimination between methylene (blue signals in the spectrum in *Figure 9.4*) and methine units (red signals in the *Figure 9.4*)

In *Figure 9.4*, DEPT-ed HSQC spectrum of γCD is shown: the frequency at ~60 ppm can be assigned to the unsubstituted carbons 6 (C6, blue circle). As the γCD is unmodified, only one type of methylene unit is expected for this molecule. The frequencies at around 80 ppm are typical for the H4 protons, they usually possess the highest carbon chemical shifts. The H2 and H3 frequencies can be promptly assigned by COSY analysis [71], while the H5 frequencies usually resonate in the same region (as they are not involved in chemical modification of the CD ring) and they are identified as last, remaining frequencies.

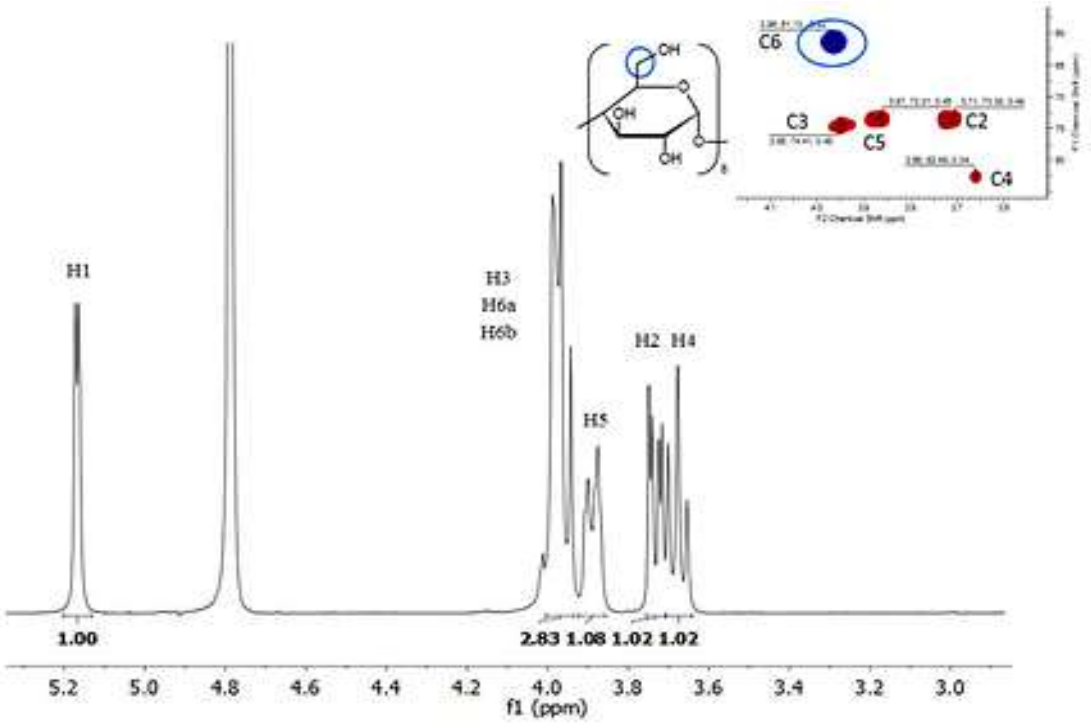


Figure 9.4: ^{1}H Deuterium oxide and HSQC-DEPT spectra of gamma CD. ^{1}H-NMR spectrum of native γCD (D_2O, x MHz, TEMPERATURE in kelvin).

[71] Eliadou, Kyriaki, et al. "Synthesis of 6-mono-6-deoxy-β-cyclodextrins substituted with isomeric aminobenzoic acids. Structural characterization, conformational preferences, and self-inclusion as studied by NMR spectroscopy in aqueous solution and by X-ray crystallography in the solid state." *The Journal of organic chemistry* 68.22 (2003): 8550-8557.

In *Figure 9.5*, ^{1}H spectrum of compound xx is shown the assignement of protons is described below: the frequencies of the anomeric proton resonates at ~6.11 ppm while the signals at ~5.34 pmm and at 4.36-4.30 ppm can be asigned to the H3 and H5, respectively.

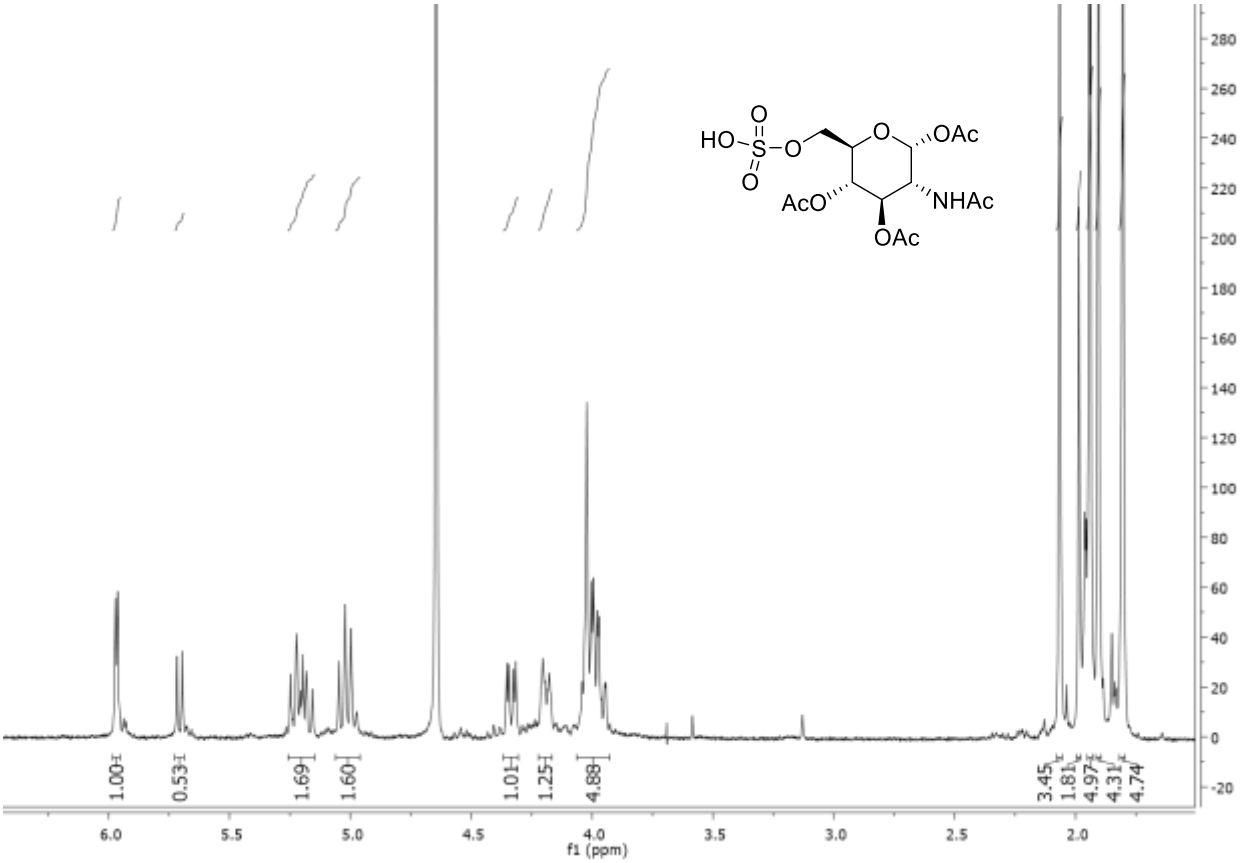


Figure 9.5: ^{1}H NMR (400 MHz, Deuterium Oxide) δ 5.97 (d, *J* = 3.6 Hz, 1H, **H-1α**), 5.71 (d, *J* = 8.8 Hz, 0.6H, **H-1β**), 5.27 – 5.14 (m, 1.6H, **H-3αβ**), 5.06 – 4.97 (m, 1.6H, **H-4αβ**), 4.33 (dd, *J* = 11.0, 3.7 Hz, 1H, **H-2α**), 4.22 – 4.16 (m, 1H, **H-5α**), 4.06 – 3.93 (m, 4.4H, **H-2β**, **5β**, **6abαβ**), 2.07 (s, **Ac**), 1.99 (s, **Ac**), 1.94 (s, **Ac**), 1.91 (s, **Ac**), 1.81 (s, **Ac**).

Finally, in *Figure 9.6* the ^{1}H spectrum of the complex γCD-1D is shown: the signals of the molecules constituting the complex can be easily recognized. The intensities of the CD counterpart are more pronounced than those of the guest molecule, still the chemical shifts of the host and guest are slightly modified confirming some interactions:

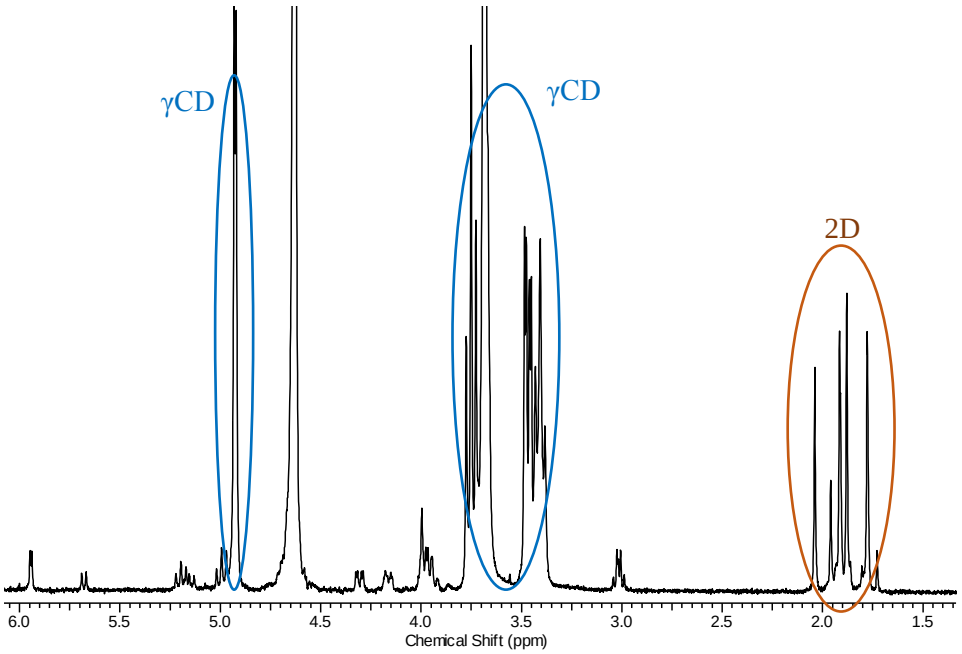
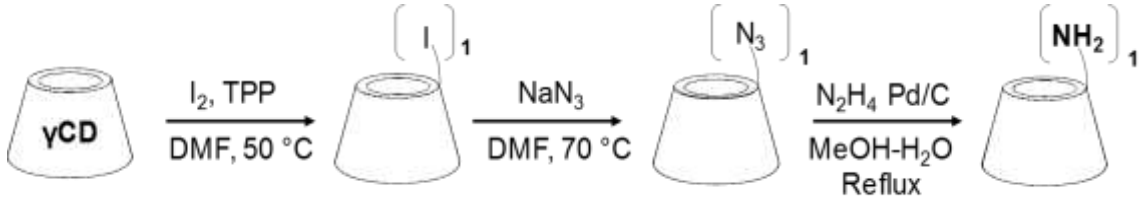


Figure 9.6: ^{1}H Deuterium oxide of complex.

In *Figure 9.6*, at ppm 5 and from 3.2 to 3.9 ppm there are signals of the cyclodextrins. The signals of compound 2D are shifted compared the isolated molecule spectrum. In conclusions, the interaction detected by CE between γCD and the 2D compound was confirmed by NMR spectroscopy.

9.5. Cyclodextrins synthesis

In order to get γCD as Trojan Horse to deliver the enzymatic inhibitors inside cancer cells we functionalized *gamma cyclodextrins* with biotin, selective targeting cancer cell molecule.[72,73] In order to introduce a single biotin moiety on the γCD scaffold, a single amino group was installed on the primary side of the macrocycle. The utilized synthetic strategy is of general application, it involves a three steps reaction and it is shown in *Scheme 9.1*.



Scheme 9.1: general scheme for the synthesis of 6-monoamino-γCD.

First, one single glucose unit of the cyclodextrin was selectively modified on the primary side with an effective leaving group such as iodine. The primary-side monohalogenation of CD was achieved by strictly controlling the temperature under the classical reaction conditions described first by *Gadelle et al. (1991)* [74].

CD can be directly and effectively halogenated by reacting the macrocycle with triphenylphosphine and bromine/iodine in DMF. [75] The monosubstitution can be achieved by controlling the temperature of the reaction mixture as described by *Malanga et al.*[76]

The halogenation is selective for the primary side, if the crude product is a mixture of mono- di- and trisubstituted isomers, then he purification is based on chromatography to obtain the pure, 6-monosubstituted derivative. 6-monoazido CDs can be prepared by displacement of the halogen with azide in DMF. The conversion is effective, quantitative and the purification consists of removal of the solvent under reduced pressure and crystallization. The reduction of the azido group is performed by hydrogenolytic method with palladium on carbon as catalyst, hydrazine as hydrogen donor in refluxing methanol aqueous mixture. The method is fast (the reaction is completed in 30 minutes), the conversion is quantitative and the yield good (85%).

---

[72] Yang, Wenjun, et al. "Targeting cancer cells with biotin–dendrimer conjugates." *European journal of medicinal chemistry* 44.2 (2009): 862-868.

[73] Yellepeddi, Venkata K., Ajay Kumar, and Srinath Palakurthi. "Biotinylated poly (amido) amine (PAMAM) dendrimers as carriers for drug delivery to ovarian cancer cells in vitro." *Anticancer research*29.8 (2009): 2933-2943.

[74] Gadelle, Andrée, and Jacques Defaye. "Selective Halogenation at Primary Positions of Cyclomaltooligosaccharides and a Synthesis of Per-3, 6-anhydro Cyclomaltooligosaccharides." *Angewandte Chemie International Edition* 30.1 (1991): 78-80.

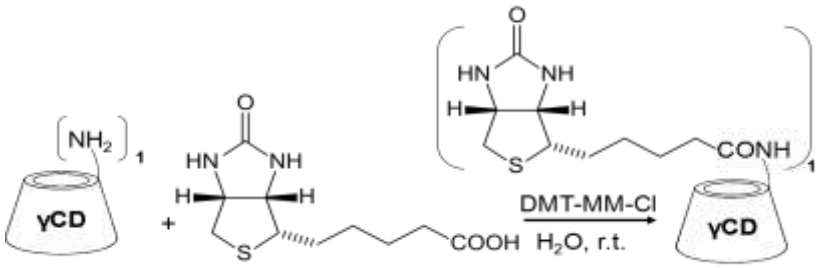
[75] Chmurski, Kazimierz, and Jacques Defaye. "An improved synthesis of 6-deoxyhalo cyclodextrins via halomethylenemorpholinium halides Vilsmeier-Haack type reagents." *Tetrahedron letters* 38.42 (1997): 7365-7368.

[76] Malanga, Milo, et al. "Synthetic strategies for the fluorescent labeling of epichlorohydrin-branched cyclodextrin polymers." *Beilstein journal of organic chemistry* 10 (2014): 3007.

The purification consists of filtration (to remove the catalyst), azeotropic distillation (to remove the solvents and most of the hydrazine), ion exchange resin treatment and precipitation. The procedure can be effectively scale-up and generates the 6-monosubstituted CD as free base.

- **Biotin CD synthesis**

In *Scheme 9.2* is shown the coupling reaction between biotin and 6-monoamino-γCD.



Scheme 9.2: synthetic scheme for 6-monobiotinyl-γCD.

The condensation reaction occurs in water at room temperature with the aid of the coupling agent4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride (DMT-MM-Cl). DMT-MM-Cl is an effective reagent widely utilized in peptide chemistry and its byproducts can be easily removed by selective precipitation as they are water soluble, By using a slight molar excess of the condensing agent, the reaction can be pushed to completeness. The work-up/purification is based on partial removal of the solvent under reduced pressure, ion exchange resin treatment, clarification an activated charcoal and selective precipitation of then 6-monobiotinylated γCD. The industrial scale-up of the procedure is feasible and the yield is satisfactory (75%).

The compound was characterized by 1D and 2D NMR spectroscopy. In *Figure 9.7*, the ^{1}H and the DEPT-ed HSQC spectra of the 6-monobiotinyl-γCD are shown:

i) around 5.1 ppm, the anomeric region with a sharp doublet of the H1;
ii) the region between 4.1 and 3.8 ppm that comprises the frequencies of the H3, H5 and H6 protons;
iii) the set of frequencies between 3.8 and 3.6 ppm, these resonances can be assigned to the protons H2 and H4;
iv) from 1.5 to 3 ppm the resonance of aliphatic signals of biotin.

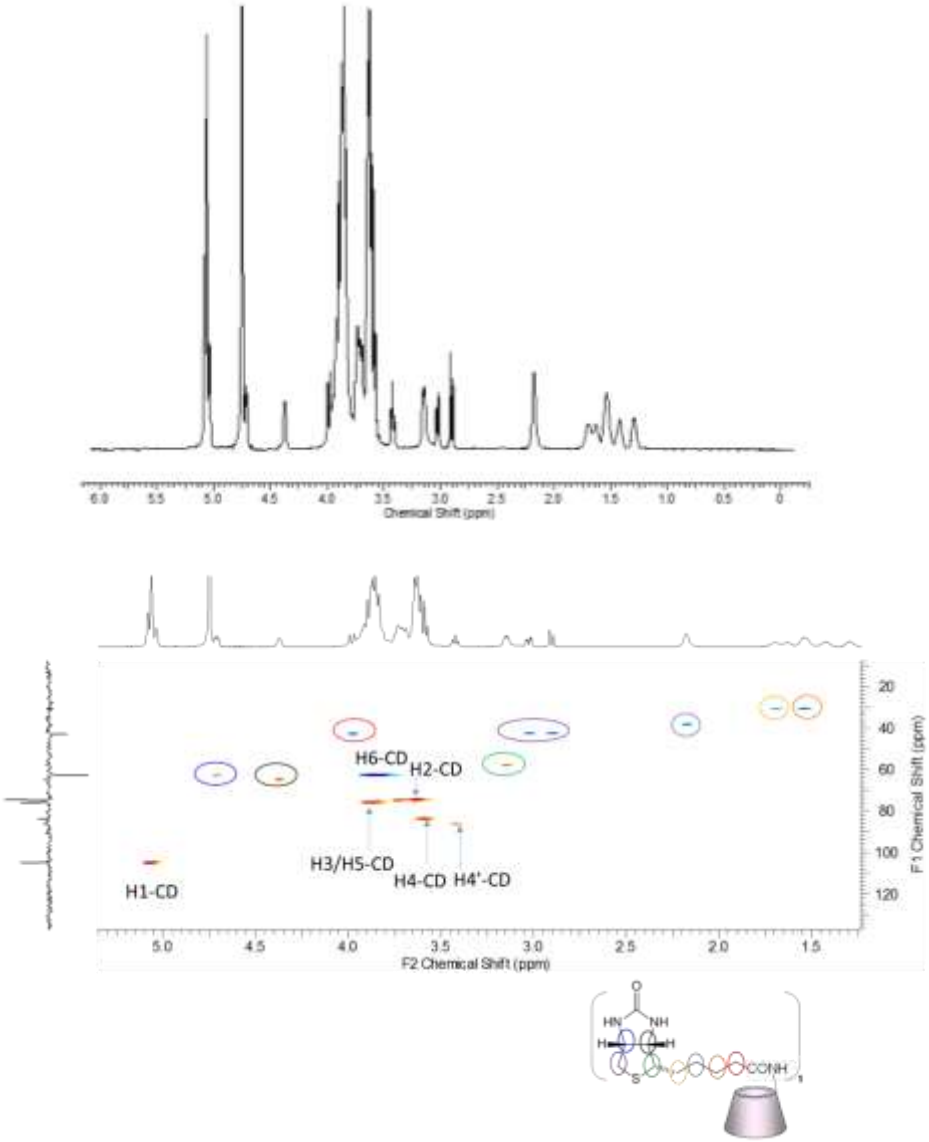


Figure 9.7, the ^{1}H and the DEPT-ed HSQC spectra of the 6-monobiotinyl-γCD.

In conclusion, the structure of the 6-monosubstituted CD has been proven.

The complex between the 6-monobiotinyl-γCD and 2D was prepared by kneading according to well described procedure [77]

[77] Loftsson, Thorsteinn, and Marcus E. Brewster. "Cyclodextrins as functional excipients: methods to enhance complexation efficiency." *Journal of pharmaceutical sciences* 101.9 (2012): 3019-3032.

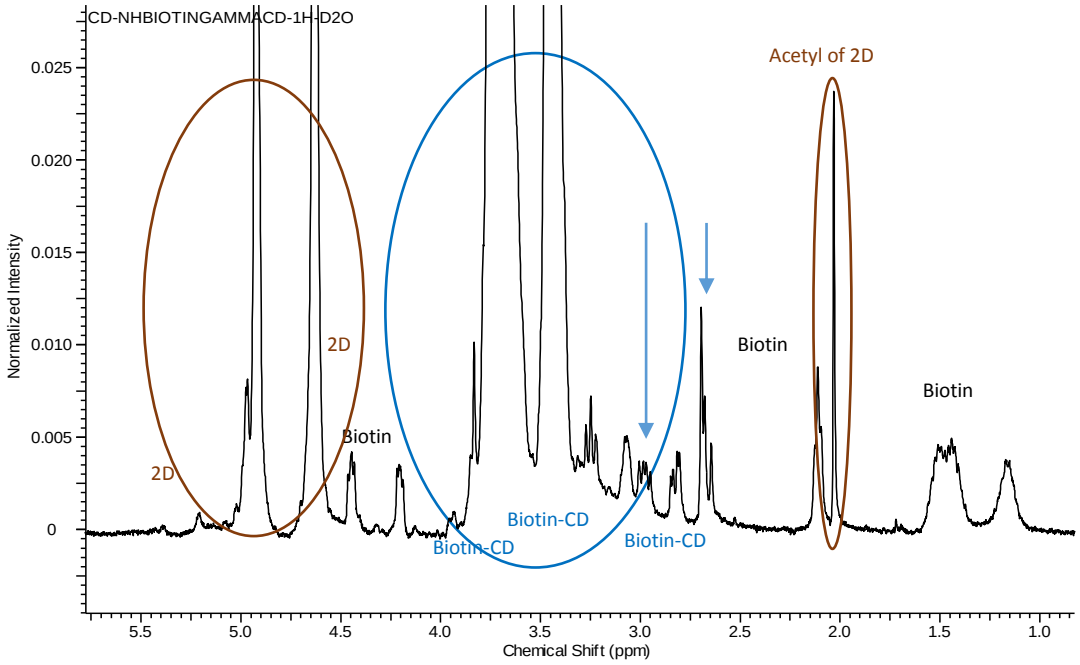
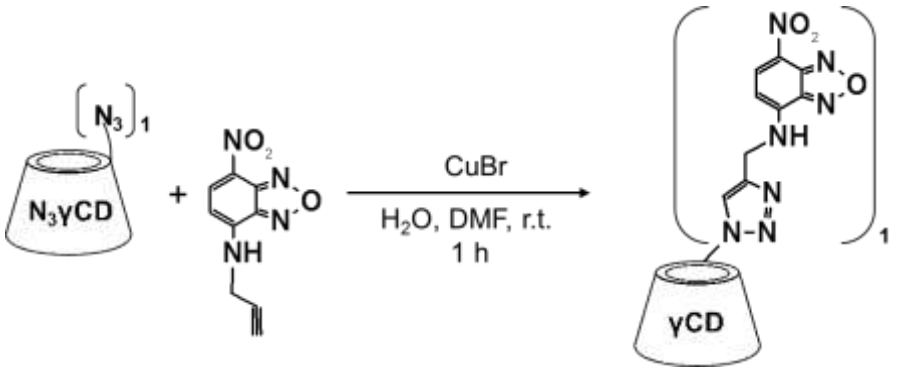


Figure 9.8: ^{1}H Deuterium oxide of complex γCD-NHbiotin and 2D.

In *Figure 9.8* is reported the ^{1}H spectrum of compound 2D-γCD-biotin complex: some signals are typical of the free biotin: from 2.5 to 4.0 ppm there are signals of γCD-biotin and signals of compound 2D are shifted compared the isolated molecule spectrum; in particular the multiplicity of signals indicates by the arrow is different than in the isolated biotin- γCD spectrum and it is a proof of interaction. In conclusions, the interaction detected by CE between and the 2D compound was confirmed by NMR spectroscopy.

➢ **NBF-Triazole γCD synthesis**

In *Scheme 9.3* is shown the coupling reaction between nitrobenzofuran-*N*-propargyl and 6-monoazido-γCD. The reaction occurs at room temperature in a 50% DMF-aqueous solution with the aid of copper bromide as catalyst. *In Figure 9.9*, the 1D and 2D spectra of the 6-mono-nitrobenzofurazan-triazolyl-γCDderivative are shown.



Scheme 9.3: coupling reaction between nitrobenzofuran-*N*-propargyl and 6-monoazido-γCD.

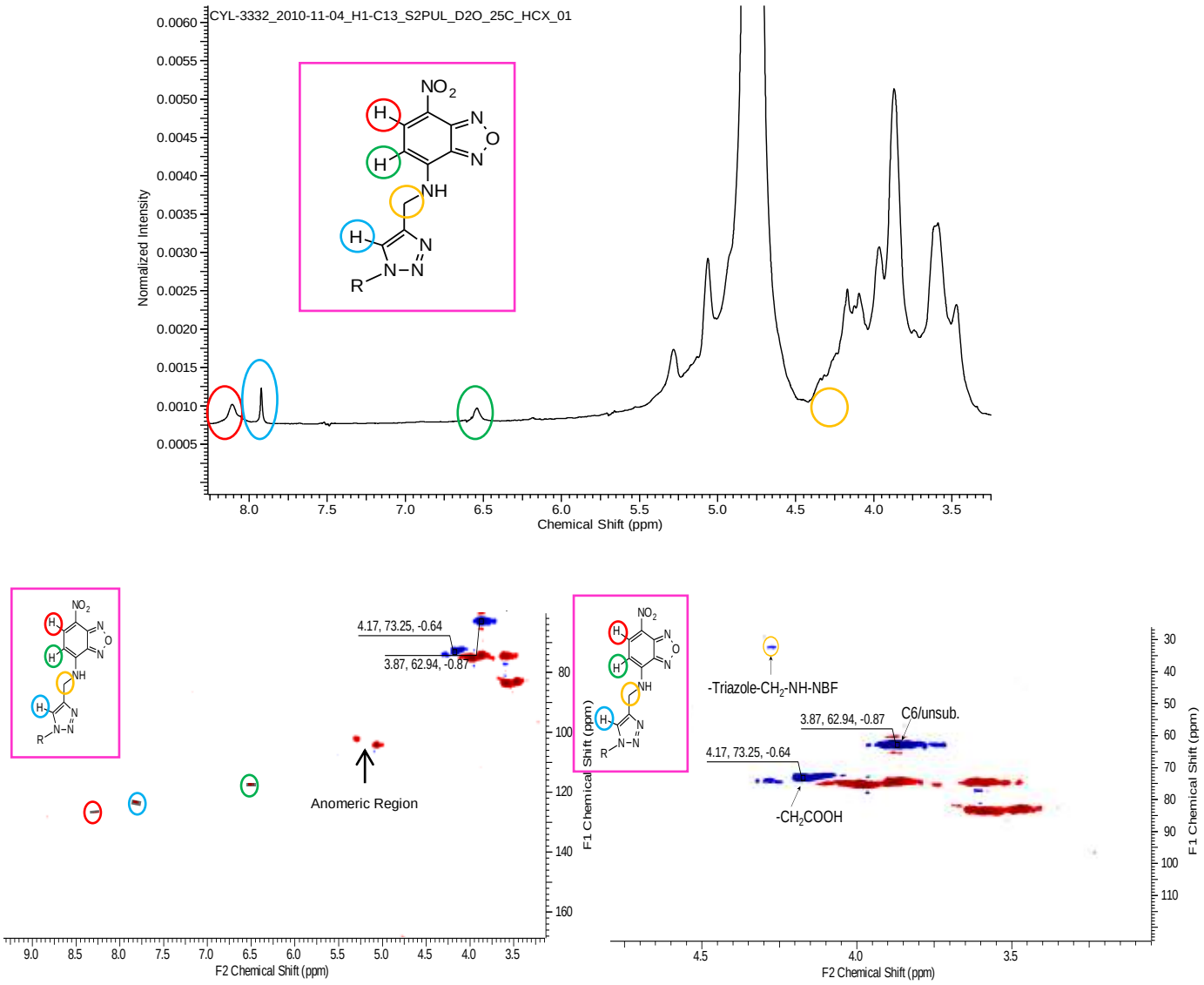


Figure 9.9, the 1D and 2D spectra of the 6-mono-nitrobenzofurazan-triazolyl-γCD derivative.

i) the anomeric region is around 5.2 ppm representing the eight anomeric protons (H1). This is the starting point for all the NMR-based CD assignment;

ii) the regions between 4.1 and 3.8 ppm and between 3.8 and 3.6 ppm comprise the frequencies of the H2, H3, H4, H5 and H6 protons.

In order to check the interactions between compound 2D and NBF-triazole-γCD the 1D and 2D spectra have been recorded and analyzed (*Figure 9.10*):

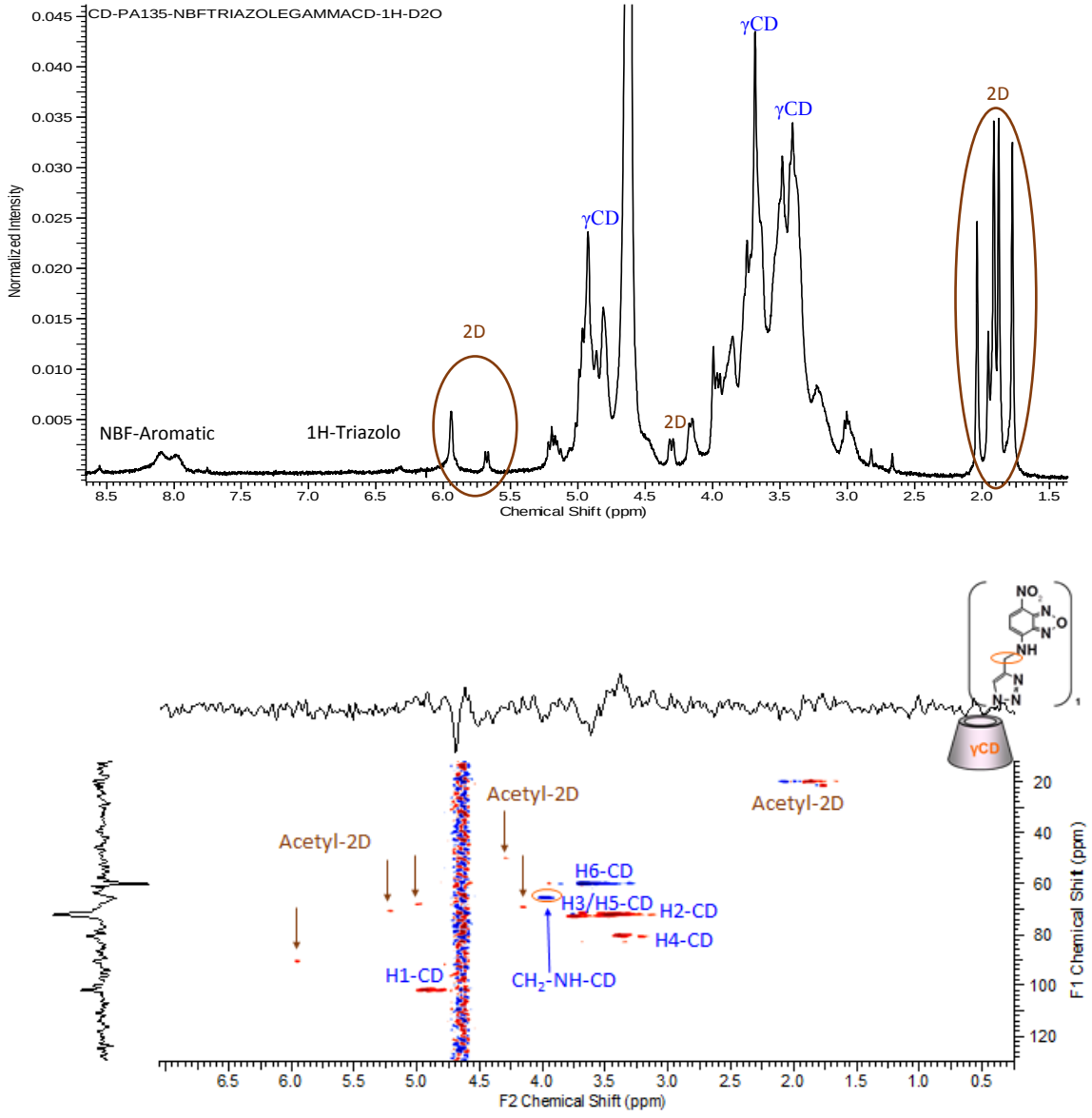


Figure 9.10: interactions between compound 2D and NBF-triazole-γCD the 1D and 2D spectra.

In *Figure 9.10* is reported the ^{1}H spectrum of compound 2D-γCD-biotin complex:

i) at 8.0 ppm there's the resonance of aromatic protons of NBF;

ii) from 3.0 to 5.0 ppm there are the signals of the γCD;

iii) from 1.8 to 2.0 and from 5.5 to 6.0 ppm there are signals of compound 2D, shifted compared the isolated molecule spectrum. In conclusions, the interaction detected by CE between and the 2D compound was confirmed by NMR spectroscopy.

9.6 Conclusions

During the erasmus period at *CycloLab*, the synthesis towards amino-bearing cyclodextrins and biotin-appended cyclodextrin were accomplished.

The introduction of a single amino group on the cyclodextrin scaffolds (alfa/beta/gamma) involved a three steps reaction. First, one single glucose unit of the cyclodextrin was selectively modified on the primary position with iodine (6-monosubstitution procedure), followed by substitution of the halogen with azido group and exhaustive reduction to amino group. The characterization of the amino cyclodextrins derivatives was done by NMR techniques (^{1}H and HSQC-DEPT-ed).

The amino-bearing CD derivatives were then functionalized with biotin and with the green fluorescent dye nitrobenzofurazane. The compounds were characterized by ^{1}H and HSQC-DEPT-ed NMR spectroscopy.

Suitable methods for selecting the best *N*-acetyl glucosamine derivatives's host were *ad-hoc* developed. Several cyclodextrin derivatives were screen by capillary electrophoresis in order to choose the most favorable interaction between the *N*-acetyl glucosamine derivatives and the cyclodextrins.

The gamma-CD derivatives were used to encapsulate drug 2D and the host-guest complexes were characterized by NMR techniques.

The project carried out during the 3 years-Ph.D. has had the objective to identify and synthesize new glycomimetics as molecular tools to study the *Hexosamine Biosynthetic Pathway* (HBP), which role is to regulate the proliferation and survival of cancer cells. The project has been funded by AIRC and the principal aim was to identify the *Adenocarcinoma of the Pancreatic Duct* (PDAC) as the target of research.

Some studies in last few years, have confirmed that both glucose and glutamine are critical for the PDAC progression, to coordinates the metabolism and to support the growth of pancreatic cancer by limiting oxygen availability.[1] HBP and PDAC are deeply linked together, in fact HBP uses glucose, glutamine, acetyl CoA, glucose 1,6-bisphosphate and uridine to synthesize *Uridine diphosphate-N-acetil-D-glucosamine* (UDP-GlcNAc), the substrate of the enzymes catalyzing the reactions of *N*- and *O*- glycosylation of the proteins. The synthesis of innovative chemical tools helps the understanding of the HBP pathway and its response in PDAC: new potential inhibitors, which are similar to the natural substrate of enzyme, can be recognized but trick the enzyme and block its activity in order to decrease the UDP-GlcNAc production and consequently modify the protein glycosylation. Tumoral cells follow important physiological processes such as proliferation, adhesion, motility, resistance to stress and metastasis, which are linked to glycosylation. Due to the important role of the HBP in the cells, alteration of this pathway can bring to alteration of *N*- and *O*-*glycosylation* and activate the *Unfolded Protein Response* (UPR) during the E*ndoplasmic Reticulum* (ER) stress. The behavior of normal and cancer cells stress is different and can be exploited to induce a secondary response in cancer cells that leads them to apoptosis.

The description of the research target helps the understanding of the design of molecular tools: the focus point is the inhibition of the enzyme *N-acetylglucosamine-phosphate mutase* (AGM1): its inhibition could represent the way to induce apoptosis in cancer cells.

Through the *Molecular Design*, descripbed in Chapter 4, a rational design of potential inhibitors has been done. This design is based on the similarity with the structures of the natural substrate of enzyme AGM1, with some modifications. All of the drawn structures have been used for M*olecular Docking* in order to get a first virtual screening on the compounds library: structures 6C and 4B seem to get better interaction than the natural substrate, while the others have docking score similar or worst with respect to the natural substrate. These preliminary results identify structures which are most likely capable to bind to the target enzyme: the information obtained by theoretical calculus are compared to the experimental results from the in *vitro test*. Starting from preliminary results of theoretical approach, the synthesis of compounds have been done following three different synthetic strategies: in *strategy 1*.

Modifications at the –OH in position C-1 and, at the same time, at the phosphate in position C-6, critical for the interaction with the enzyme and directly involved in the catalytic mechanism, have been done thanks to the removal of anomeric hydroxyl group and then the introduction of sulphate and sulphonate groups in position C-6. *Strategy 2.* introduces a change only in position C-6 with phosphoramidate, sulphate and sulphonate groups. *Strategy 3.* Different synthetic steps lead to the replacement of the anomeric hydroxyl group with the oxazoline ring. In Chapter 6 all the steps and reaction condition are described in details and in the Appendix (A.3) are shown the characterization (^{1}H, ^{13}C NMR spectra, *m/z* not reported) of all the synthetized compound.

The optimization of the analytical method on *High Performance Liquid Chromatography* is necessary in order to achieve experimental data on the ability of the designed compounds to inhibit the target enzyme, data to be compared to those obtained through a computational theoretical approach. Initial, experiments carried out directly using commercially available enzyme AGM1 didn't afford any reliable results. This suggested the idea to verify the inhibition activity of the compounds through the quantification of the end product of the HBP, the UDP-GlcNAc, using a cellular extract containing all HBP enzymes.

To this aim an HPLC method has been set-up for the quantification of UDP-GlcNAc produced using the cellular extract as enzyme source, and carrying out the reaction with the natural substrates GlcNAc-6P, UTP in the presence or not of the test molecules.

To settled the enzymatic reaction different procedures have been considered: first, the volume of cellular extract to be used for the reaction has been determined, Then the amount of substrate has been identified, finally the inhibitors have been tested. As reported, using 10 and 30 µL of extract, compounds 3B, 6C and 1F, lead to a decrease of production of UDP-GlcNAc .

The computational data "describes" the interaction between the enzyme and the molecules: compounds 6C and 3B have good affinity due to higher docking score, higher than the natural substrate, this data is in agreement with the enzymatic assay, in which these compounds show a capacity to decrease the production of UDP-GlcNAc.

Subsequently, the calculation of C Log*P* has confirmed the most apolar character of compound 3B in the acetylated form (compound 2B) in comparison with the strong polarity of phosphoramidate group in the acetylated form of compound 6C (compound 7C) and led us to study the kinetic of reaction of compound 3B. Some preliminary evaluation of the effect of compound 2B in a Triple Negative Breast Cancer (TNBC) cell model has been carried out.

In conclusion, the study of the target of this research, the HBP pathway, and the focus on the inhibition of AGM1 are the starting point for a complete project, that includes at first the design of a library of compound based on the structural properties of the natural substrate. the "in silico" evaluation of their interaction with the target enzyme, the synthesis and the screening through an enzymatic assay.

The tuning of the strategy of synthesis is important to obtain the compound for the in *vitro test*. The analytical method with HPLC gives results comparable to the docking scores, and then, after a calculation of C Log*P*, the test on cells gives the final results of potency of compound 3B (2B the acetylated form).

The last part, Chapter 9, describes the collaboration with *CycloLab* (Budapest): some compounds of the library possess chemical-physical characteristics that make their passage through cell membrane very harsh: they are very polar and some of them possess negative charges (sulphate, sulphonates, phosphoramidate). The generation of suitable delivery systems helps to target the enzymatic inhibitors inside cancer cells. To this aim the idea was to exploit opportunely functionalized cyclodextrins as Trojan Horse. After a preliminary study carried out on a selected compound (compound 2D) of the interaction between ligand-cyclodextrines, necessary to have information of the best cavity for compound 2D, the selected cyclodextrin was externally functionalized with ligands overexpressed on tumor cell lines such as biotin. A fluorescent tag has been also conjugated to allow cellular uptake studies. This preliminary work is still in progress.

Appendix
Experimental data

A.1 Technical features of Molecular Docking

All molecular docking calculations were performed using the AGM1 crystal structure from Candida Albicans (Protein Data bank code: 2dkc) co-crystallized with the natural substrate (GlcNAc-6-P).
The sequence identity over the entire protein between human AGM1 (Hs-AGM1) and AGM1 of Candida Albicans (Ca-AMG1) is 47%.
The docking scores of the project were computed with the software *Schrodinger 10.1 Maestro* and the docking calculations were performed using the *Glide Docking* module (Friesner et al., 2004) (Jorgensen et al., 1988). The force field used is OPLS_2005, preferred for biological systems and organic molecules. The docking of the potential inhibitors into the protein has been setted at pH 7 (± 0,3) because of the cellular pH, and sampling a box (18x18x18$Å^3$) centered on the enzyme active site.
All ligands were docked with the extra precision (XP) method and explicitly taking into account the conformational flexibility of ligands.

In order to be used as a docking receptor, protein structure should be processed: the first step was the preparation of the protein through the process of the workspace structure with some typical operations including addition of hydrogens atoms, assignment of atomic charges and bond orders, elimination of water molecules that are not involved in ligand binding. The refined tab of the structure is important for the optimization of the H-bond assignment and optimized within the Protein Preparation Wizard using force field OPLS_2005. After that, the ligand preparation is complete: in this case it must be used the same force field and pH previously set to get the lowest energy of conformations which may be generated in the absence of the receptor and subsequently docked. The next step was the generation of the grid: it contains part of the active site where the ligand can interact with the receptor.

The last part of the analysis is the docking of the ligand: after plugging the ligand inside the receptor grid and the definition of the precision of the docking, we got the pose of the ligand inside the enzymatic active site with its value of interaction.

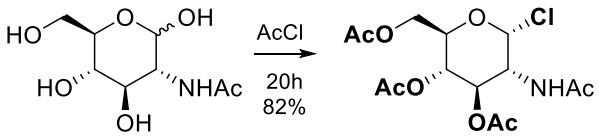
A.2 Synthetic Techniques and characterization methods

Chemical reactions were conducted under nitrogen in anhydrous solvents unless otherwise stated. Anhydrous solvents were purchased from commercial sources and were used without additional purification except for: (i) Acetone (Me_2CO) which was further dried over 3 Å molecular sieves; and (ii) tetrahydrofuran (THF). All other reagents obtained from commercial suppliers were used without further purification. Thin-layer chromatography (TLC) was performed on silica gel 60 F254 plates (Merck) with detection using UV light when possible, or by charring with a solution of concd. H_2SO_4/EtOH/H_2O (5:45:45) or a solution of $(NH_4)_6Mo_7O_{24}$ (21g), $Ce(SO_4)_2$ (1g), concd. H_2SO_4 (31mL) in water (500mL).

Normal-phase flash and gravity column chromatography were performed using silica gel (200–425 mesh 60 Å pore size - Merck) and ACS grade solvents. ^{1}H NMR spectra were recorded at 25°C, with a Varian Mercury 400 MHz instrument, and processed with MestReNova v6.0.2-5475 software. Chemical shift assignments, reported in ppm, are referenced to the corresponding solvent peaks (omitted in the peak assignment). Peak locations were referenced using either tetramethyl-silane (TMS) or residual nondeuterated solvent as an internal standard. ^{13}C NMR chemical shifts are reported to the first decimal place unless peaks are very close wherein for such instances values are reported to a second decimal place.

A.3 Steps of synthesis

Scheme A - *N*-acetylglucosamine-1-deoxy from 6-phosphoramidate, derivative from GlcNAc

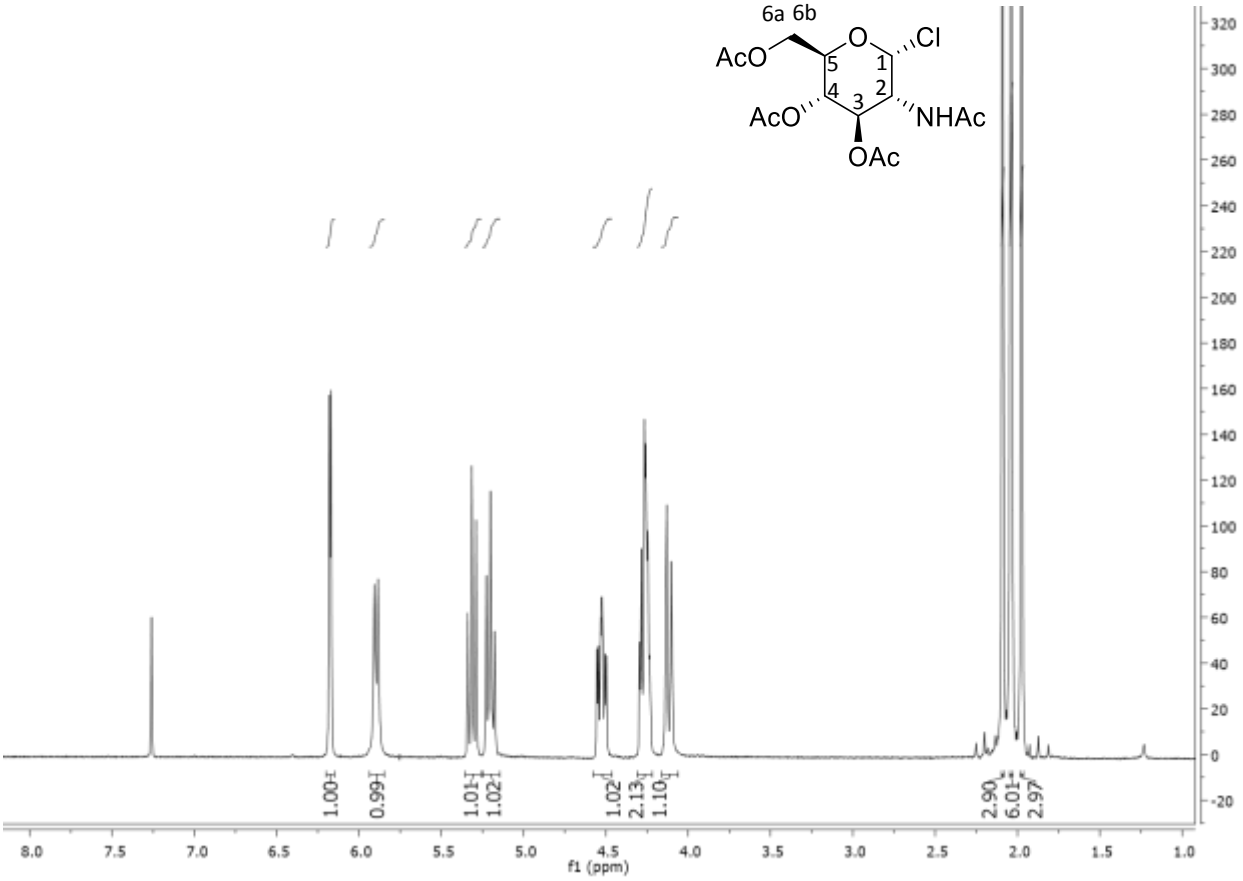


GlcNAc 1A

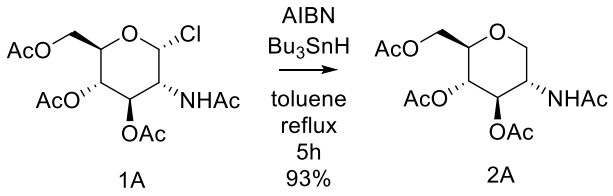
	g	MW	mmol	eq	d (g/mL)	V (mL)	[M]
GlcNAc	10,07	221	45,52	1			
AcCl						35	

Acetylchloride was added drop wise to the *N*-acetyl-D-glucosamine under stirring at r.t. for 20h. at complete disappearance of the starting material, the product was obtain by extraction with H_2O-ice/CH_2Cl_2 and by washing the organic phase with $NaHCO_3$/CH_2Cl_2. The purification was done with flash chromatography Pet Et:AcOEt 3:7, rf 0,35 to afford pure compound 1A (10,38 g – 82% yield), *m/z* 365,09.

The characterization of compound 1A (^{1}H spectrum reported – ^{13}C, 2D and *m/z* spectra not reported) is in according to the literature.[33]



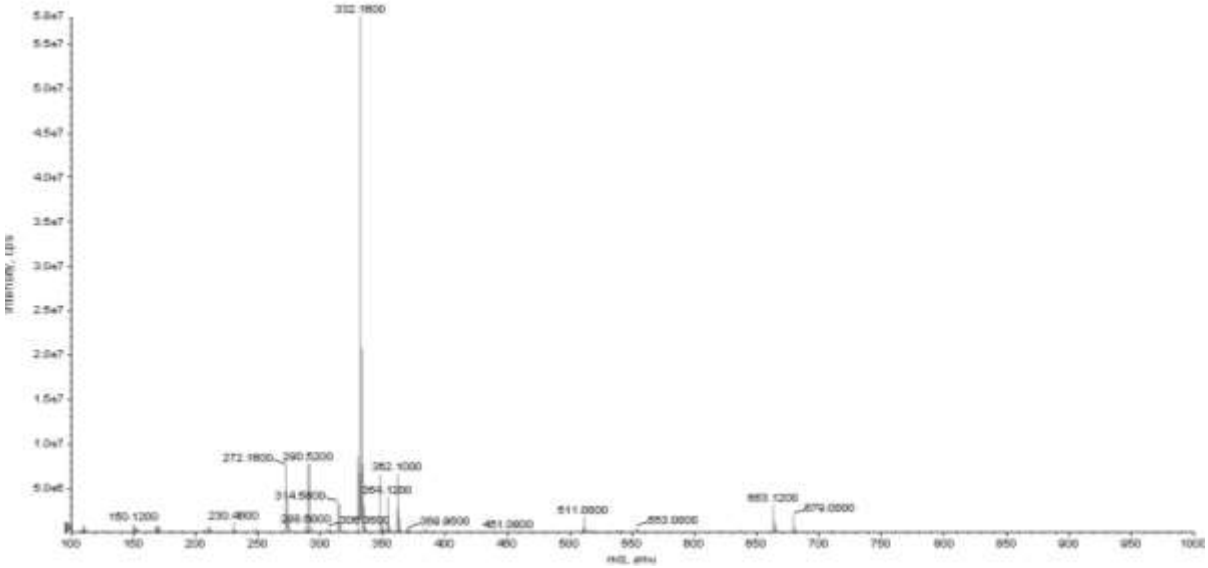
[1]H NMR (400 MHz, Chloroform-d) δ 6.17 (d, *J* = 3.6 Hz, 1H, **H-1**), 5.90 (d, *J* = 8.6 Hz, 1H, ***N*H**), 5.31 (t, *J* = 10.1 Hz, 1H, **H-4**), 5.20 (t, *J* = 9.8 Hz, 1H, **H-3**), 4.53 (ddd, *J* = 10.6, 8.8, 3.7 Hz, 1H, **H-2**), 4.31 – 4.22 (m, 2H, **H-5**, **6a**), 4.12 (bd, *J* = 10.6 Hz, 1H, **H-6b**), 2.09 (s, 3H, **Ac**), 2.04 (s, 6H, **2Ac**), 1.98 (s, 3H, **Ac**).

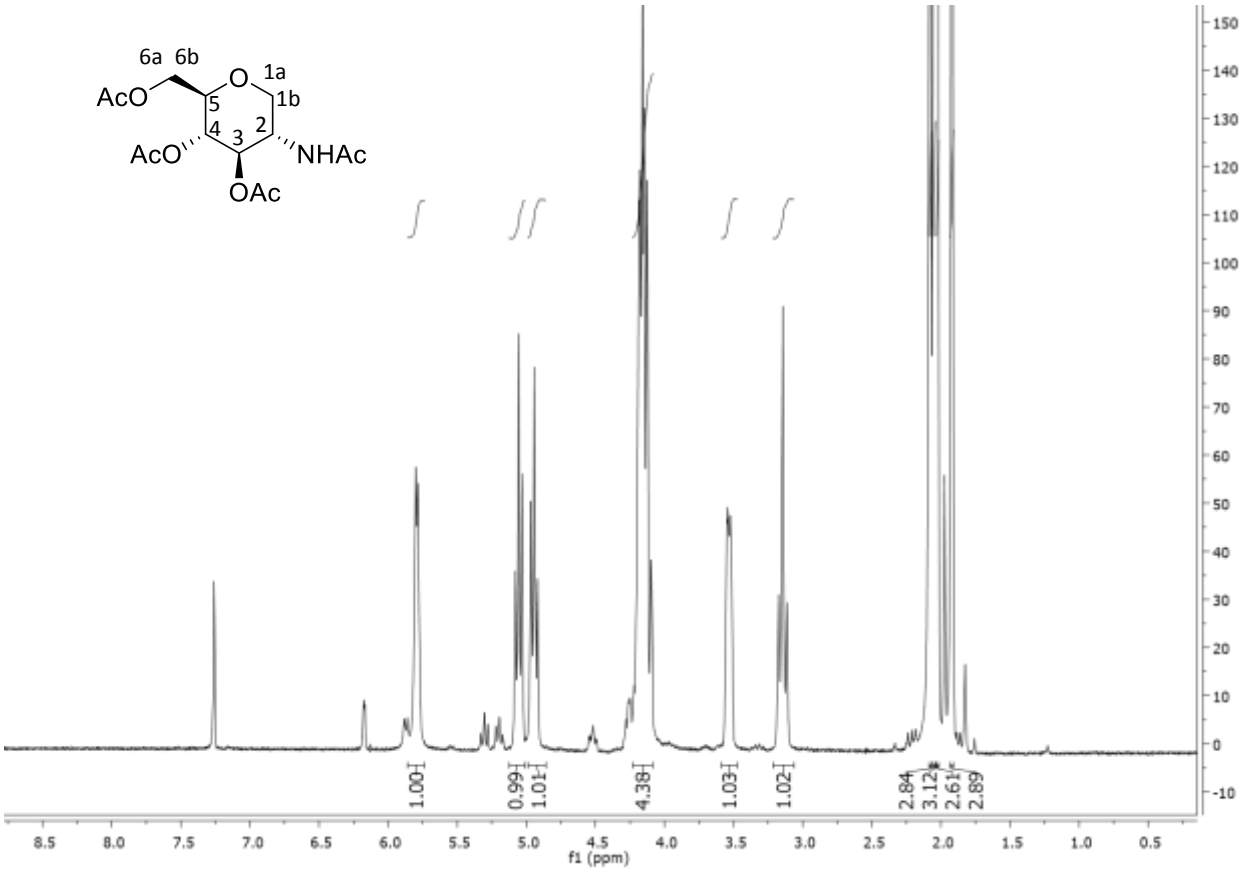


	g	MW	mmol	eq	d (g/mL)	V (mL)	[M]
1A	10,15	365,09	27,80	1			
Bu_3SnH	12,14	291,06	41,7	1,5	1,08	11,23	
AIBN	cat						
toluene						108	

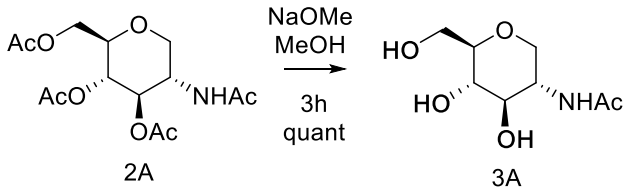
To a suspension of 1A in dry toluene, under inert atmosphere, AIBN was added. Then Bu_3SnH was added and the reaction was stirred at 120°C for 12h. The reaction was quenched with KF 2M. The product was extracted with CH_2Cl_2/H_2O. The organic layer was dried over anhydrous $NaSO_4$, filtered and the organic layer removed under reduced pressure. The purification was done with flash chromatography AcOEt, rf 0,35 to afford pure 2A (8,61 g – quant), m/z_{calc} 331.13 (100.0%), 332.13 (15.1%), 333.13 (1.6%), 333.13 (1.1%), m/z $[M + H]^+$ = 332.16, $[M + Na]^+$ = 354.12.

The characterization of compound 2A (^1H spectrum reported – ^{13}C and 2D spectra not reported) is in according to the literature. [33]





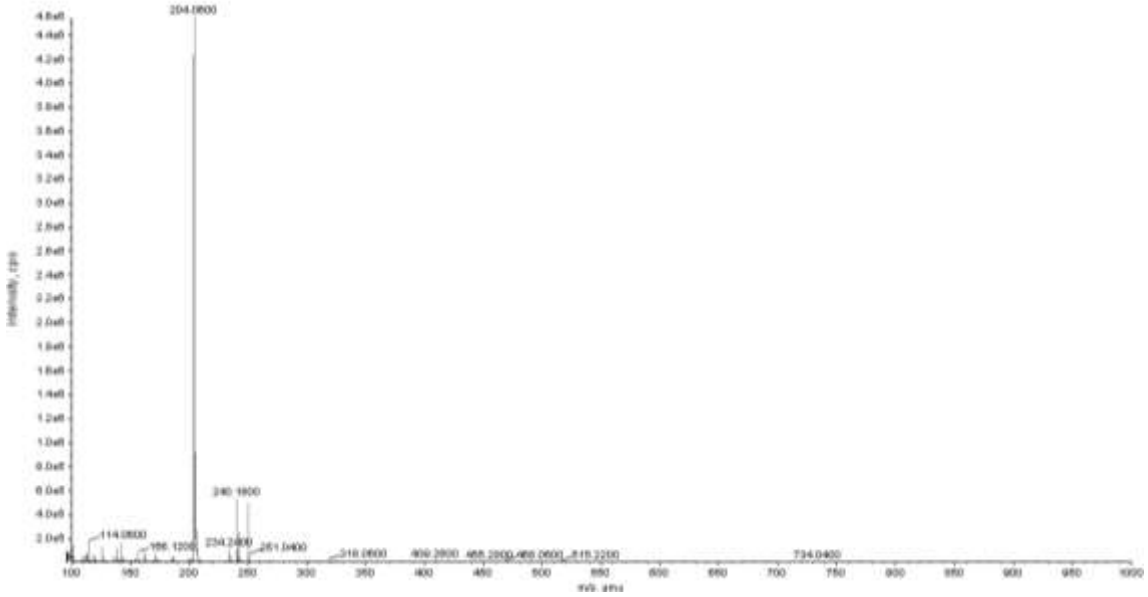
[1]H NMR (400 MHz, Chloroform-d) δ 5.78 (t, *J* = 6.9 Hz, 1H, NH), 5.06 (t, *J* = 9.6 Hz, 1H, H-4), 4.94 (t, *J* = 9.7 Hz, 1H, H-3), 4.23 – 4.10 (m, 4H, **H-2**,**6a**, **1a**, **1b**), 3.54-3.51 (m, 1H, **H-5**), 3.14 (t, *J* = 12.0 Hz, 1H, **H-6b**), 2.08 (s, 3H, **Ac**), 2.05 (s, 3H, **Ac**), 2.02 (s, 3H, **Ac**), 1.92 (s, 3H, **Ac**).

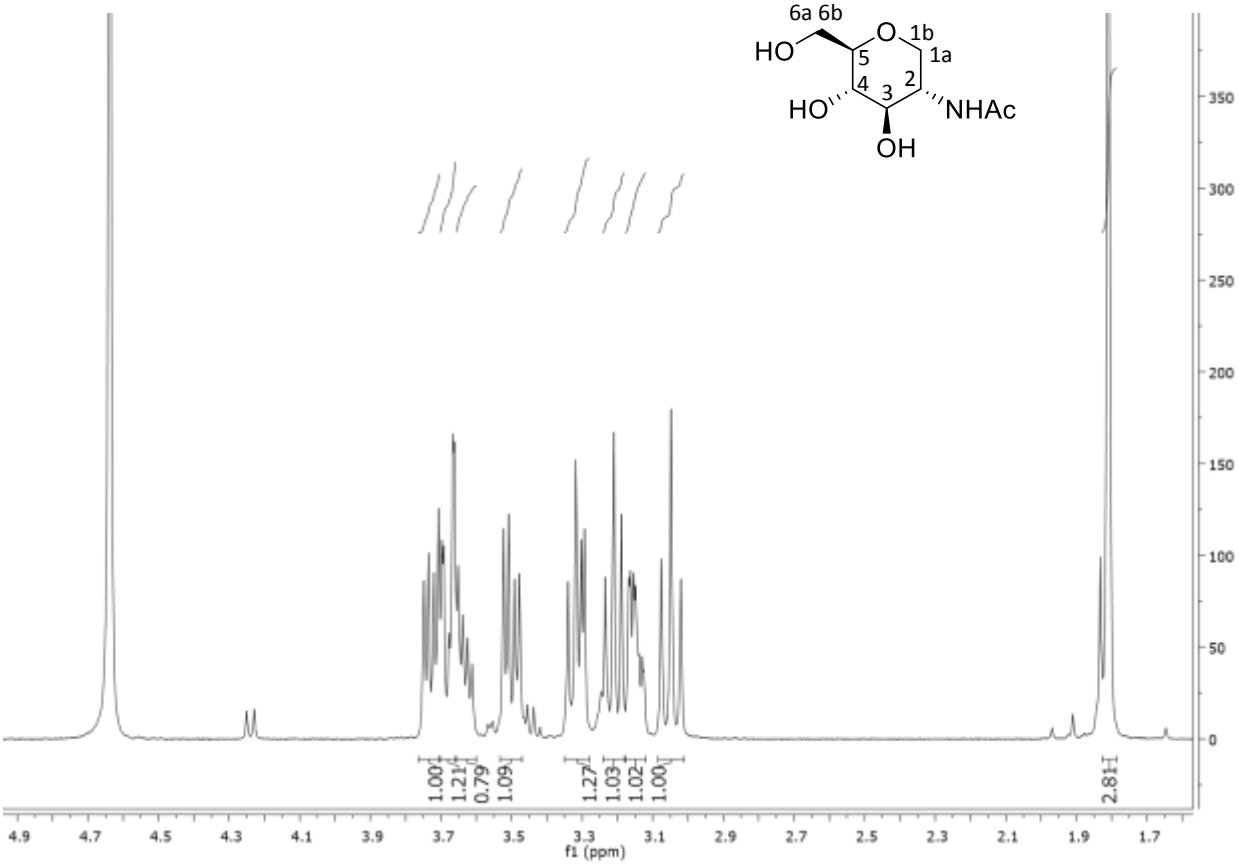


	g	MW	mmol	eq	d (g/mL)	V (mL)	[M]
2A	4,01	331,13	12,12	1			
NaOMe	0,540	54,02	9,99	1			
MeOH						100	0,1

Compound 2A was dissolved in dry MeOH under Argon atmosphere, and NaOMe was added. After 3h the stirring was stopped and the reaction was neutralized with Amberlite 5% HCl to pH 6. After filtration the solvent was removed under reduced pressure affording compound 3A (2,68 g – quant), m/z_{calc}: 205.10 (100.0%), 206.10 (8.7%), 207.10 (1.0%), m/z $[M - H]^-$ = 204.06.
The compound doesn't need purification.

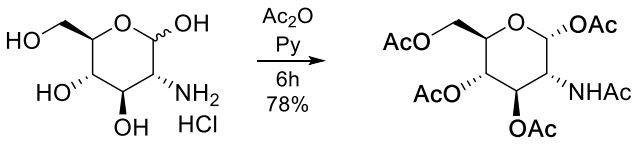
The characterization of compound 3A, ***final compound*** (^{1}H spectrum reported – ^{13}C and 2D spectra not reported) for in vitro test, is in according to the literature.[33]





^{1}H NMR (400 MHz, Deuterium Oxide) δ 3.73 (dd, *J* = 11.2, 5.2 Hz, 1H, **H-1a**), 3.68 (dd, *J* = 12.3, 1.9 Hz, 1H, **H-6a**), 3.64 (dt, *J* = 10.7, 5.6 Hz, 1H, **H-2**), 3.50 (dd, *J* = 12.3, 5.7 Hz, 1H, **H-6b**), 3.32 (t, *J* = 9.3 Hz, 1H, **H-3**), 3.21 (t, *J* = 9.1 Hz, 1H, **H-4**), 3.15 (m, 1H, **H-5**), 3.05 (t, *J* = 11.1 Hz, 1H, **H-1b**), 1.81 (s, 3H, **Ac**).

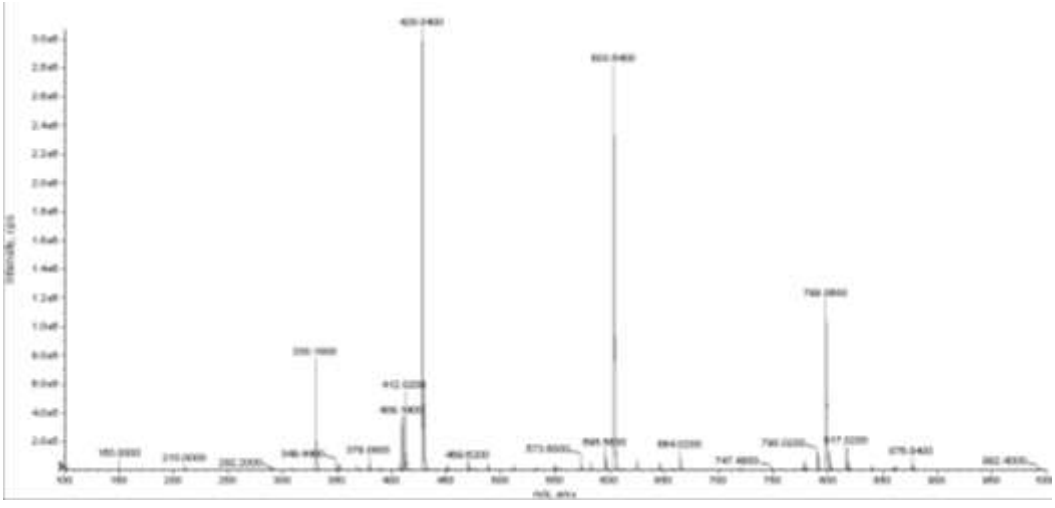
Scheme B - 2-Methyl-(3,4,6-tri-O-acetyl-1,2-dideoxy-α-D-glucopyrano)-[2,1-d]-2-oxazoline [41]



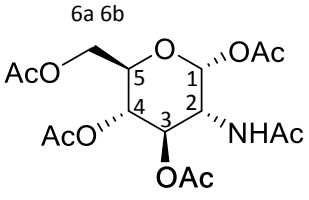
Glucosamine HCl 1B

	g	MW	mmol	eq	d (g/mL)	V (mL)	[M]
GlcNH$_2$-HCl	1,06	215,63	4,92	1			
Ac$_2$O	5,02	102,09	49,20	10	1,08	13,88	
Py	13,62	79,10	172,20	35	0,981	4,65	

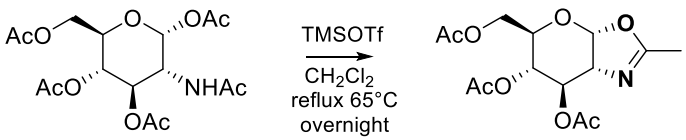
Glucosamine hydrochloride was dissolved in 66 mL on Py; acetic anhydride was added dropwise at 0°C. The reaction was stirred for 4 hours at r.t. The product was diluted with AcOEt and washed with HCl 5% solution, then with H_2O. The organic layer was dried over anhydrous $NaSO_4$, filtered and the solvent was removed under reduced pressure. The crude product was crystallized from isopropanol to afford 1B (1,55 g – 78% yield), m/z_{calc} 389.13 (100.0%), 390.14 (17.3%), 391.14 (2.1%), 391.14 (1.4%), *m/z* $[M+Na]^+$= 412.02, $[M+K]^+$= 428.04.



^{1}H NMR (400 MHz, Chloroform-d) δ 6.14 (d, *J*= 3.6 Hz, 1H, **H1**), 5.56 (d, *J*=2 9.0 Hz, 1H, ***N*H**), 5.24-5.14 (m, 2H, **H3**, **H4**), 4.49-4.41 (m, 1H, **H2**), 4.22 (dd, *J*= 4.2,12.6 Hz, 1H, **H6a**), 4.03 (dd, *J*=2.4, 12.3 Hz, 1H, **H6b**), 3.98-3.93 (m,1H, **H5**), 2.16 (s, 3H, OCOC**H$_3$**), 2.05 (s, 3H, OCOC**H$_3$**), 2.02 (s, 3H, OCOC**H$_3$**) 2.01 (s, 3H, OCOC**H$_3$**), 1.90 (s, 3H, *N*HCOC**H$_3$**)



The characterization of compound 1B is in according to the literature. [42]

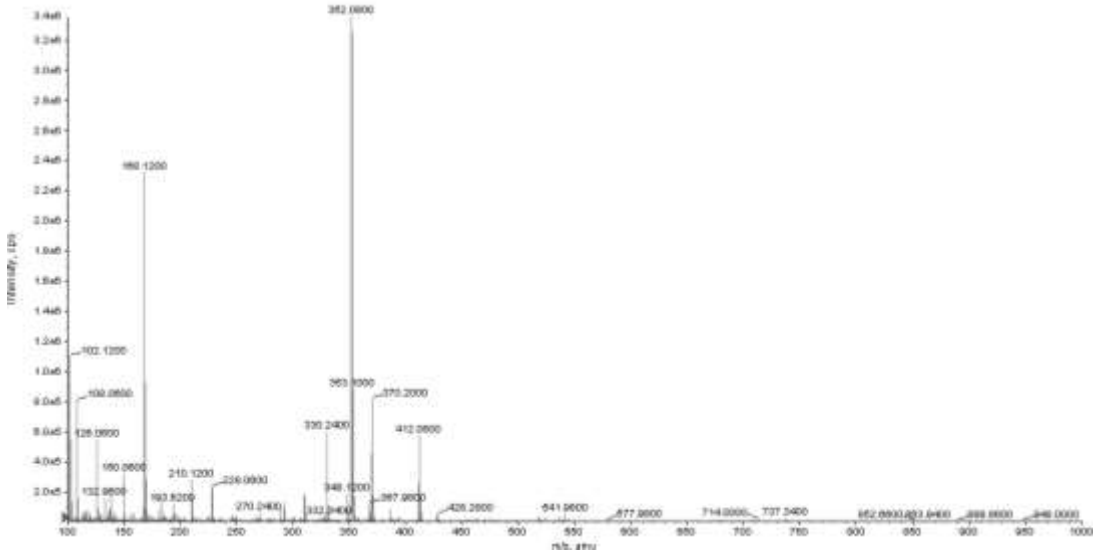


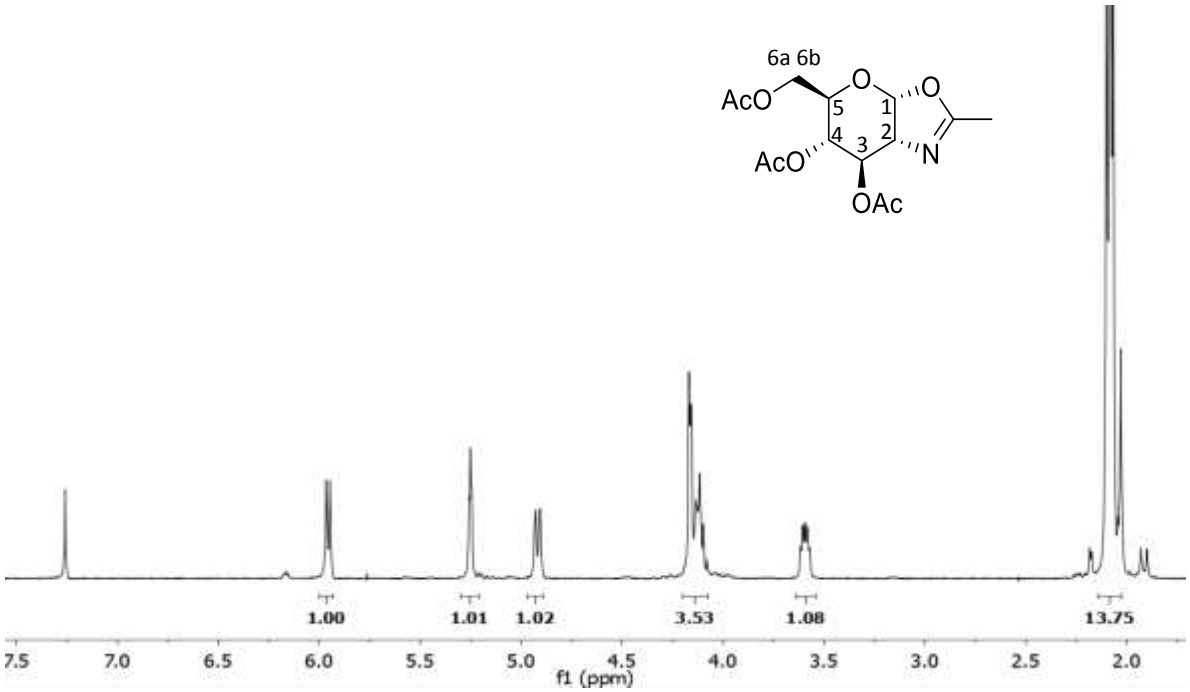
1B 2B

	g	MW	mmol	eq	d (g/mL)	V (mL)	[M]
1B	1,55	389,13	3,97	1			
TMSOTf	1,41	222,29	6,36	1,6	1,23		
CH_2Cl_2 dry						30	0,1

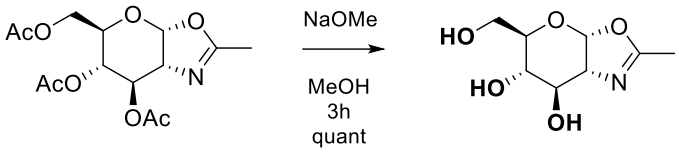
Compound 1B was dissolved in dry CH_2Cl_2 under inert atmosphere and the reaction was cooled to 0°C, then TMSOTf was added and the reaction was left stirring overnight at 65°C. The reaction was quenched though the addition of 2 mL of triethylamine and the solvent was removed under reduced pressure. The crude compound was purified with flash chromatography Et/AcOEt 8:2, rf 0,31 affording 1,05 g of 2B (yield 81%), m/z_{calc}: 329.11 (100.0%), 330.11 (15.1%), 331.12 (1.6%), 331.12 (1.1%), *m/z* $[M + Na]^+$= 352.08, $[M + K]^+$=367.98.

The characterization of compound 2B, ***final compound*** ready for the *cellular test*, is in according to the literature. [41]





^{1}H NMR (400 MHz, Chloroform-d) δ 5.95 (d, *J* = 7.5, 1.9 Hz, 1H, **H1**), 5.30 – 5.21 (m, 1H, **H3**), 4.92 (d, *J* = 9.2, 1.8 Hz, 1H, **H4**), 4.20 – 4.08 (m, 3H, **H2**, **H6a**, **H6b**), 3.64 – 3.54 (m, 1H), 2.14 – 2.02 (m, 12H, 4 **CH_3**).



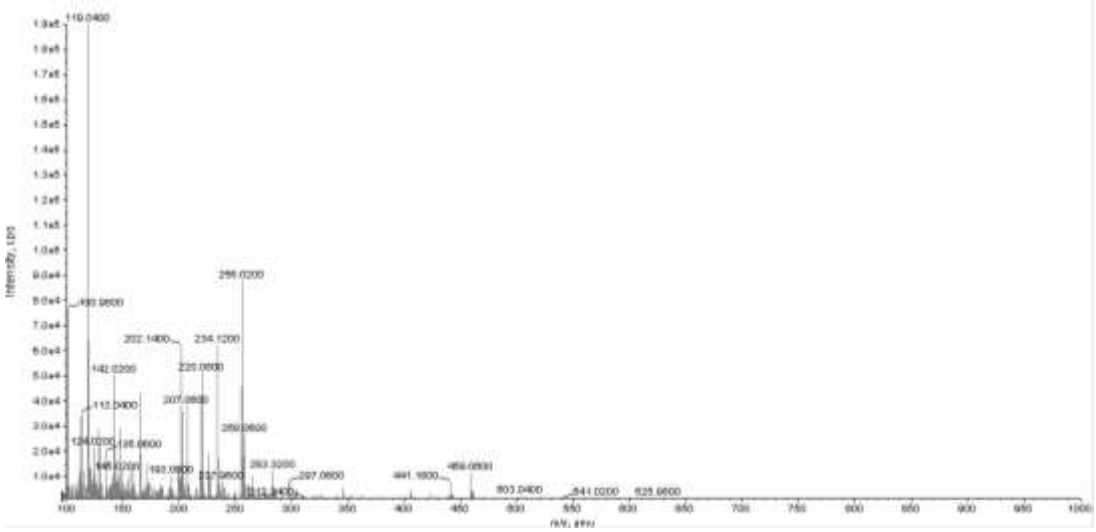
2B 3B

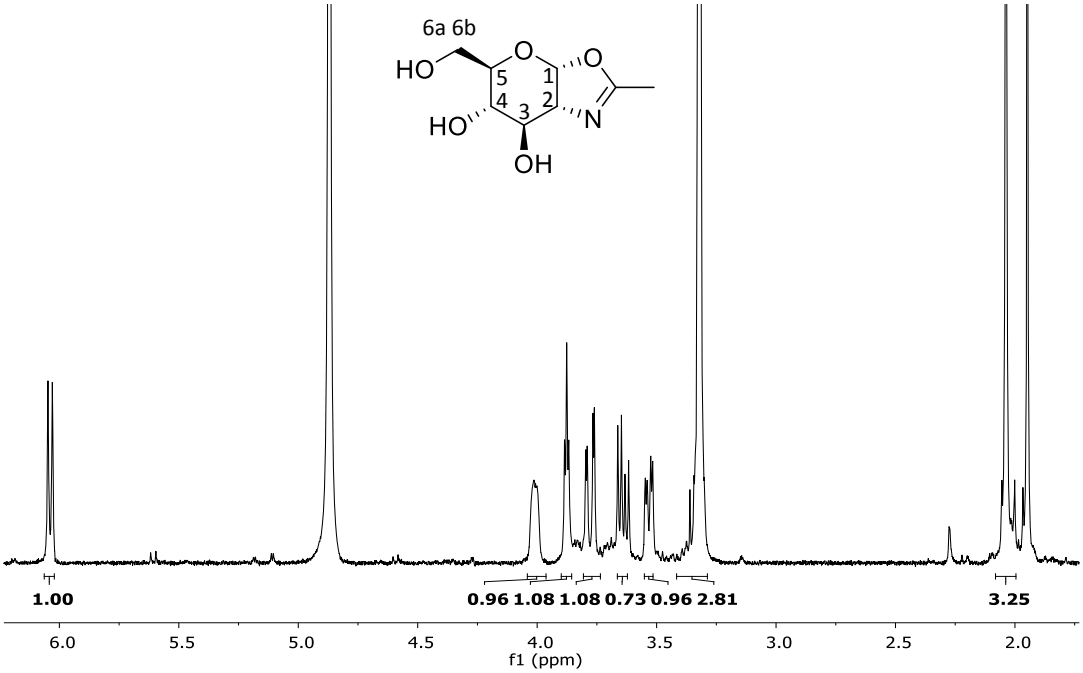
	g	MW	mmol	eq	d (g/mL)	V (mL)	[M]
2B	1,05	329,11	3,03	1			
NaOMe	0,16	54,02	3,03	1			
MeOH dry						30	0,1

Compound 2B was dissolved in 30 mL of dry MeOH under inert atmosphere, then NaOMe was added to the solution and the reaction was stirred for 3h. The reaction was neutralized with Amberlite 5% HCl to pH 7. After filtration the solvent was removed under reduced pressure affording compound 3B (0,85 g – quant), m/z_{calc}: 203.08 (100.0%), 204.08 (8.7%), 205.08 (1.0%), *m/z* $[M - H]^-$ = 202.14.

The compound doesn't need purification.

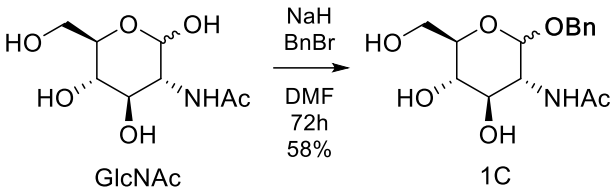
The characterization of compound 3B, ***final compound*** (^{1}H spectrum and *m/z*) for *enzymatic test*, is in according to the literature. [41]





^{1}H NMR (400 MHz, Methanol-d) δ 6.03 (d, *J* = 7.3 Hz, 1H, **H1**), 3.99 (bs, 1H, **H2**), 3.87 (t, *J* = 3.7 Hz, 1H, **H3**), 3.77 (dd, *J* = 12.0, 2.5 Hz, 1H, **H6a**), 3.67 – 3.61 (dd, *J* = 6.8 Hz, 1H, **H6b**), 3.55 – 3.50 (m, 1H, **H4**), 3.33 (ddd, *J* = 12.0 Hz, 2.4Hz, 6.4 Hz, 1H, **H5**), 2.03 (s, 3H, **CH_3**).

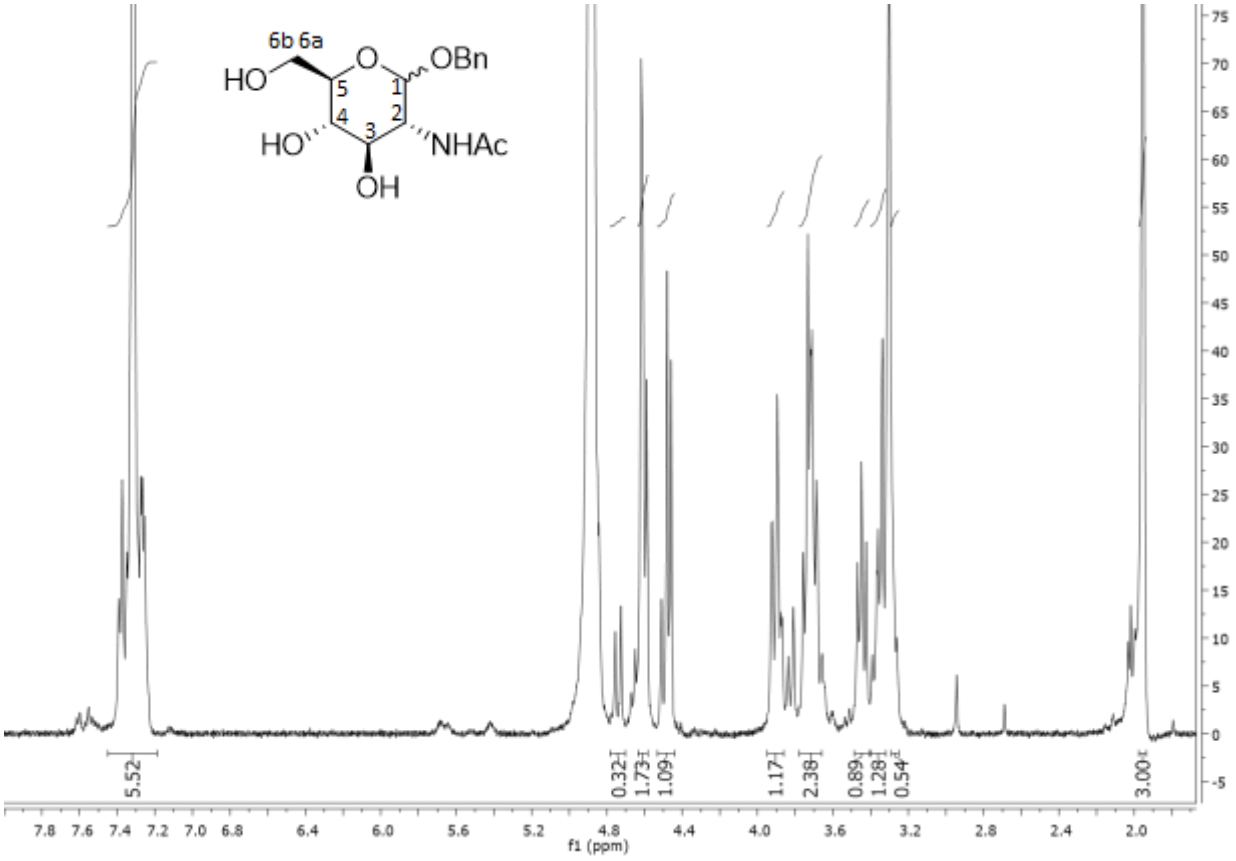
Scheme C - 6-phosphoramidate derivate from *N*-acetylglucosamine



	g	MW	mmol	eq	d (g/mL)	V (mL)	[M]
GlcNAc	10	221	45,24	1			
DMF		73,10				50	
NaH	1,41	24	58,81	1,3			
BnBr	15,47	171,03	90,48	2	1,44	10,74	

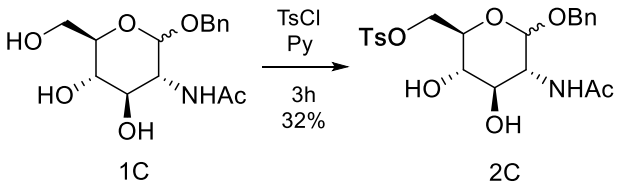
The NaH was added to a solution of GlcNAc in DMF dry under inert atmosphere, then benzylbromide was added to the mixture at 0°C. The reaction was left stirring overnight at room temperature. Then the reaction was quenched with MeOH and the solvent evaporated under reduced pressure. The product was washed with diethylether and purified with flash chromatography DCM:MeOH 9:1, rf 0,35 (8,16 g – 58% yield), m/z_{calc} 311.14 (100.0%), 312.14 (16.2%), 313.14 (1.2%), 313.14 (1.2%).

The characterization of compound 1C (^{1}H spectrum reported – ^{13}C, 2D and *m/z* spectra not reported) is in according to the literature.[37]



Mixture α/β ~2/8, major isomer peeks β

^{1}H NMR (400 MHz, Methanol-d) δ 7.47 – 7.19 (m, 5H), 4.87 – 4.83 (m, 1H, C**H**Ph), 4.60 (d, *J* = 11.8 Hz, 1H, C**H**Ph), 4.47 (d, *J* = 8.3 Hz, 1H, **H-1**), 3.91 (bd, *J* = 10.9 Hz, 1H, **H-6a**), 3.78 – 3.67 (m, 2H, **H-6b**,**2**), 3.48 – 3.42 (m, 1H, **H-5**), 3.39 – 3.28 (m, 2H, **H-3**,**4**), 1.96 (s, 3H, **Ac**).

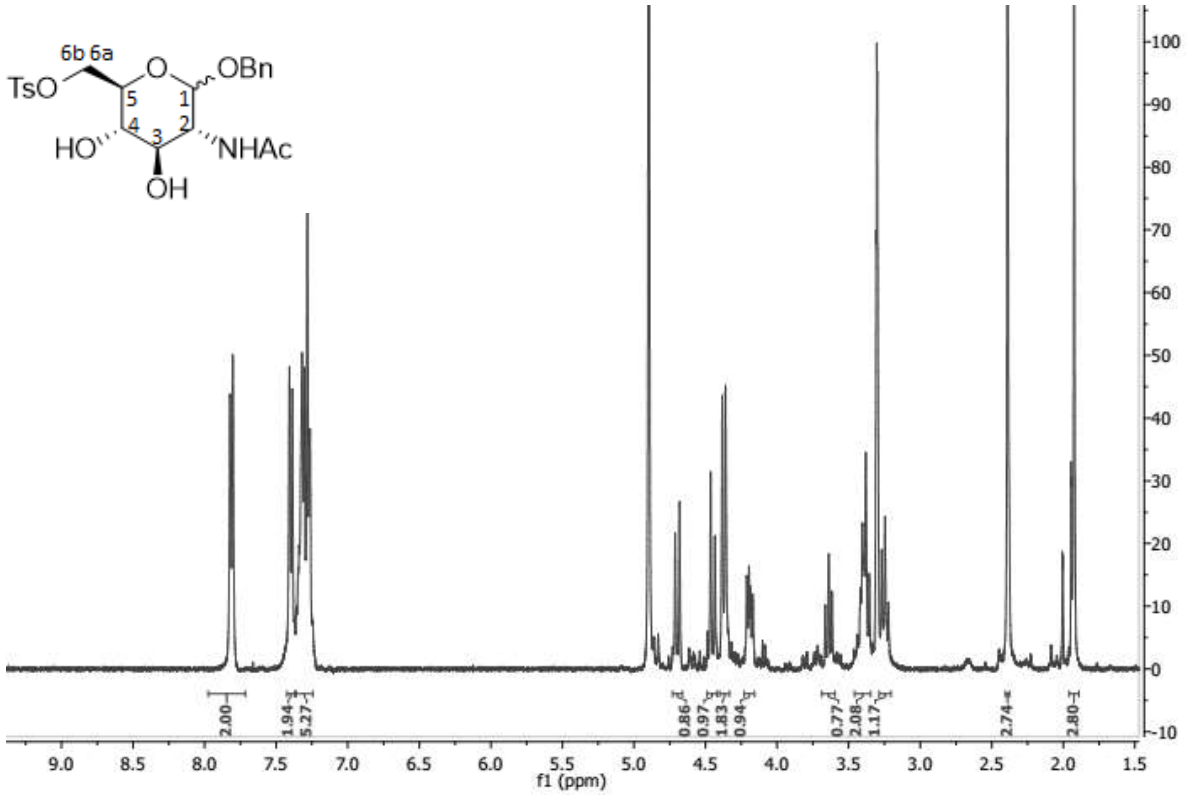


1C 2C

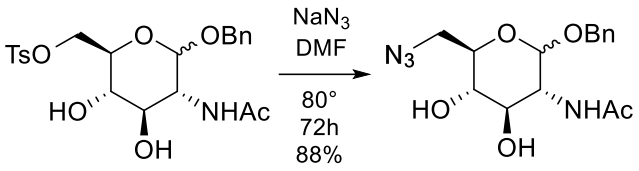
	g	MW	mmol	eq	d (g/mL)	V (mL)	[M]
1C	2	311,33	6,40	1			
TsCl	3,66	190,65	19,21	3			
Pyridine						32	
DMAP	0,078	122,17	0,64	0,1	1,44	10,74	

Compound 1C was dissolved in dry Py (20mL), under inert atmosphere, and DMAP was added. to TsCl was dissolved in Py (12mL) and added dropwise at the solution. After 3h the reaction was quenched with EtOH, was dissolved in $CHCl_3$ and washed and the final compound was purified with flash chromatography DCM:MeOH 9:1, rf 0,32 (0.8 g – 32% yield), m/z_{calc} 465.15 (100.0%), 466.15 (23.8%), 467.14 (4.5%), 467.15 (2.7%), 467.15 (1.6%), 468.14 (1.1%).

The characterization of compound 2C (^{1}H spectrum reported – ^{13}C, 2D and *m/z* spectra not reported) is in according to the literature.[38]



^{1}H NMR (400 MHz, Methanol-d) δ 7.81 (d, *J* = 7.3 Hz, 2H, **HAr**), 7.40 (d, *J* = 7.9 Hz, 2H, HAr), 7.29 (m, 5H, **HAr**), 4.70 (d, *J* = 12.3 Hz, 1H, C**H**Ph), 4.45 (d, *J* = 12.3 Hz, 1H, C**H**Ph), 4.37 (bd, *J* = 9.5 Hz, 1H, **H-1**, **6a**), 4.19 (dd, *J* = 10.4, 5.8 Hz, 1H, **H-6b**), 3.64 (t, *J* = 9.2 Hz, 1H, **H-2**), 3.45 – 3.34 (m, 2H, **H-3**,**4**), 3.25 (t, *J* = 9.2 Hz, 1H, **H-5**), 2.39 (s, 3H, C**H**$_3$Ts), 1.93 (s, 3H, **Ac**).

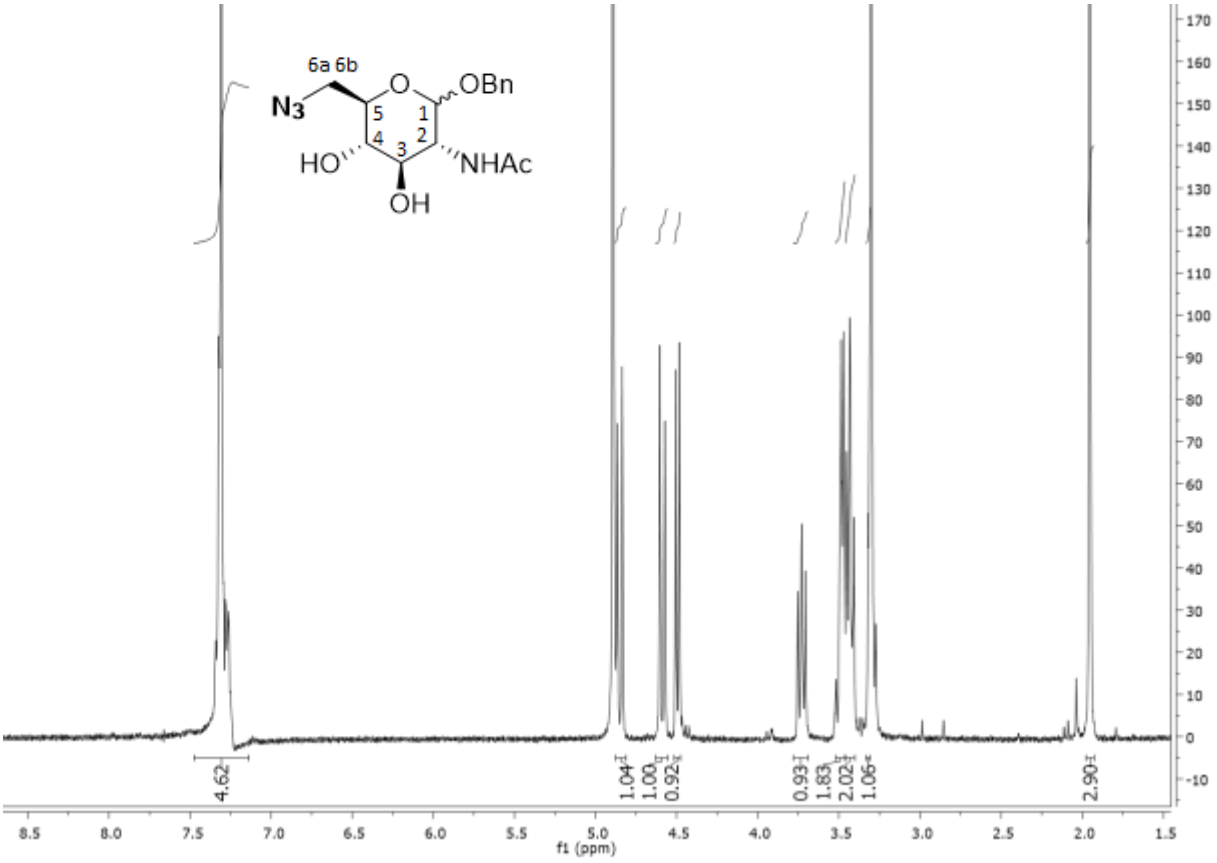


2C 3C

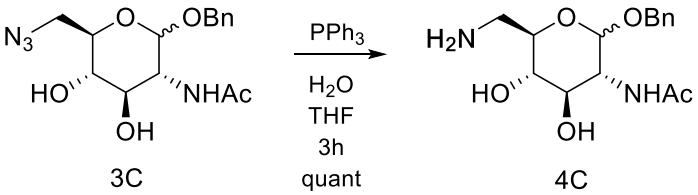
	g	MW	mmol	eq	d (g/mL)	V (mL)	[M]
2C	0,242	465,14	0,64	1			
NaN_3	0,0806	65,01	1,24	2			
DMF						3,72	0,1

Product 2C was dissolved in 19 mL of DMF and NaN_3 was added. The reaction was stirred at 80°C for 72 h. Then solvent was evapored and the crude compound was purified with flash chromatography AcOEt:MeOH 9:1, rf 0,28 (0,18 g – 88% yield), m/z_{calc} 336.14 (100.0%), 337.15 (16.2%), 337.14 (1.5%), 338.15 (1.2%), 338.15 (1.0%).

The characterization of compound 3C (^{1}H spectrum reported – ^{13}C, 2D and *m/z* spectra not reported) is in according to the literature.[38]



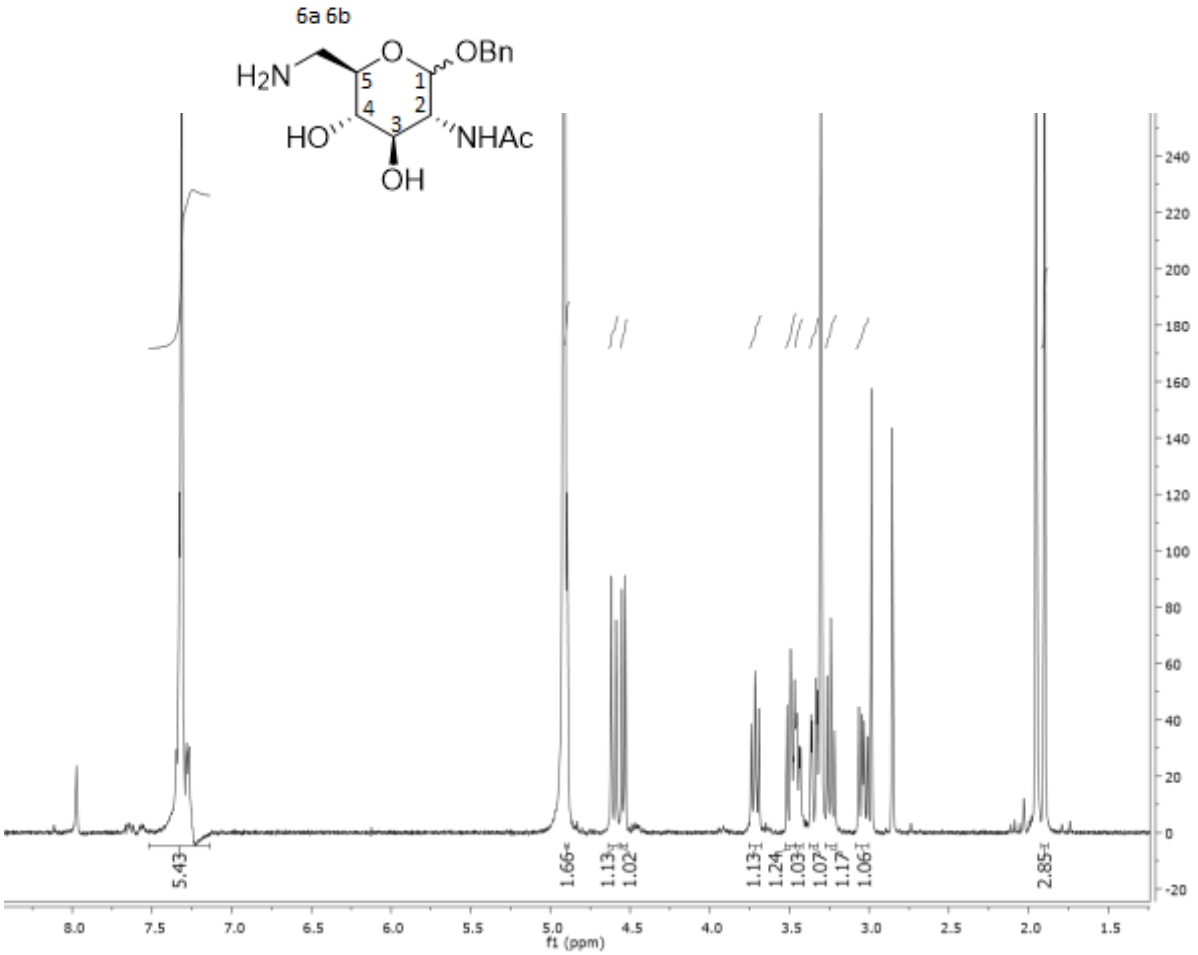
^{1}H NMR (400 MHz, Methanol-d) δ 7.45 – 7.16 (m, 5H, **H-**Ar), 4.85 (d, *J* = 12.1 Hz, 1H, C**H**Ph), 4.59 (d, *J* = 12.0 Hz, 1H, C**H**Ph), 4.50 (d, *J* = 8.4 Hz, 1H, **H-1**), 3.78 – 3.69 (m, 1H, **H-2**), 3.54 – 3.39 (m, 4H, **H-3**,**4**,**6a**,**6b**), 3.31 (m, 1H, **H-5**), 1.95 (s, 3H, **Ac**).



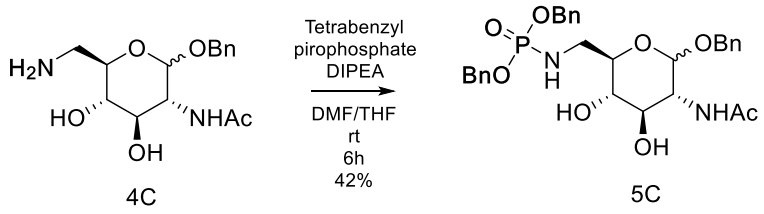
	g	MW	mmol	eq	d (g/mL)	V (mL)	[M]
3C	0,128	336,14	0,38	1			
PPh_3	0,199	262,30	0,76	2			
H_2O						0,20	
THF							3

Compound 3C was dissolved in THF and PPh_3 and H_2O were added. The reaction was stirred overnight at 60°C. Then the reaction was quenched with MeOH and the solvent evapored. The product was extracted with H_2O/AcOEt, the organic layer was dried over anhydrous $NaSO_4$, filtered and the solvent was removed under reduced pressure. The crude was purified by flash chromatography AcOEt:MeOH:NH_3 6:4:0,5, rf 0,35 to afford product 4C (0,11 g – quant), m/z_{calc} 310.15 (100.0%), 311.16 (16.2%), 312.16 (1.2%), 312.16 (1.0%).

The characterization of compound 4C (^{1}H spectrum reported – ^{13}C, 2D and *m/z* spectra not reported) is in according to the literature. [38]



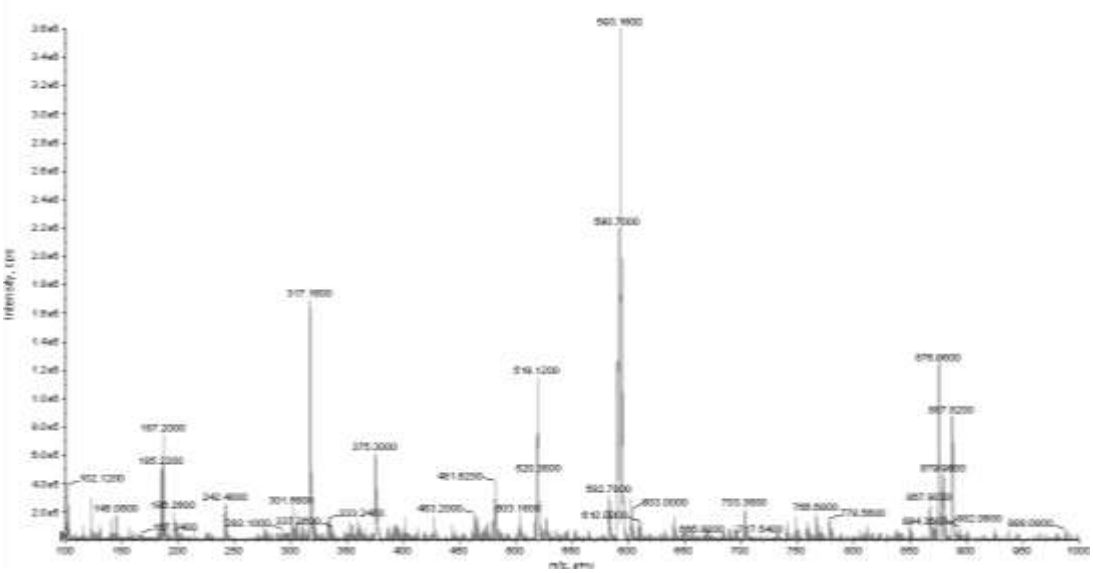
^{1}H NMR (400 MHz, Methanol-d) δ 7.41 – 7.24 (m, 5H, **H**-Ar), 4.90 (1H, C**H**Ph), 4.60 (d, *J* = 12.2 Hz, 1H, C**H**Ph), 4.54 (d, *J* = 8.5 Hz, 1H, **H-1**), 3.80 – 3.67 (m, 1H, **H-2**), 3.53 – 3.42 (m, 2H, **H-3**,**5**), 3.34 (dd, *J* = 13.2, 3.0 Hz, 1H, **H-6a**), 3.24 (t, *J* = 9.2 Hz, 1H, **H-4**), 3.04 (dd, *J* = 13.2, 8.4 Hz, 1H, **H-6b**), 1.90 (s, 3H, **Ac**).

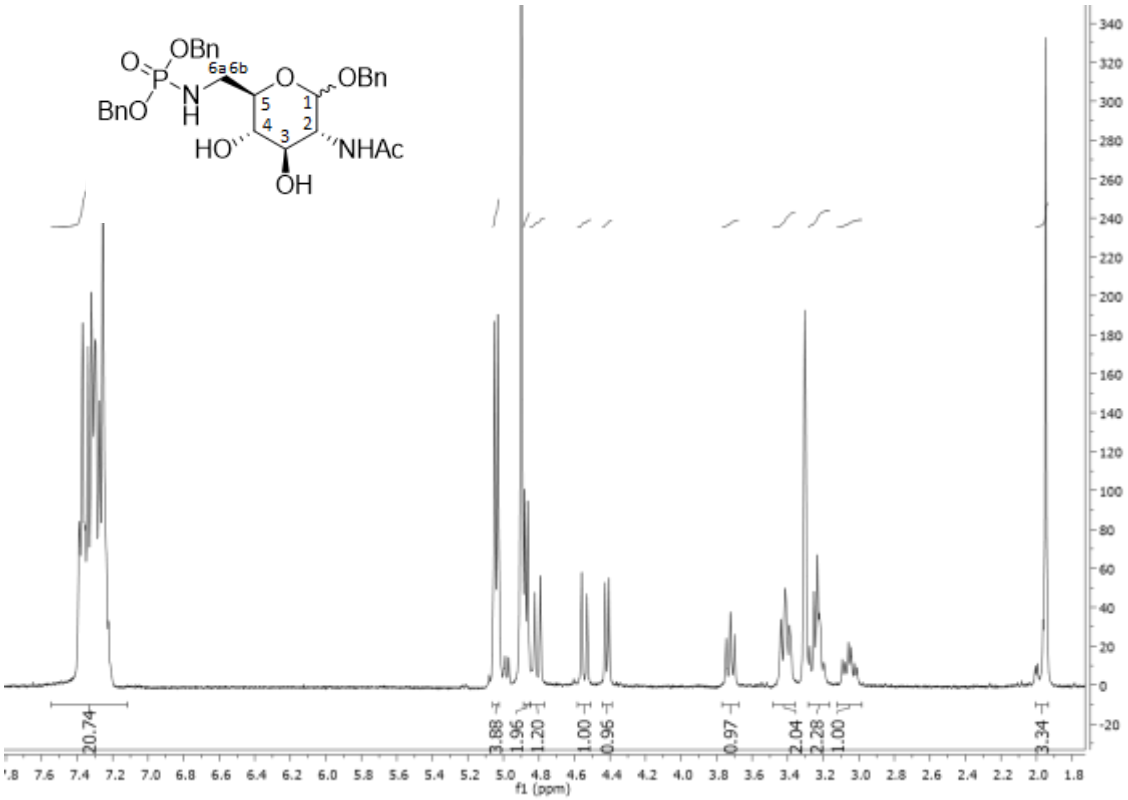


	g	MW	mmol	eq	d (g/mL)	V (mL)	[M]
4C	0,150	310,35	0,483	1			
Tetrabenzyl pirophosphate	0,312	538,34	0,58	1,2			
DIPEA	0.094	129,24	0,72	1,5	0,75	0,12	
DMF/THF						6	

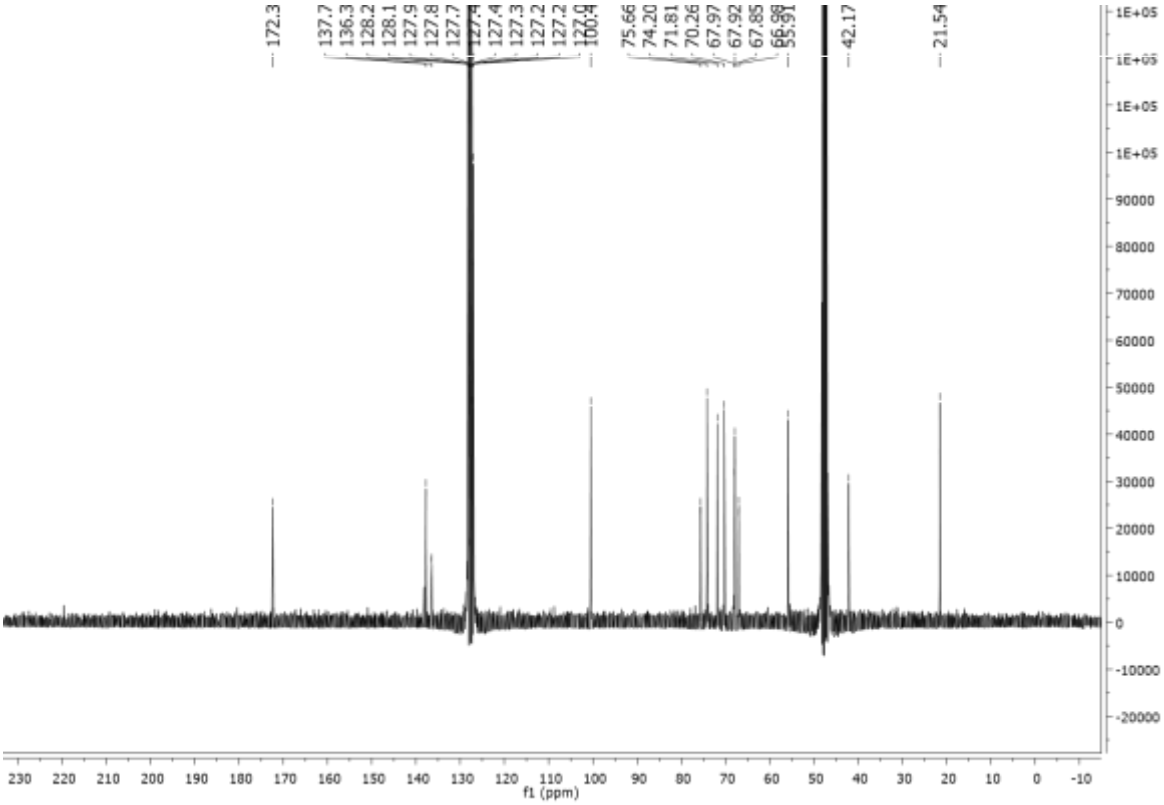
The product 4C was dissolved in DMF/THF under inert atmosphere; the solution was cooled to 0°and DIPEA and a solution of tetrabenzylpirophosphate in DMF was added drop wise. The reaction was left stirring for 6h at r.t. The solvent was evaporated the crude product was purified by flash chromatography AcOEt:MeOH 9:1, rf 0,33 (0,11 g – yield 42%), m/z_{calc} 480.17 (100.0%), 481.17 (23.8%), 482.17 (2.7%), 482.17 (1.6%), *m/z* $[M + Na]^+$ = 593.16.

New compound 5C have been characterized through ^1H and ^{13}C spectra.

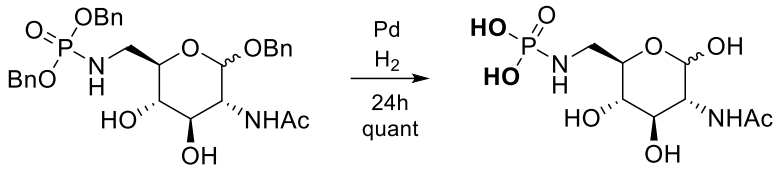




[1]H NMR (400 MHz, Methanol-d) δ 7.44 – 7.19 (m, 20H, **H**-Ar), 5.04 (bd, *J* = 7.6 Hz, 3H, 3POC**H**Ph), 4.87 (d, *J* = 6.4 Hz, 1H, POC**H**P), 4.81 (d, *J* = 12.2 Hz, 1H, C**H**Ph), 4.54 (d, *J* = 12.3 Hz, 1H, C**H**Ph), 4.42 (d, *J* = 8.4 Hz, 1H, **H-1**), 3.75 – 3.69 (m, 1H, **H-2**), 3.43-3.40 (m, 2H, **H-3**,**6a**), 3.28 – 3.19 (m, 2H, **H-4**,**5**), 3.05 (td, *J* = 13.5, 6.4 Hz, 1H, **H-6b**), 1.95 (s, 3H, **Ac**).



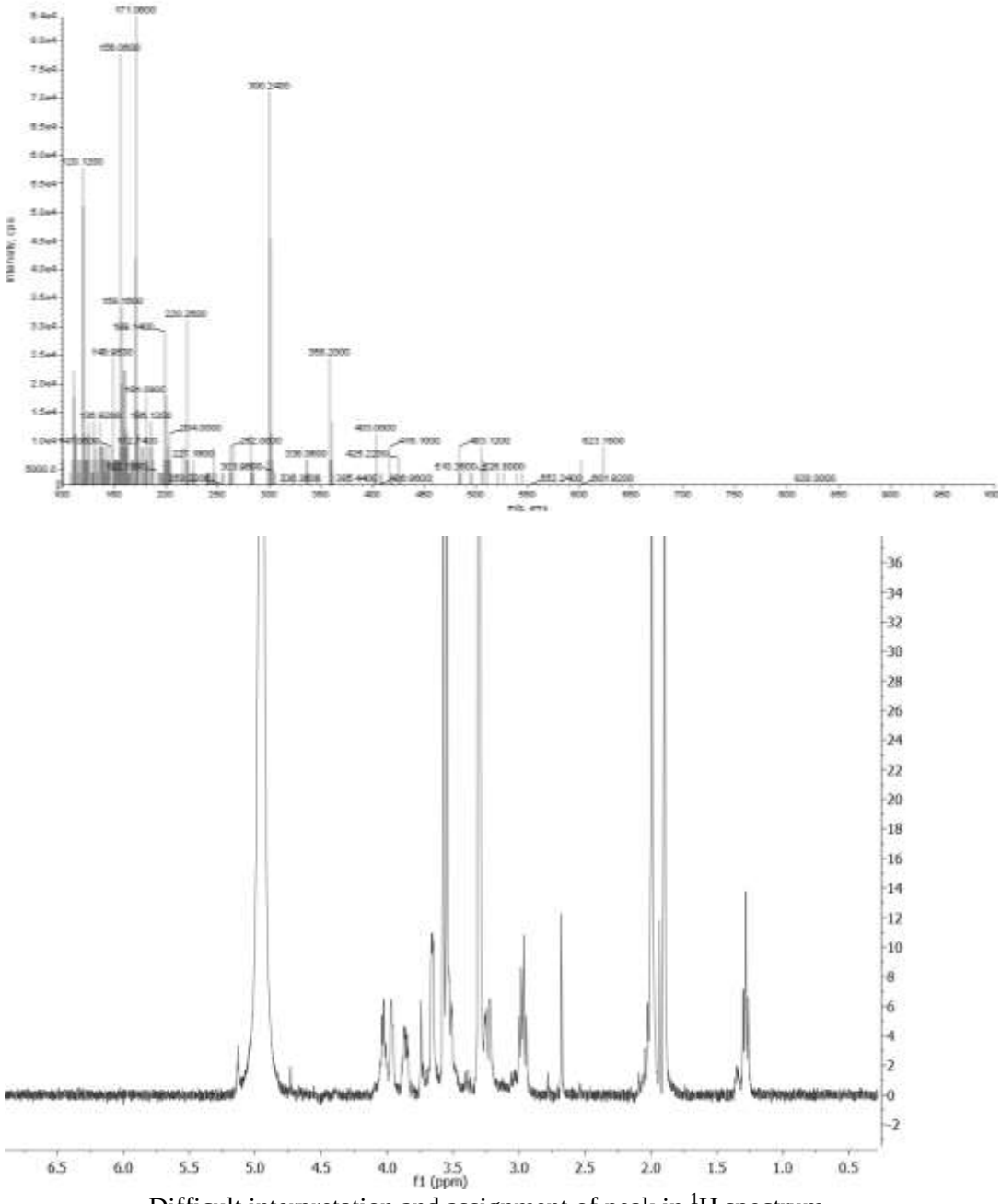
[13]C NMR (404 MHz, Methanol-d) δ 172.32 (**C**O), 137.72, 136.46, 136.38 (Cq**Ar**), 128.24-127.06 (CH**Ar**), 100.48 (**C1**), 75.66, 75.60, 74.20 (**C 3**, **4**, **5**), 71.81, 70.26 (CH2 **Bn**), 67.97, 67.92, 67.85, 67.03 (impurities), 66.98, 55.91 (C**2**), 42.17 (**C6)**, 21.54 (**Ac**).



5C 6C

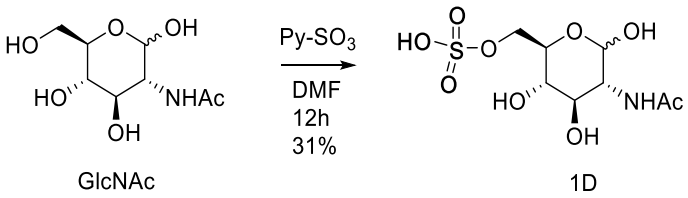
Compound 5C was dissolved in MeOH and the solvent was degassed. Then Pd/C was added and the reaction was stirred overnight under H_2 atmosphere.

The catalyst was removed by filtration through a celite filter, followed by filtration through Nylon 0.45 µm filter . After solvent evaporation compound 6C was obtained (0,045 g – quant), m/z_{calc} 300.07 (100.0%), 301.08 (8.7%), 302.08 (1.6%), *m/z* 300.07.



Difficult interpretation and assignment of peak in ^{1}H spectrum

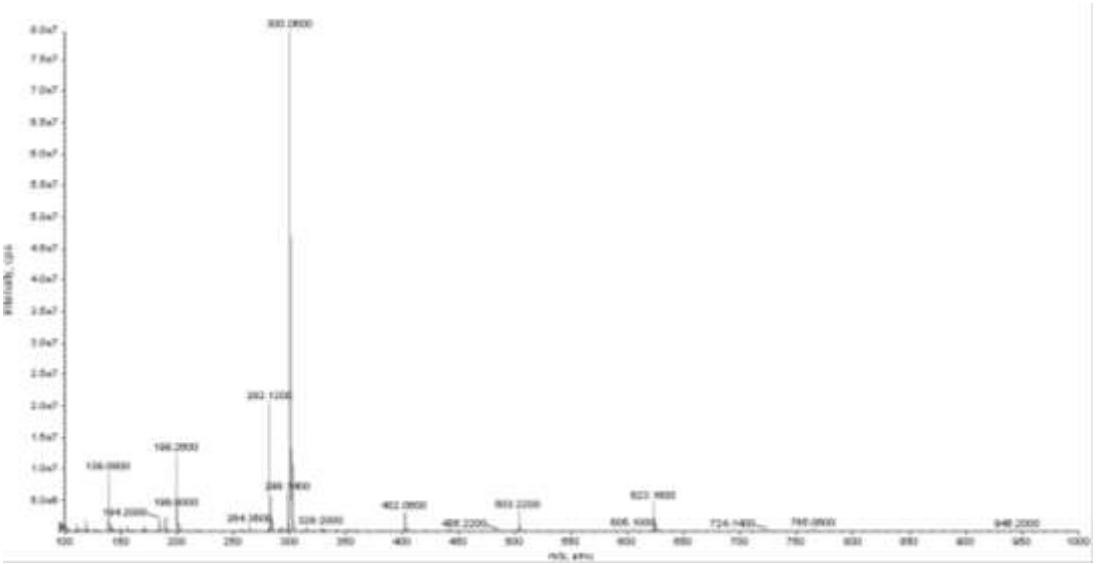
Scheme D - 6-sulfate derivative from *N*-acetylglucosamine

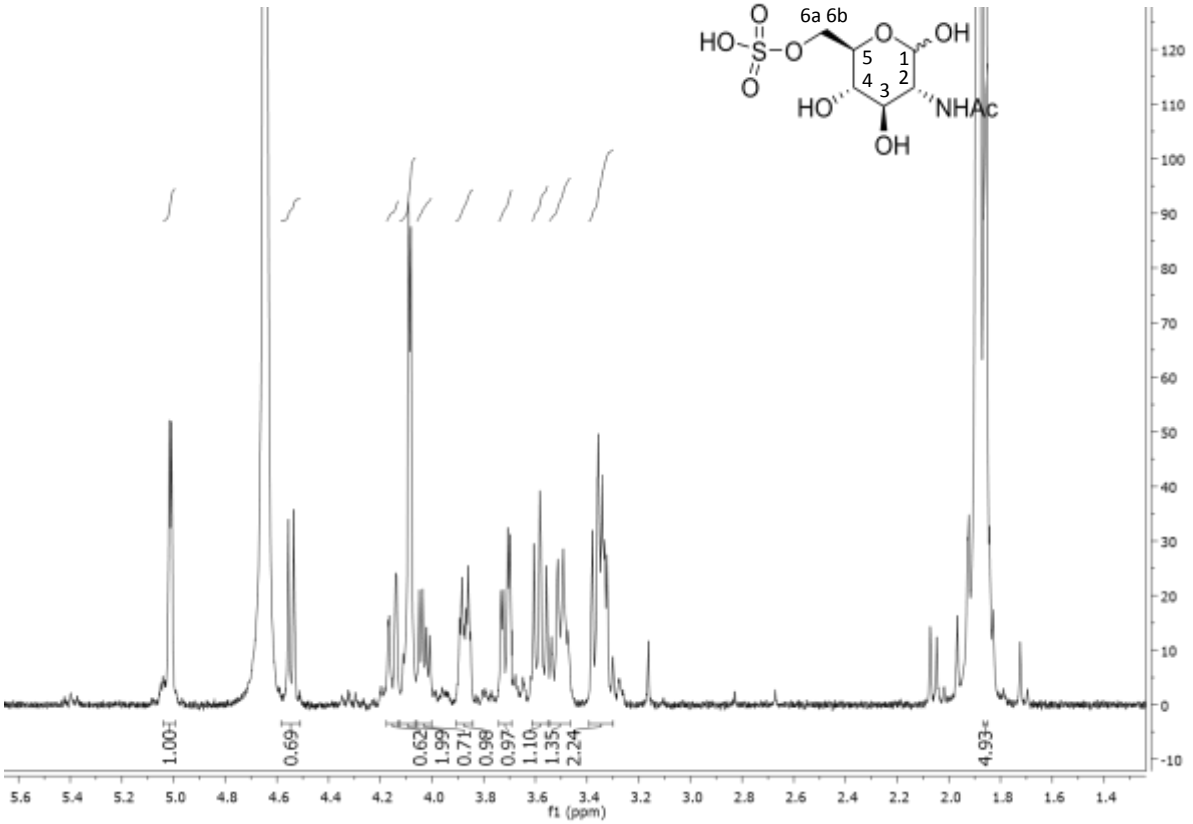


	g	MW	mmol	eq	d (g/mL)	V (mL)	[M]
GlcNAc	0,500	221	2,26	1			
Py-SO_3	0,349	159,16	2,19	1,5			
DMF				1		20	0,1

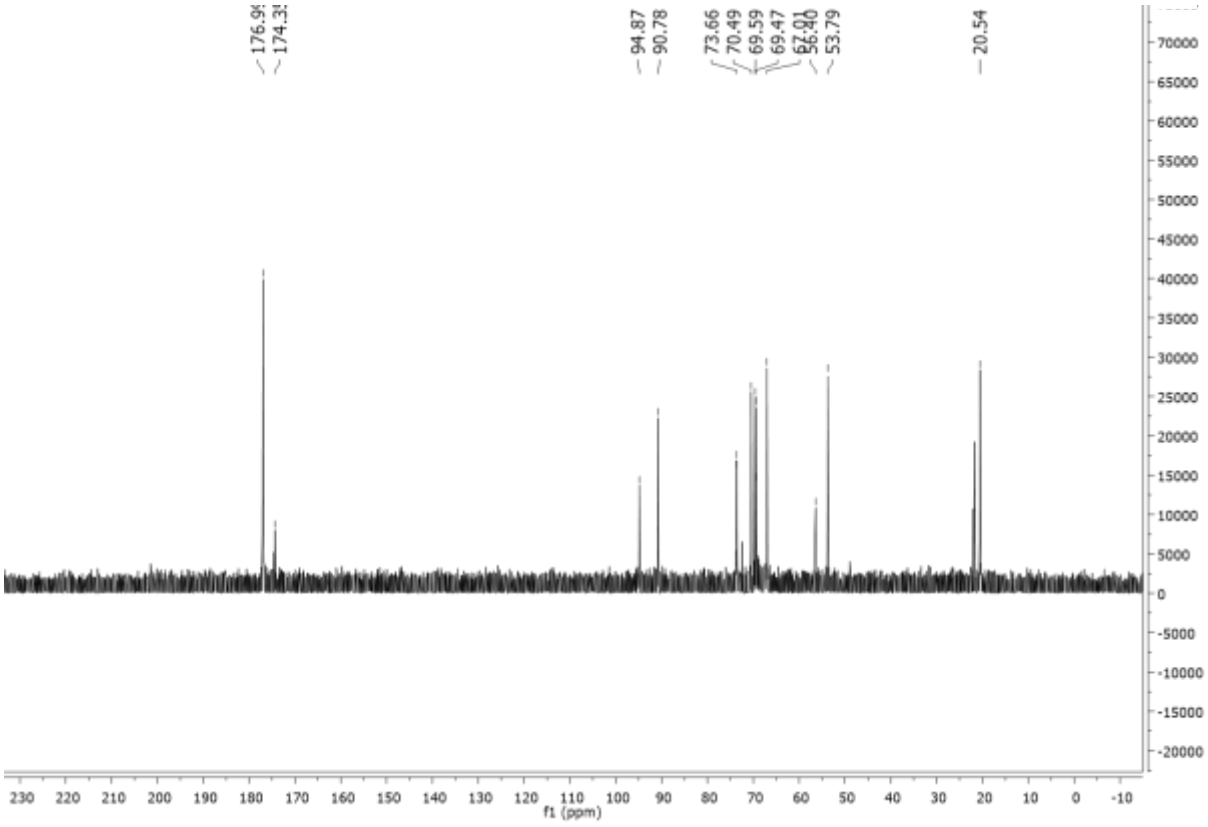
Pyridine – sulfur trioxide in DMF (0,1M) was added to a solution of *N*-acetylglucosamine dissolved in DMF over 3h at r.t. The mixture was stirred for another hour and concentrated in high vacuum at 40°C to remove DMF and pyridine. The pH was adjusted to 7.5 with KOH 2N and concentrated to remove pyridine. The mixture was purified with flash chromatography AcOEt:MeOH:H_2O 7:3:0,5:0,5 rf 0,32 (1,37 g – 31% yield), m/z_{calc} 301.05 (100.0%), 302.05 (8.7%), 303.04 (4.5%), 303.05 (1.8%), *m/z* $[M - H]^-$ = 300.06.

The characterization of compound 1D (^1H and ^{13}C spectra), ***final compound*** ready for *enzymatic test*, is in according to the literature. [34,35,36]

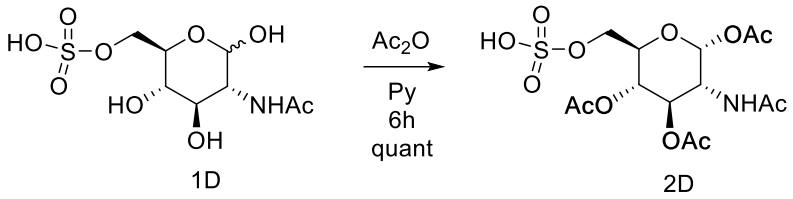




[1]H NMR (400 MHz, Deuterium Oxide) δ 5.01 (d, *J* = 3.3 Hz, 1H, **H-1α**), 4.54 (d, *J* = 8.4 Hz, 0.6H, **H-1β**), 4.15 (dd, *J* = 11.1, 1.6 Hz, 0.6H, **H-6aβ**), 4..01-4.08 (m, 2H, **H-6abα**), 4.03 (dd, *J* = 11.1, 5.2 Hz, 1H, **H-6bβ**), 3.87 (dt, *J* = 9.9, 3.5 Hz, 1H, **H-5α**), 3.72 (dd, *J* = 10.6, 3.4 Hz, 1H, **H-2α**), 3.58 (bt, *J* = 10.0, 1H, **H-3α**), 3.55 – 3.46 (m, 1.2H, **H-2β**, **5β**), 3.40 – 3.29 (m, 2.2H, **H-4α**, **3β**, **4β**), 1.86 (s, 4H, **Ac α and β**).



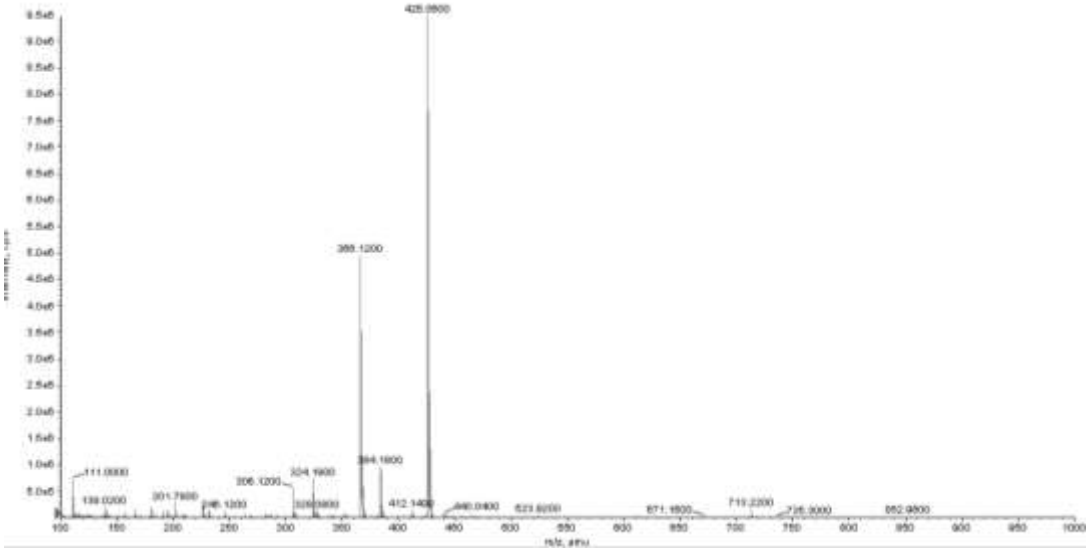
[13]C NMR (404 MHz, Deuterium Oxide) δ 174.35 (**C**O), 94.87 (**C1β**), 90.78(**C1α**), 73.66, 73.60, 70.49, 69.59, 69.47, 69.36 (**C3,4,5 αβ**), 67.01 (**C6αβ**), 56.40(**C2β**), 53.79(**C2α**), 20.54 (**Ac**).

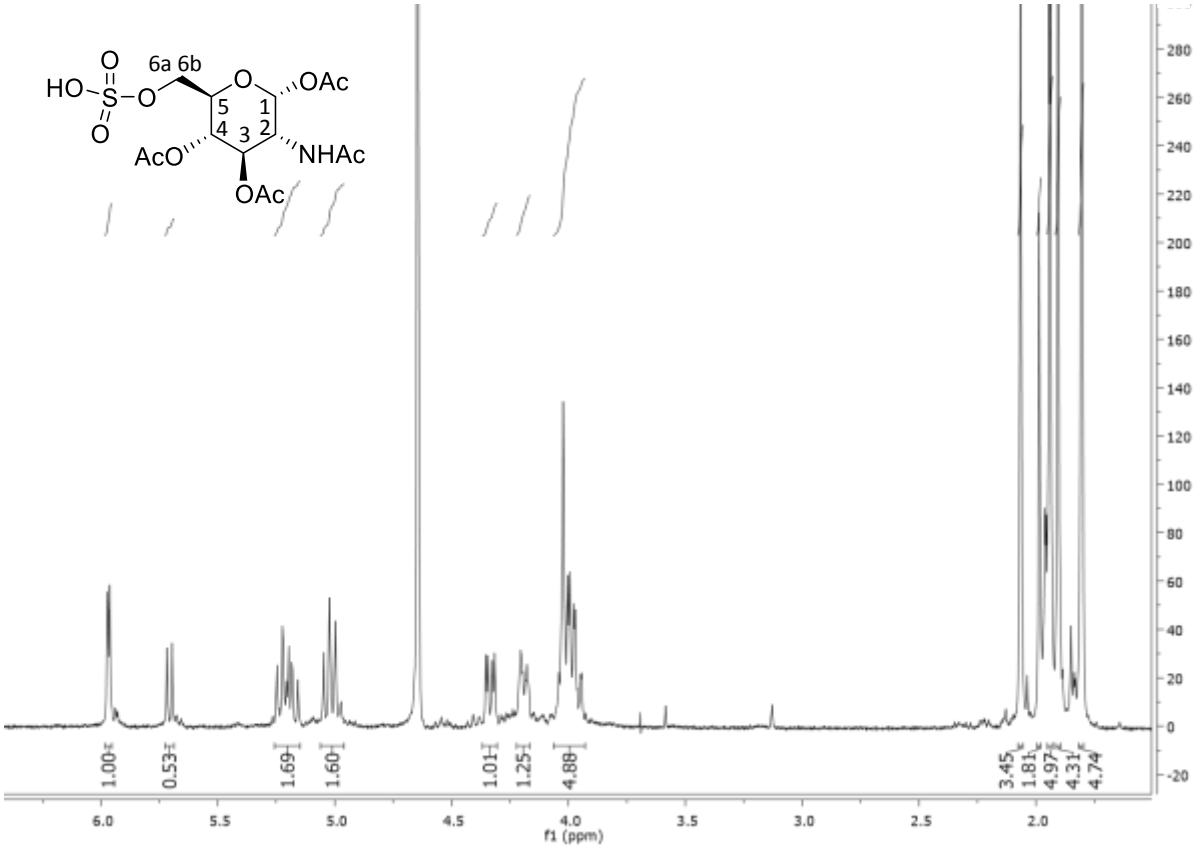


	g	MW	mmol	eq	d (g/mL)	V (mL)	[M]
1D	0,500	301,27	1,66	1			
Ac_2O	1,71	102,9	16,60	10	1,84	1,08	
Py	4,59	79,1	58,1	35	0,982	4,60	

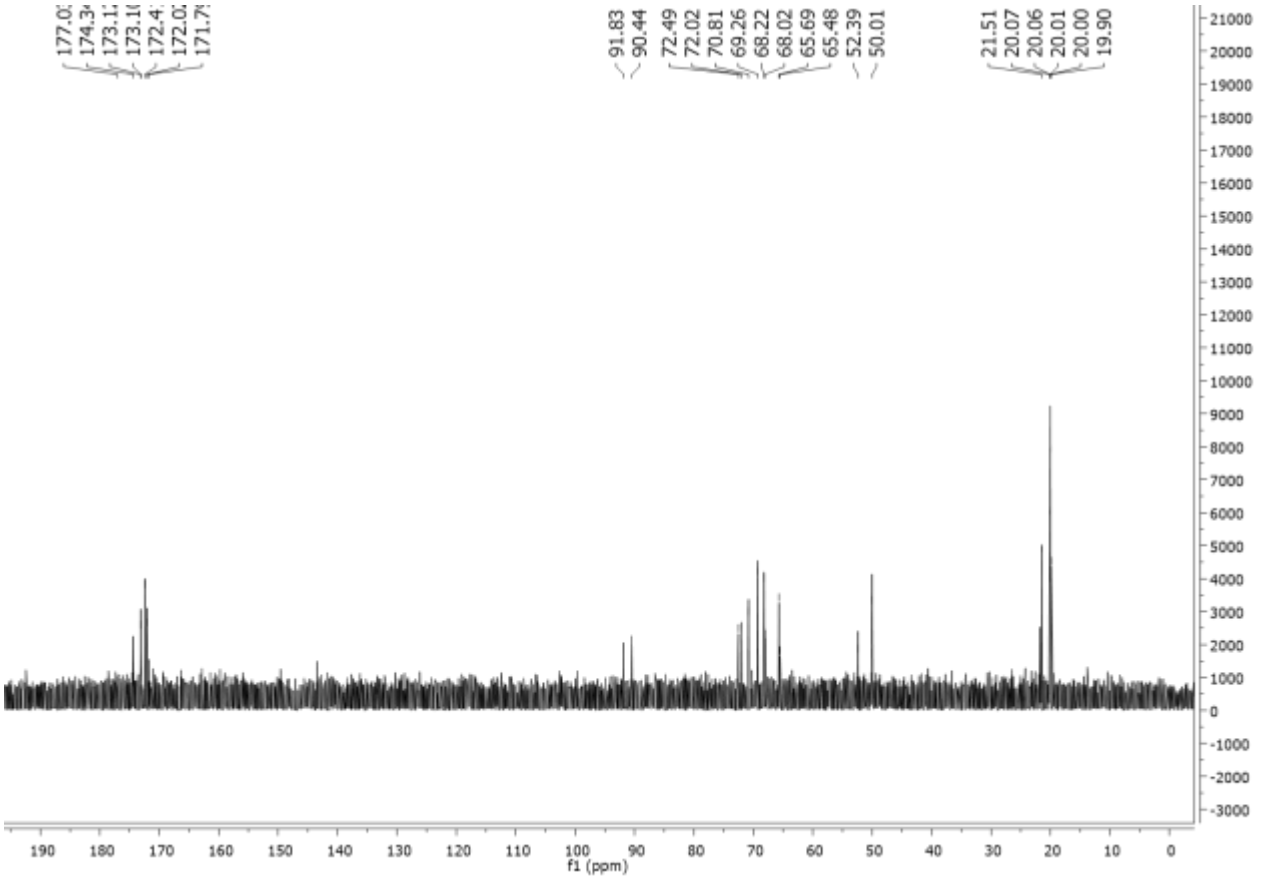
The sulfate 1D was suspended in Py, and the reaction cooled to 0°C, then acetic anhydride was added. The mixture was stirred at room temp. for 4 h. The reaction was quenched by the addition of MeOH (3 mL) and the solvent was removed. The crude was purified by flash chromatography AcOEt:MeOH 8:2, rf 0,29 (0,67 g – quant), m/z_{calc} 426.07 (100.0%), 427.07 (15.1%), 428.07 (4.5%), 428.08 (2.5%), 428.08 (1.1%) *m/z* $[M - H]^-$ = 426.06.

New compound 2D have been characterized through ^1H and ^{13}C spectra.



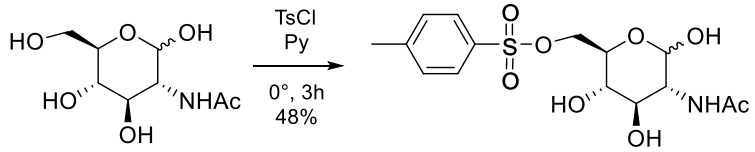


[1]H NMR (400 MHz, Deuterium Oxide) δ 5.97 (d, *J* = 3.6 Hz, 1H, **H-1α**), 5.71 (d, *J* = 8.8 Hz, 0.6H, **H-1β**), 5.27 – 5.14 (m, 1.6H, **H-3αβ**), 5.06 – 4.97 (m, 1.6H, **H-4αβ**), 4.33 (dd, *J* = 11.0, 3.7 Hz, 1H, **H-2α),** 4.22 – 4.16 (m, 1H, **H-5α**), 4.06 – 3.93 (m, 4.4H, **H-2β**, **5β**, **6abαβ**), 2.07 (s, **Ac**), 1.99 (s, **Ac**), 1.94 (s, **Ac),** 1.91 (s, **Ac**), 1.81 (s, **Ac**).



[13]C NMR (101 MHz, Deuterium Oxide) δ 174.34, 173.12, 173.10, 172.41, 172.02, 171.79 (**CO**), 91.83 (**C1β**), 90.44(**C1α**), 72.49, 72.02, 70.81, 69.26, 68.22, 68.02 (**C 3,4,5 αβ**), 65.69, 65.48 (**C6αβ**), 52.39(**C2β**), 50.01 (**C2α**), 21.51, 20.07, 20.06, 20.01, 20.00, 19.90 (**Ac**).

Scheme E - 6-sulfonate derivative from *N*-acetylglucosamine

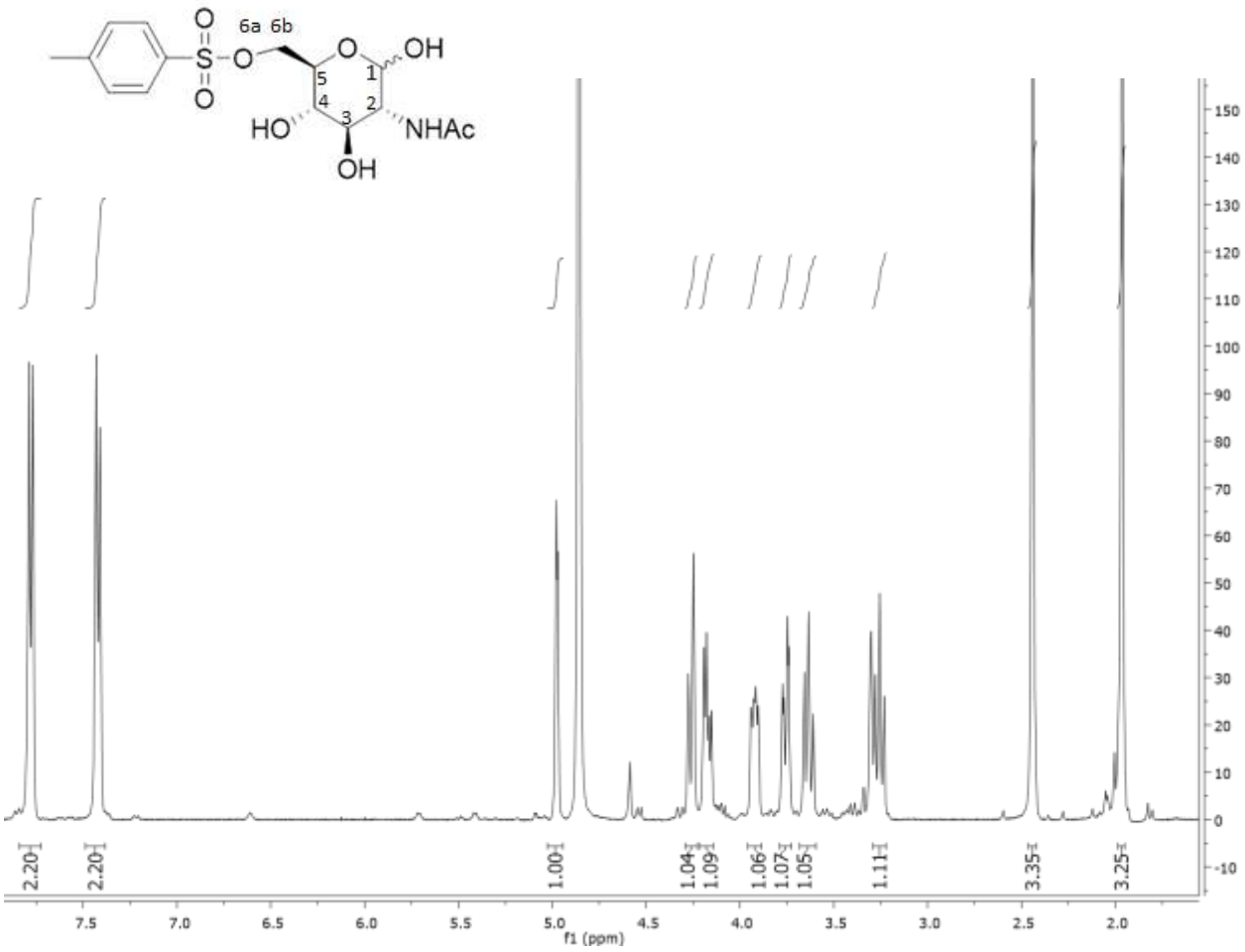


GlcNAc 1E

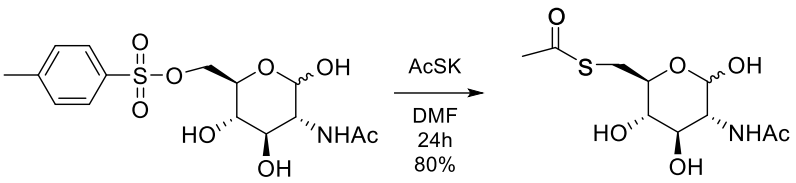
	g	MW	mmol	eq	d (g/mL)	V (mL)	[M]
GlcNAc	5,00	221,26	22,62	1			
Py		79,1			0,982	40	
TsCl in Py	5,17	180,60	27,15	1,2		30	

GlcNAc was dissolved in anhydrous pyridine and cooled to 0°C, then a solution of p-toluene-sulfonyl chloride in anhydrous pyridine was added dropwise . The reaction was stirred approx. 0°C for 2.5 h. then the reaction was quenched by the addition of methanol, toluene was added and the mixture was concentrated under reduce pressure. The residue was purified by column chromatography on silica gel AcOEt:MeOH 9:1, rf 0,31 (3,63 g – 46% yield), m/z_{calc} 375.10 (100.0%), 376.10 (16.2%), 377.09 (4.5%), 377.10 (1.6%), 377.11 (1.2%).[42]

The characterization of compound 1E (^{1}H spectrum reported – ^{13}C, 2D and *m/z* spectra not reported) is in according to the literature.[42]



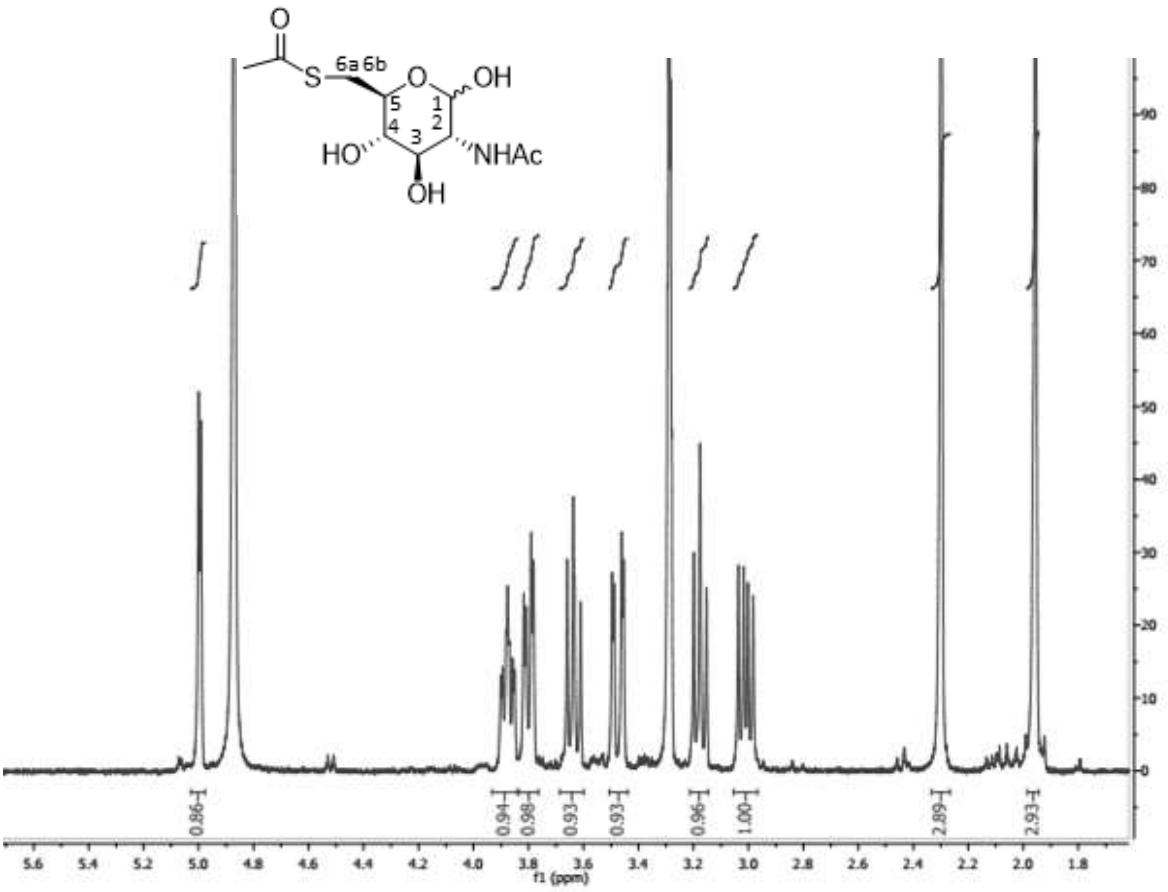
^{1}H NMR (400 MHz, Methanol-d) δ 7.78 (d, *J* = 8.2 Hz, 2H, **Ar**), 7.42 (d, *J* = 7.9 Hz, 2H, **Ar**), 4.97 (d, *J* = 3.1 Hz, 1H, **H-1**), 4.26 (bd, *J* = 10.5 Hz, 1H, **H-6a**), 4.17 (dd, *J* = 10.6, 5.5 Hz, 1H, **H-6b**), 3.96 – 3.89 (m, 1H, **H-5**), 3.76 (dd, *J* = 10.6, 3.2 Hz, 1H, **H-2),** 3.63 (t, *J* = 9.7 Hz, 1H, **H-3**), 3.25 (t, *J* = 9.4 Hz, 1H, **H-4**), 2.44 (s, 3H, **MeTs**), 1.96 (s, 3H, **Ac**).



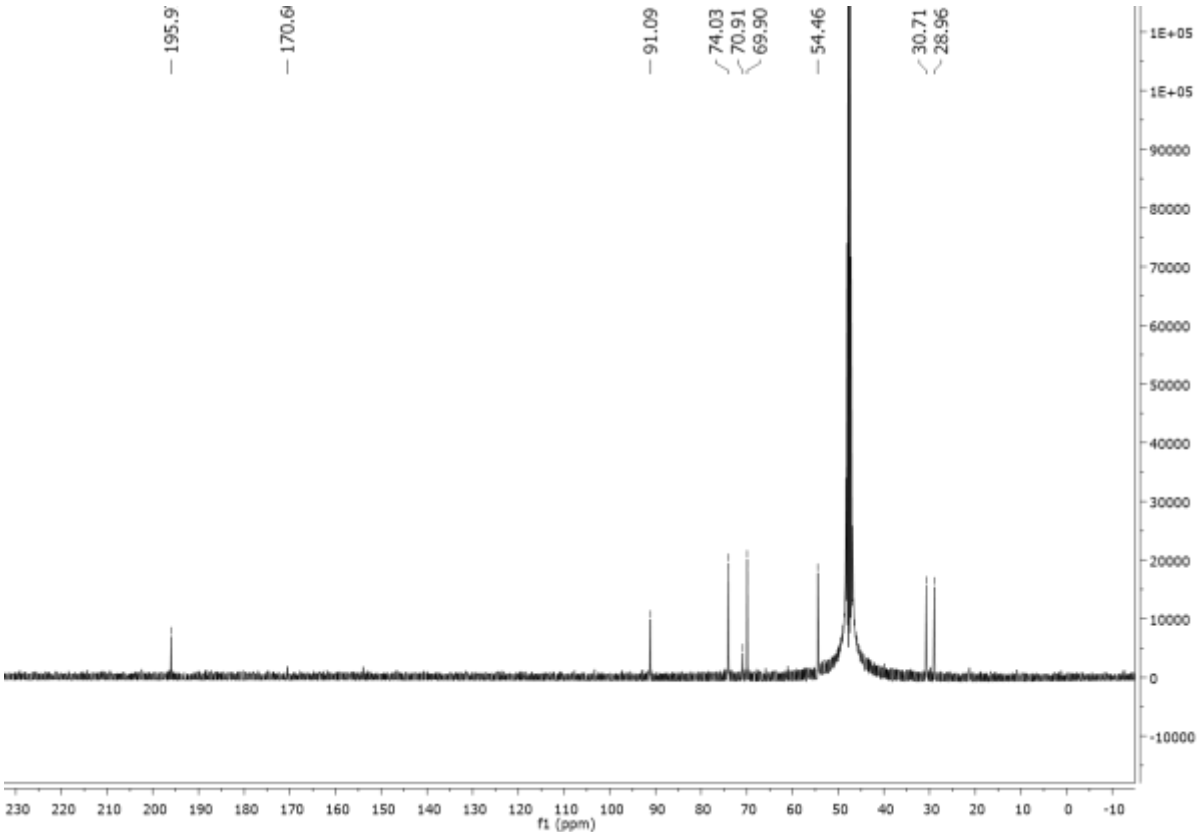
					g	MW	mmol	eq	d (g/mL)	V (mL)	[M]
1E					0,800	334,34	2,39	1			
AcSK	0,409	114,21	3,59	1,5							
DMF						20			0,1		

To a solution of 1E in DMF, AcSK was added and the reaction was stirred for 24h. Then the solvent was evaporated in vacuum pump. The residue was purified by column chromatography on silica gel AcOEt:MeOH 9:1, rf 0,31 (0,28 – 43% yield), m/z_{calc} 279.08 (100.0%), 280.08 (10.8%), 281.07 (4.5%), 281.08 (1.2%).

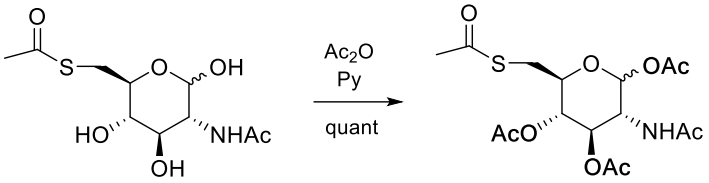
New compound 2E has been characterized through ^{1}H and ^{13}C spectra.



^{1}H NMR (400 MHz, Methanol-d) δ 5.00 (d, *J* = 3.5 Hz, 1H, **H-1**), 3.91 – 3.84 (m, 1H, **H-5**), 3.80 (dd, *J* = 10.7, 3.5 Hz, 1H, **H-2**), 3.63 (dd, *J* = 10.6, 8.9 Hz, 1H, **H-3**), 3.47 (dd, *J* = 13.8, 2.8 Hz, 1H, **H-6a**), 3.18 (t, *J* = 9.2 Hz, 1H, **H-4**), 3.01 (dd, *J* = 13.8, 7.6 Hz, 1H, **H-6b**), 2.30 (s, 3H, **SAc**), 1.96 (s, 3H, **Ac**).



^{13}C NMR (101 MHz, Methanol-d) δ 195.97(**SC**O), 170.60 (**C**O), 91.09 (**C1**), 74.03, 70.91, 69.90 (**C3,4,5**), 54.46 (**C2**), 30.71 (**C6**), 28.96 (**Ac**).

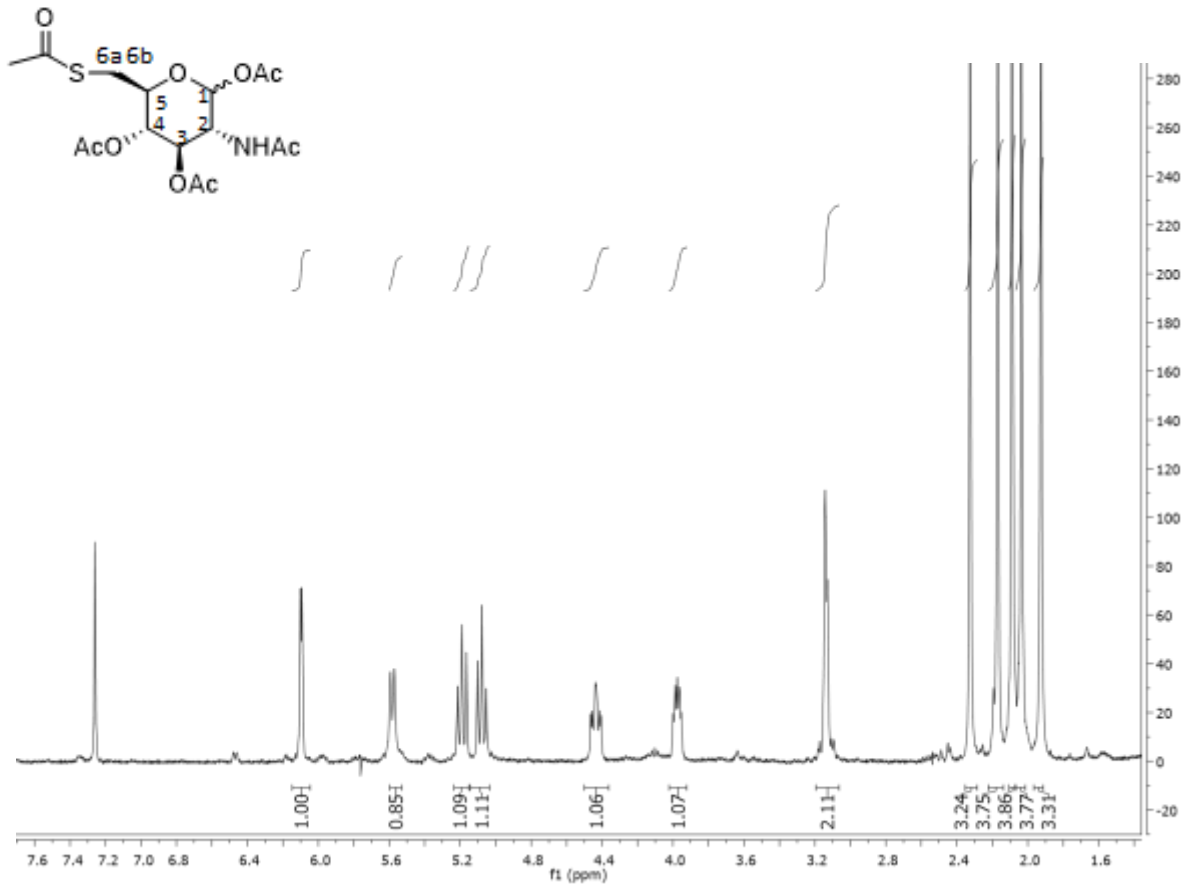


2E 3E

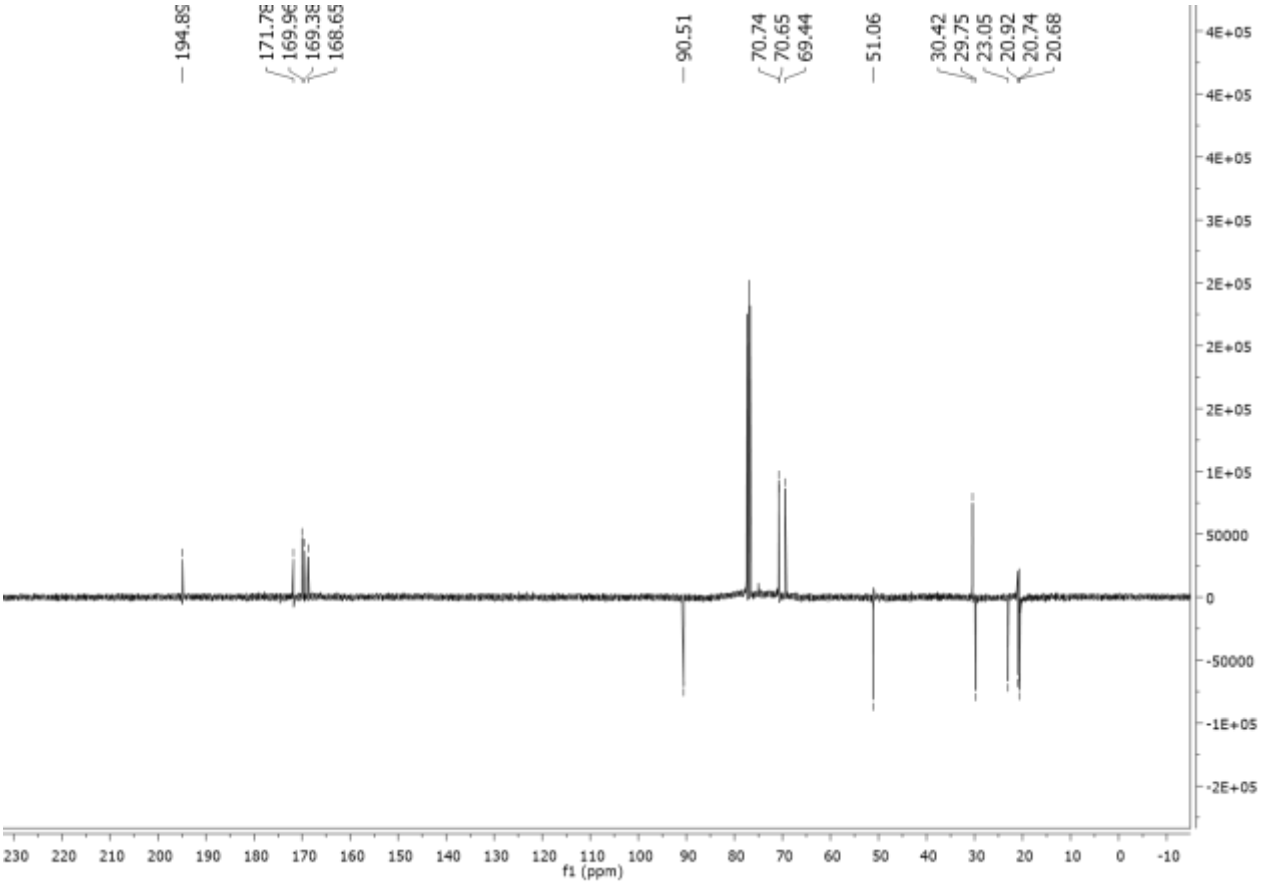
	g	MW	mmol	eq	d (g/mL)	V (mL)	[M]
3E	0,150	279,31	0,54	1			
Ac_2O		102,9		10		1,5	
Py		79,1		35		3,12	

The sulfate 2E was suspended in Py and cooled to 0°C, then acetic anhydride was added. The mixture was stirred at room temp. for 4 h. The reaction was quenched by the addition of MeOH (3 mL) and then concentrated under reduced pressure. The crude was purified by flash chromatography Et Pet:AcOEt 2:8, rf 0,22 (0,20g – quant), m/z_{calc} 405.11 (100.0%), 406.11 (17.3%), 407.11 (4.5%), 407.11 (1.8%), 407.12 (1.4%).

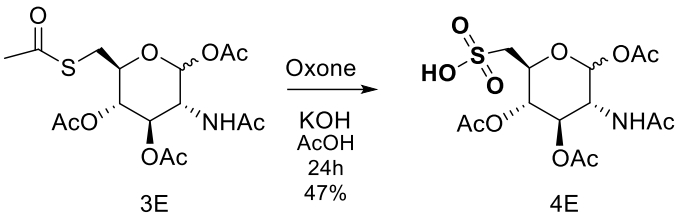
New compound 3E has been characterized through ^{1}H and ^{13}C spectra.



^{1}H NMR (400 MHz, Chloroform-d) δ 6.10 (d, *J* = 3.4 Hz, 1H, **H-1**), 5.58 (d, *J* = 8.6 Hz, 1H, ***N*H**), 5.19 (dd, *J* = 9.8, 9.6 Hz, 1H, **H-3**), 5.08 (t, *J* = 9.6 Hz, 1H, H-4), 4.48 – 4.40 (m, 1H**, H-2**), 4.02 – 3.94 (m, 1H, **H-5**), 3.20 – 3.09 (m, 2H, **H-6a**,**6b**), 2.32 (s, 3H, **Ac**), 2.17 (s, 3H, **Ac)**, 2.09 (s, 3H**, Ac**), 2.04 (s, 3H, **Ac**), 1.93 (s, 3H, **Ac**).



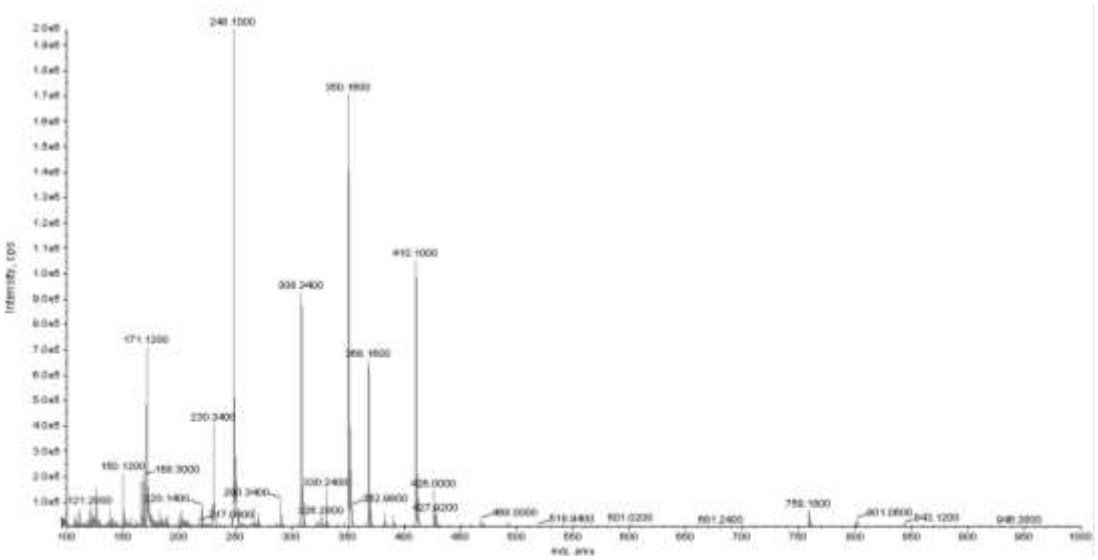
^{13}C NMR (400 MHz, Chloroform-d) δ 194.89 (**SC**O), 171.78, 169.96, 169.38, 168.65 (**C**O), 90.51 (**C1**), 70.74, 70.65, 69.44 (**C3,4,5**), 51.06 (**C2**), 30.42 (**C6**), 29.75 (**AcS**), 23.05, 20.92, 20.74, 20.68 (**Ac**).

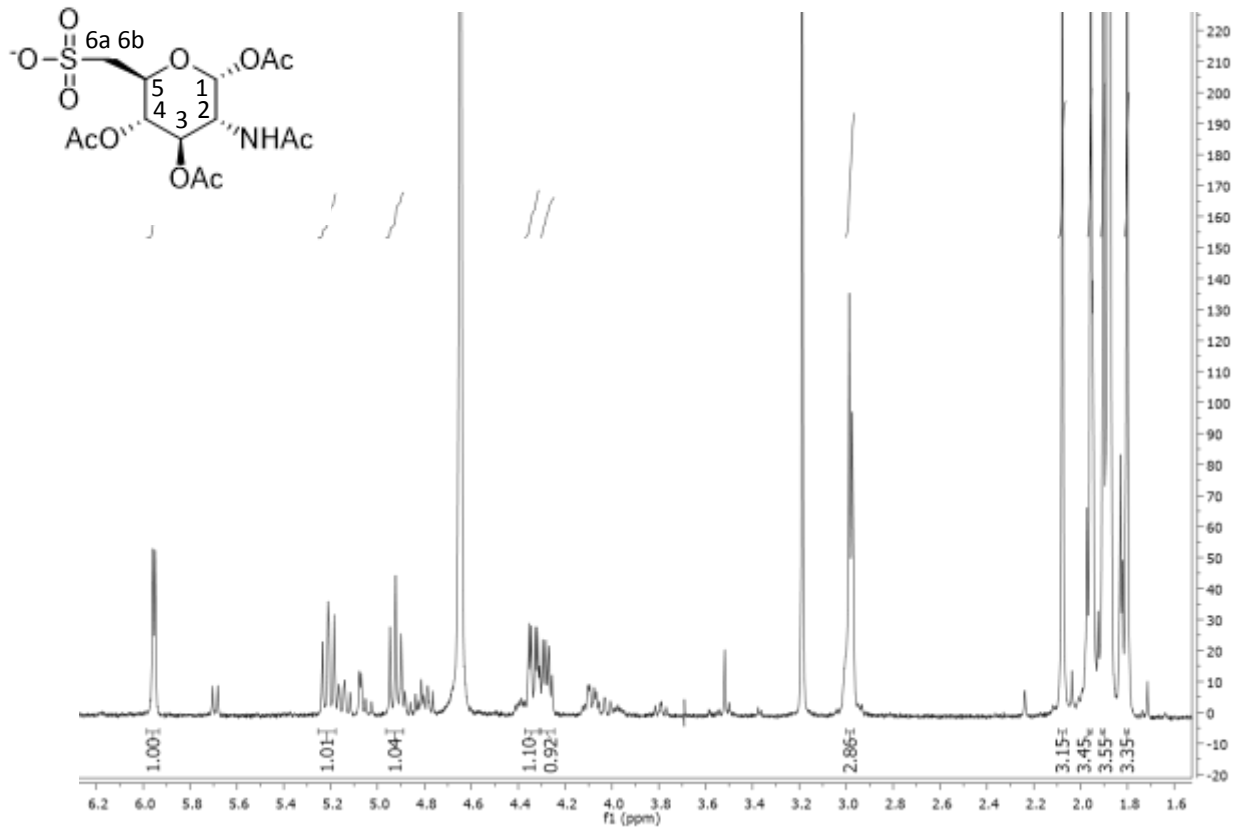


	g	MW	mmol	eq	d (g/mL)	V (mL)	[M]
3E	0,781	405,11	1,93	1			
Oxone	4,55	614,76	7,40	3,84			
KOAc	3,48	98,14		18,41			
AcOH						50	

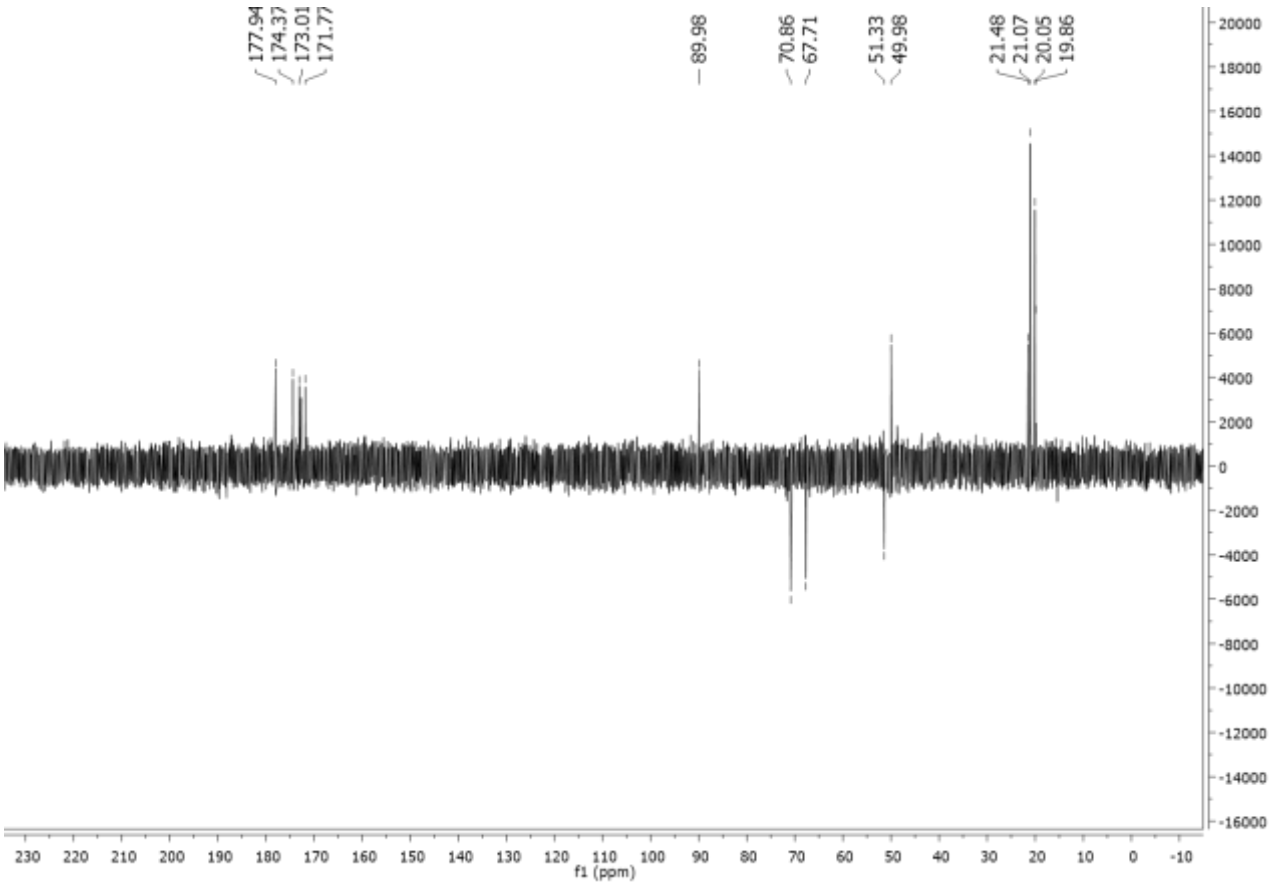
The thioacetate 3E was dissolved in AcOH, then KOAc was added followed by Oxone, and the reaction was left stirring for 24h. Then the reaction mixture was extracted with AcOEt, and the organic layer was washed with H2O. The crude product was purified by flash chromatography AcOEt:MeOH 7:3, rf 0,32 (0,37 g – 47% yield), m/z_{calc} 410.08 (100.0%), 411.08 (15.1%), 412.07 (4.5%), 412.08 (2.3%), 412.08 (1.1%) *m/z* $[M - H]^-$ = 410.10.

*The **final** and **new compound*** 4E, for in vivo test, has been characterized through ^{1}H and ^{13}C spectra.

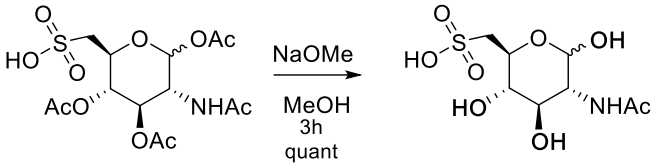




^{1}H NMR (400 MHz, Deuterium Oxide) δ 5.95 (d, *J* = 3.8 Hz, 1H, **H-1**), 5.21 (dd, *J* = 10.8, 9.4 Hz, 1H, **H-3**), 4.92 (t, *J* = 9.7 Hz, 1H, **H-4**), 4.34 (dd, *J* = 10.8, 3.7 Hz, 1H, **H-2**), 4.31 – 4.25 (m, 1H, **H-5**), 2.99-2.97 (m, 2H, **H-6a,6b**), 2.08 (s, 3H, **Ac**), 1.96 (s, 3H, **Ac**), 1.90 (s, 3H, **Ac**), 1.80 (s, 3H, **Ac**).



^{13}C NMR (101 MHz, Deuterium Oxide) δ 177.94, 174.37, 173.01, 171.77, 89.98 (**C1**), 70.86, 67.71 (**C-3,4,5**), 51.33 (**C6**), 49.98 (**C2**), 21.48, 21.07, 20.05, 19.86 (**Ac**).

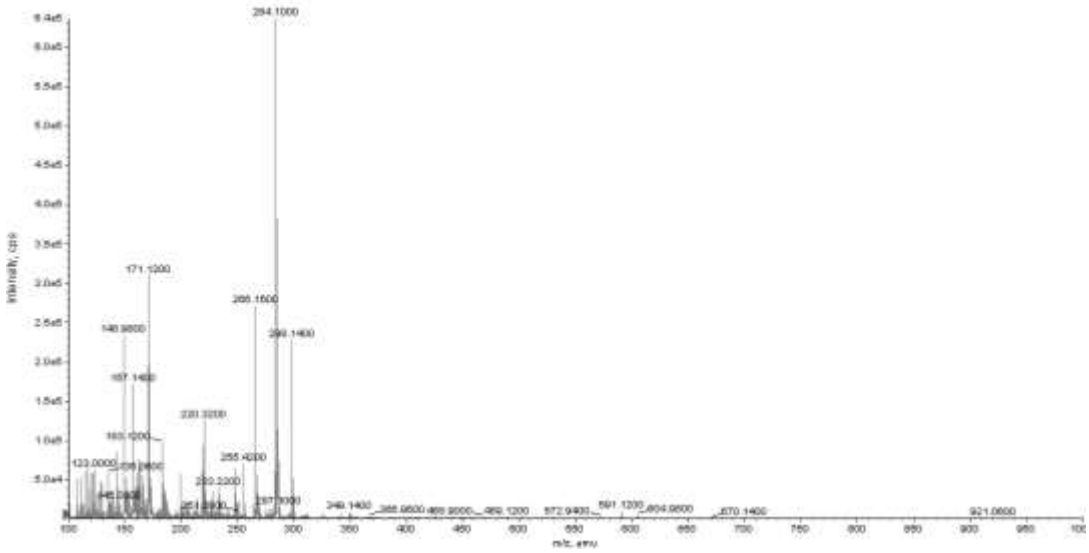


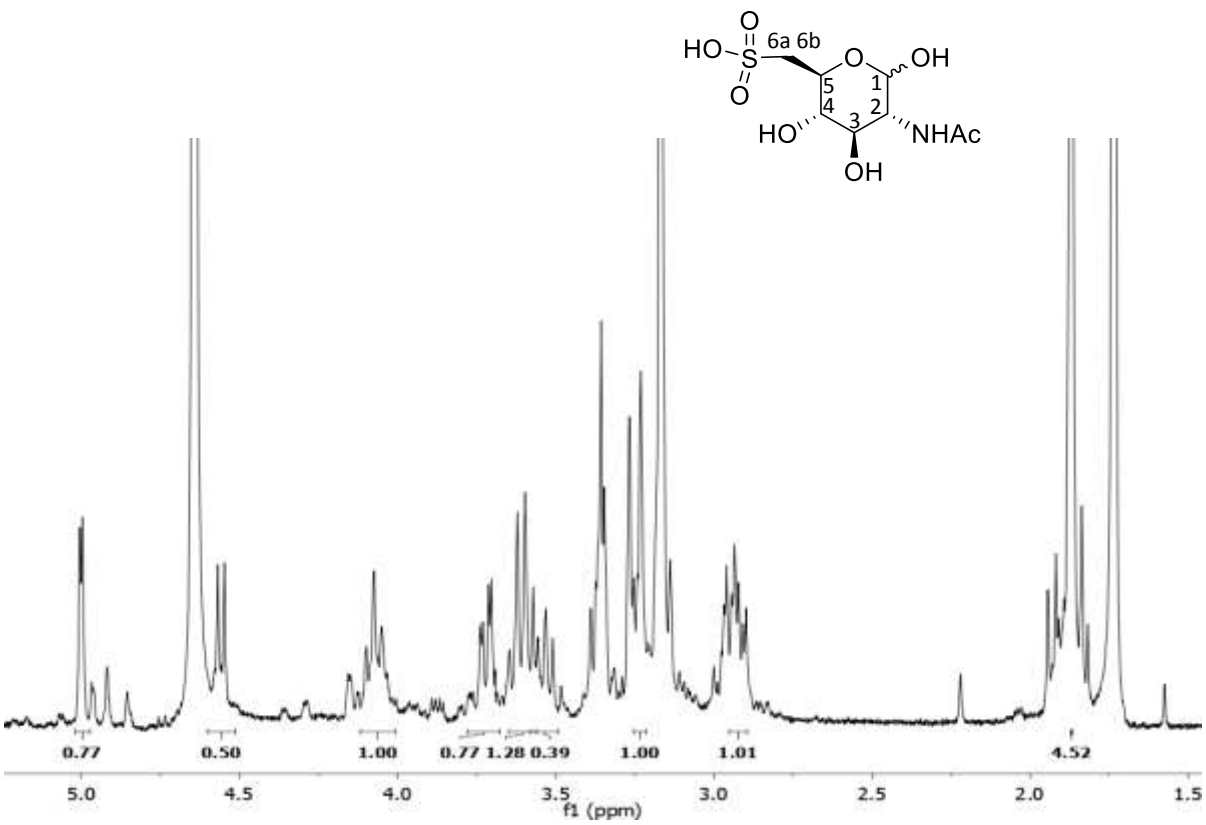
	g	MW	mmol	eq	d (g/mL)	V (mL)	[M]
5E	0,140	411,08	0,341	1			
NaOMe	0,110	54,02	2,04	6			
MeOH						7	0,1

NaOMe was added to compound 4E dissolved in dry MeOH under Argon. After 3h the stirring was stopped and the reaction treated with Amberlite 5% HCl to pH 6. The solvent was evaporated.

The compound was purification RP-C18, gradient H_2O:MeOH (0,09 g – quant), m/z_{calc} 284.04 (100.0%), 285.05 (8.7%), 286.04 (4.5%), 286.05 (1.6%) *m/z* $[M - H]^-$ =284.10.

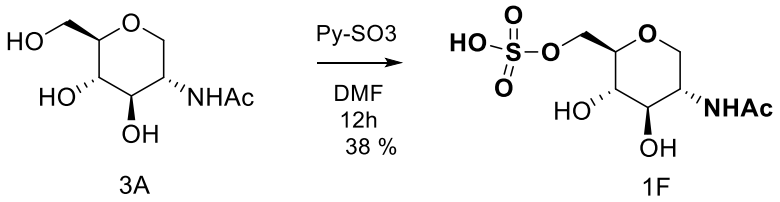
*The **final** and **new compound*** 5E, for in in vitro test, has been characterized through ^1H (difficult interpretation) and *m/z* spectrum.





^{1}H NMR (400 MHz, Deuterium Oxide): difficult interpretation

Scheme F - 6-sulfate derivative from 1-deoxy *N*-acetylglucosamine

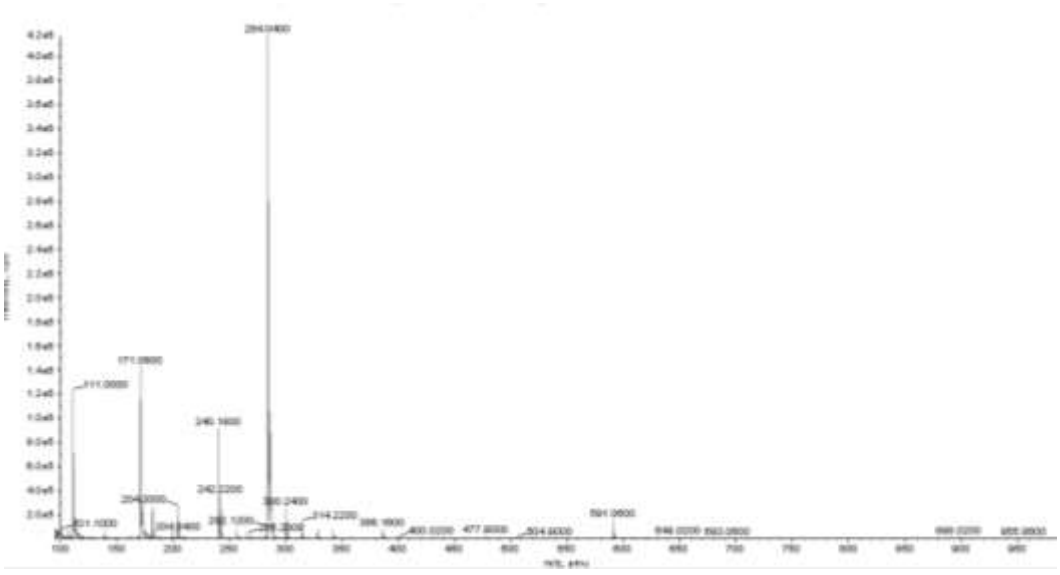


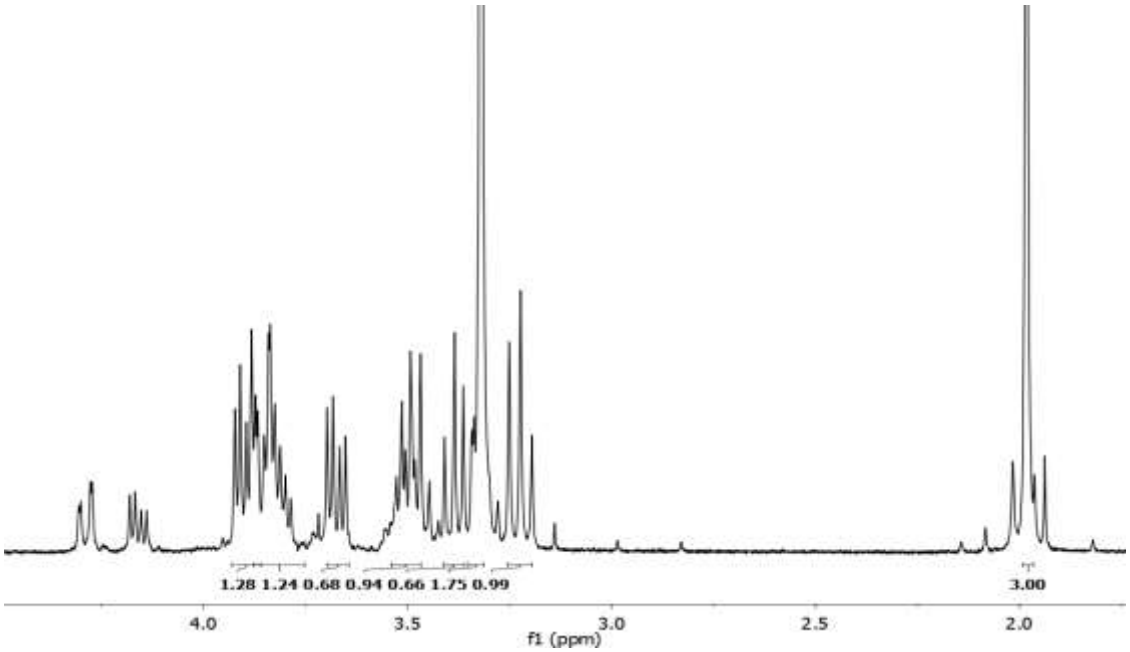
	g	MW	mmol	eq	d (g/mL)	V (mL)	[M]
3A	0,800	204	3,92	1			
Py-SO_3	0,936	159,16	5,88	1,5			
DMF				1		35	0,1

Pyridine – sulfur trioxide in DMF was added to a solution of compound 3A dissolved in DMF over 3h at r.t. The mixture was stirred for another hour and concentrated in vacuum at 40°C to remove DMF and pyridine. The pH was adjusted to 7.5 with KOH 2N and concentrated to remove pyridine.

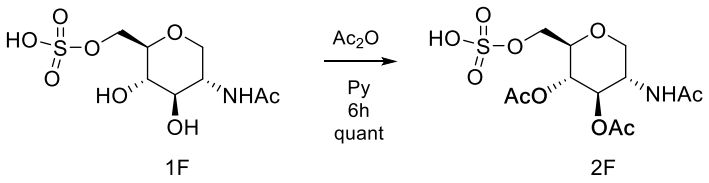
The mixture was purified with flash chromatography AcOEt:MeOH:AcOH:H_2O 7:3:0,5:0,5 rf 0,34 (0,42 g – 38% yield), m/z_{calc} 285.05 (100.0%), 286.06 (8.7%), 287.05 (4.5%), 287.06 (1.6%), *m/z* $[M - H]^-$ = 284.04.

*The **final** and **new compound*** 1F, for in in vitro test, has been characterized through ^1H and *m/z* spectra.





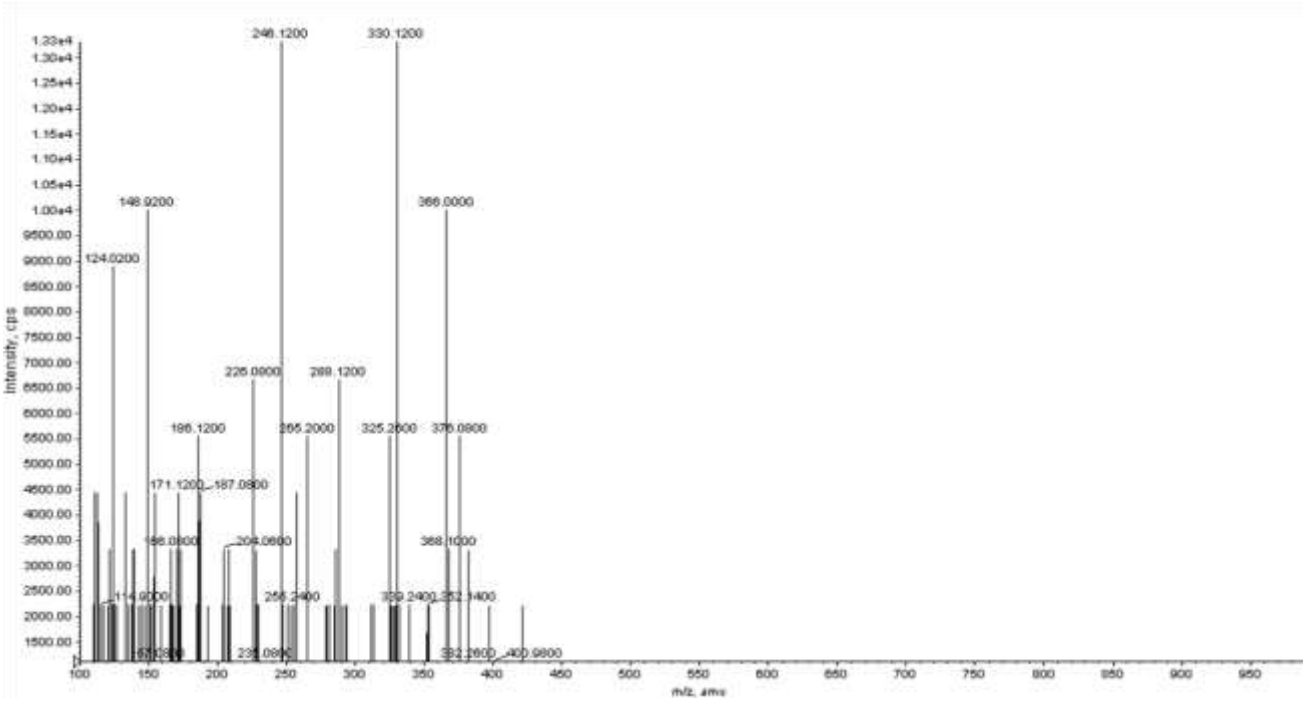
^{1}H NMR (400 MHz, Deuterium Oxide): difficult interpretation

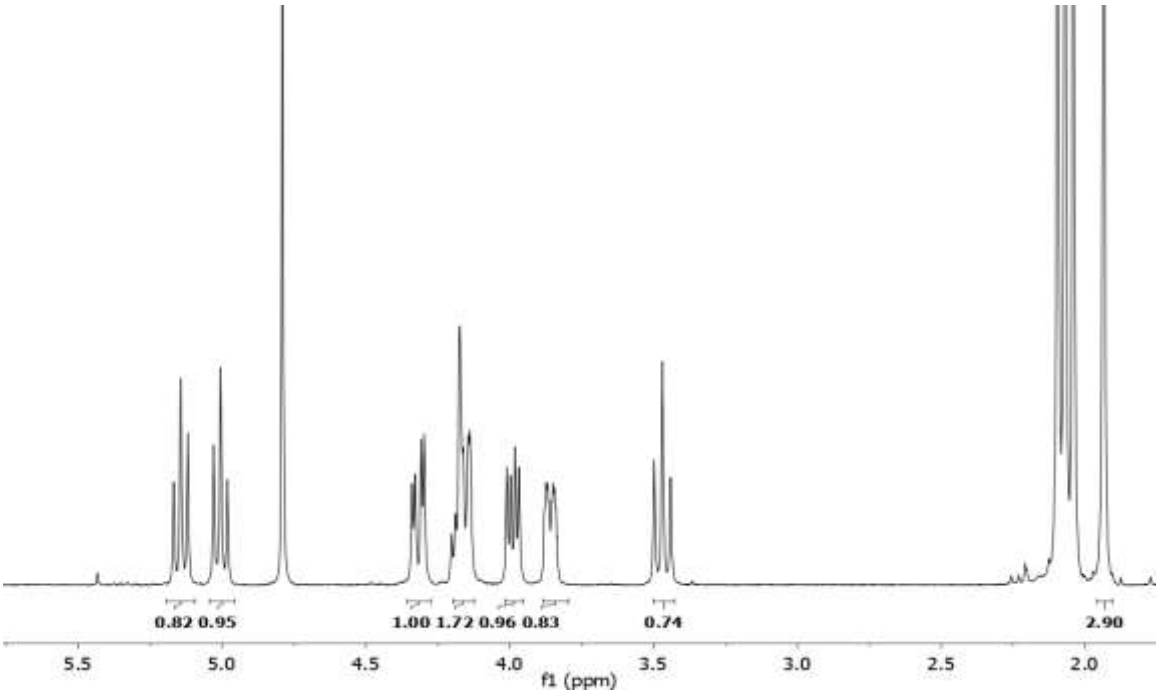


	g	MW	mmol	eq	d (g/mL)	V (mL)	[M]
1F	0,610	284,05	2,03	1			
Ac_2O	2,08	102,9	20,3	10	1,84	1,13	
Py	5,62	79,1	58,1	35	0,982	5,72	

The sulfate 1F was suspended in Py (35 eq) and acetic anhydride was added in ice bath. The mixture was stirred at room temp. for 4 h. The reaction was quenched by the addition of MeOH (3 mL) and the mixture was purified with flash chromatography Et Pet:AcOEt 2:8, rf 0,22 (0,75g – quant), m/z_{calc} 369.07 (100.0%), 370.08 (13.0%), 371.07 (4.5%), 371.08 (2.1%) *m/z* $[M- H]^-$ = 368.10.

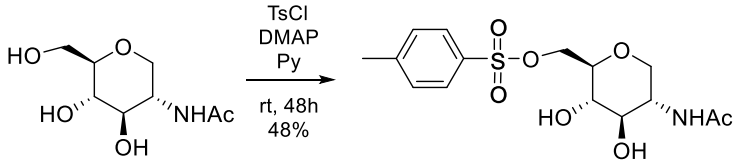
*The **final** and **new compound*** 2F, for in in vivo test, has been characterized through ^1H and *m/z* spectra.





^{1}H NMR (400 MHz, Deuterium Oxide): difficult interpretation

Scheme G - 6-sulfonate derivative from *N*-acetylglucosamine-1-deoxy

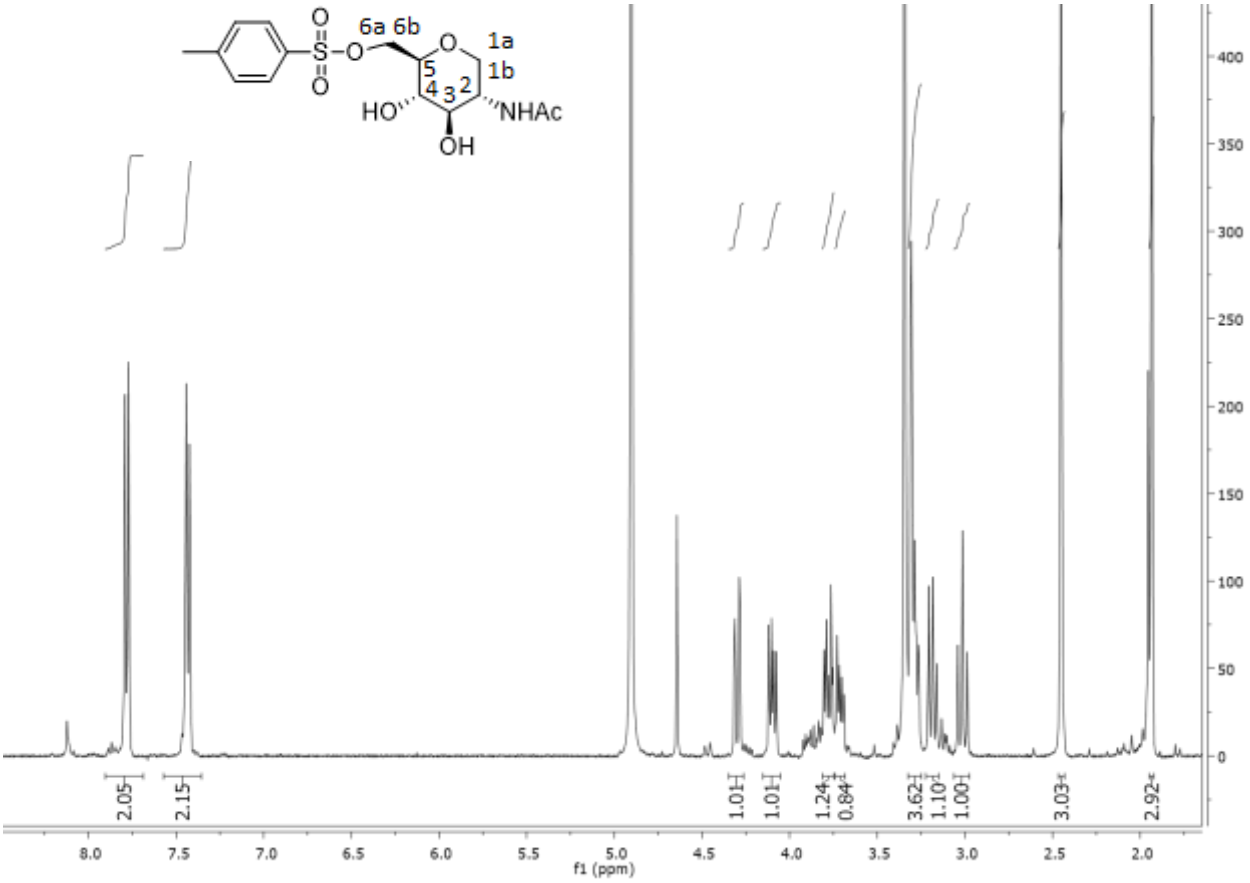


3A 1G

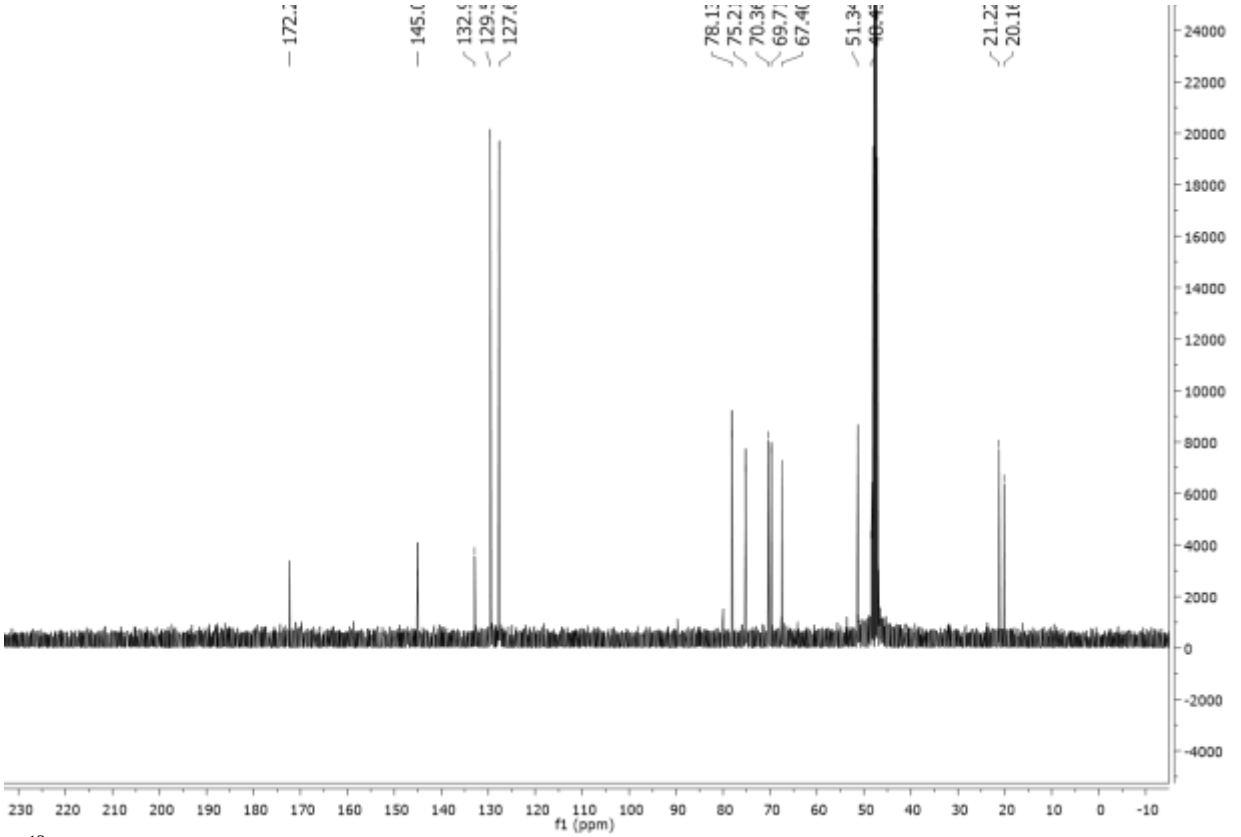
	g	MW	mmol	eq	d (g/mL)	V (mL)	[M]
3A	0,30	204,63	1,49	1			
Py		79,1			0,982	7,5	0,2
DMAP	0,02	122,17	0,15				0,1
TsCl in Py	5,17	180,60	27,15	1,2		30	

Compound 3A was dissolved in anhydrous pyridine and the solution was cooled to 0°C, then a solution of p-toluene-sulfonyl chloride in anhydrous pyridine was added dropwise and a catalytic amount of DMAP was added. The reaction was stirred approx. 0°C for 2.5 h. The reaction was quenched by the addition of ethanol, and the mixture was concentrated under reduce pressure. The residue was purified by column chromatography on silica gel AcOEt:MeOH 9:1, rf 0,33, (0,19g – 48% yield)), m/z_{calc} 359.10 (100.0%), 360.11 (16.2%), 361.10 (4.5%), 361.11 (1.4%), 361.11 (1.2%).

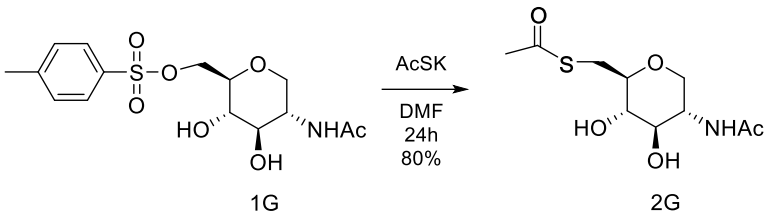
The ***new compound*** 1G has been characterized through ^{1}H and ^{13}C spectra.



[1]H NMR (400 MHz, Methanol-d) δ 7.78 (d, *J* = 8.2 Hz, 2H, **Ar**), 7.43 (d, *J* = 8.1 Hz, 2H, **Ar**), 4.30 (bd, *J* = 9.5 Hz, 1H, **H-6a**), 4.10 (dd, *J* = 10.7, 6.0 Hz, 1H, **H-6b**), 3.82 – 3.70 (m, 2H, **H-1a,2**), 3.33-3.25 (m, 3H, **H-3,4**), 3.22 – 3.15 (m, 1H, **H-5**), 3.01 (t, *J* = 10.6 Hz, 1H, **H-1b**), 2.49 (s, 3H, **MeTs**), 1.94 (s, 3H, **Ac**).

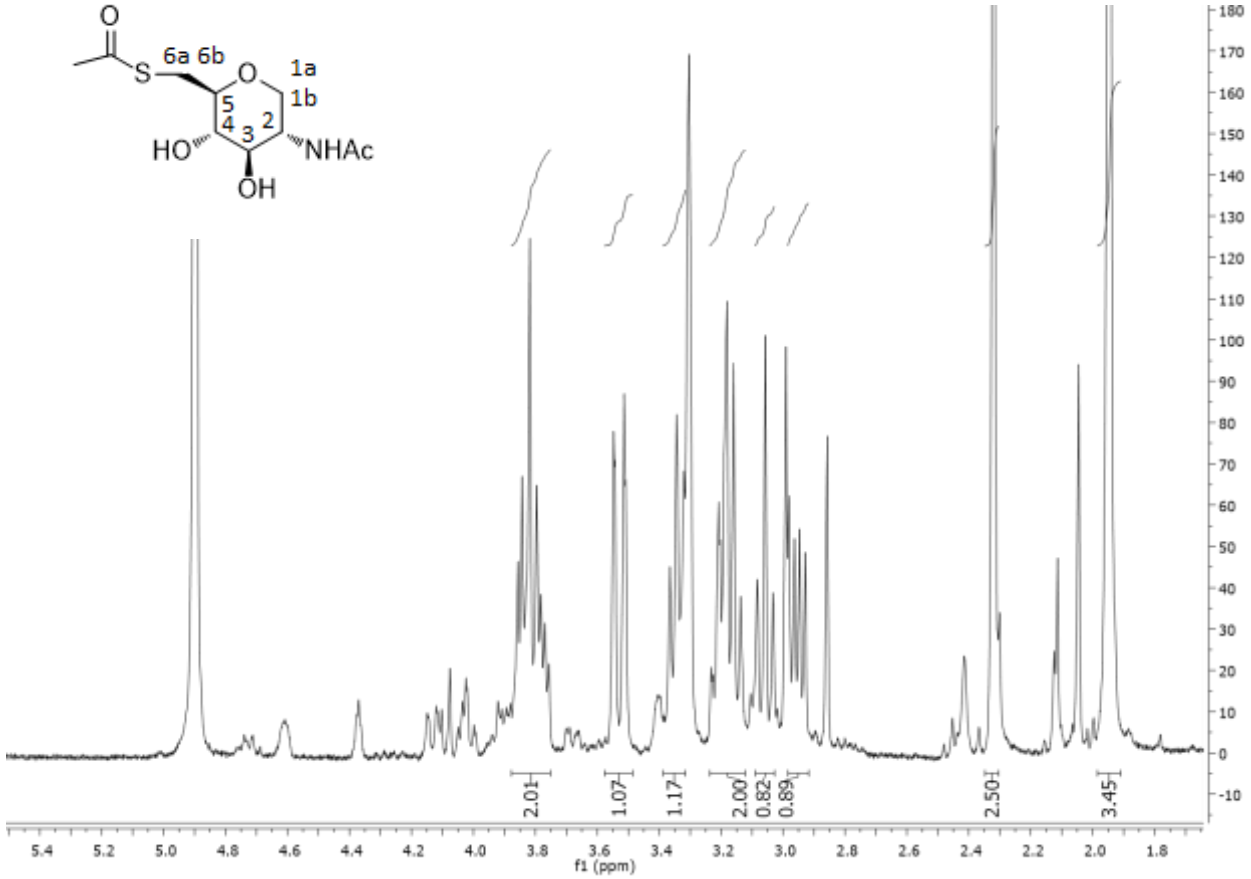


[13]C NMR (400 MHz, Methanol-d) δ 172.29 (CO), 145.06, 132.90 (CqAr), 129.55, 127.67 (CHAr), 78.13, 75.21, 70.36 (C 3,4,5), 69.71, 67.41, (C 6,1), 51.30 (C-2), 21.22, 20.16 tosyl-**CH_3**, **CH3**).

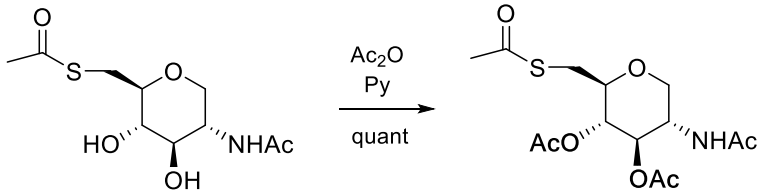


	g	MW	mmol	eq	d (g/mL)	V (mL)	[M]
1G	1.02	359.49	2.83	1			
AcSK	0,485	114.21	4.24	1,5			
DMF						28	0,1

1G was dissolved in DMF, AcSK was added to and the reaction mixture was stirred for 24h. The solvent was evaporated using a vacuum pump. The residue was purified by column chromatography on silica gel AcOEt:MeOH 9:1, rf 0,31 (0,41g – 43% yield), m/z_{calc} 263.08 (100.0%), 264.09 (10.8%), 265.08 (4.5%), 265.09 (1.0%).



[1]H NMR (400 MHz, Methanol-d) δ 3.88 – 3.75 (m, 2H, **H-1,2**), 3.53 (bd, *J* = 13.6 Hz, 1H, **H-6a**), 3.38-3.33 (m, 1H, **H-3**), 3.25 – 3.13 (m, 2H, **H-4,5**), 3.06 (t, *J* = 10.2 Hz, 1H, **H-1b**), 2.99 – 2.92 (m, 1H, **H6b**), 2.31 (s, 3H, **SAc**), 1.94 (s, 3H, **Ac**).

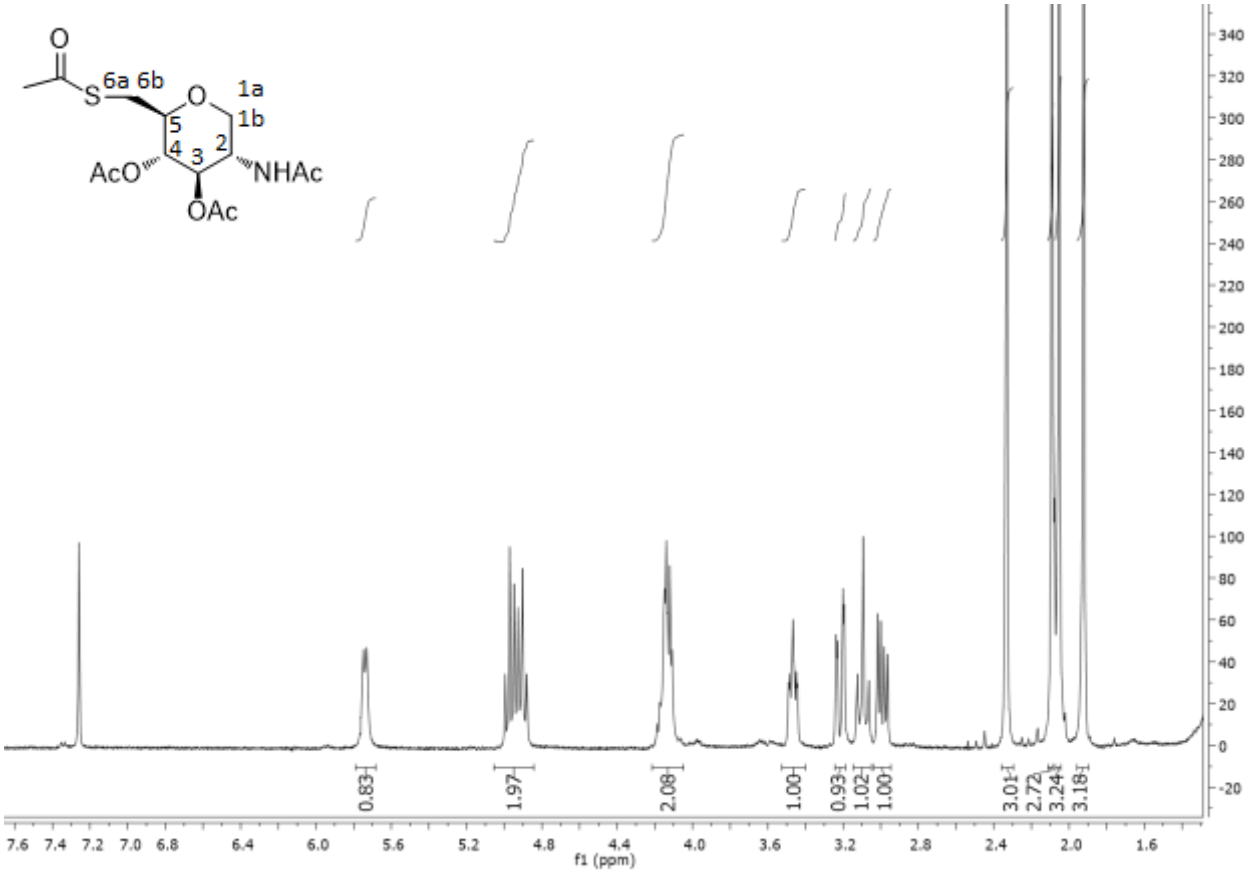


2G 3G

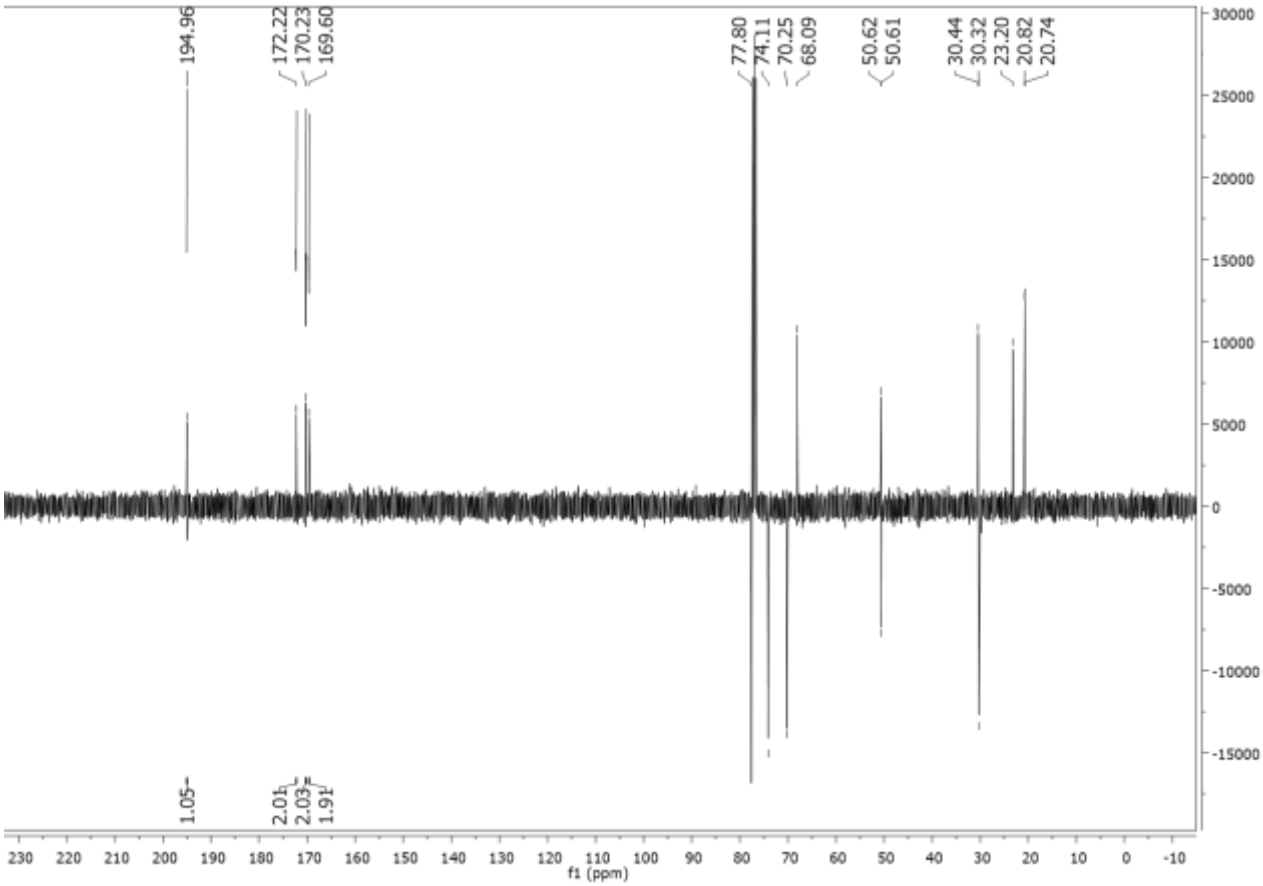
	g	MW	mmol	eq	d (g/mL)	V (mL)	[M]
2G	0,126	262,31	0,48	1			
Ac_2O		102,9		10		1,5	
Py		79,1		35		3,12	

Compound 2G was suspended in Py and cooled to 0°C, then acetic anhydride was added. The mixture was stirred at room temp. for 4 h. The reaction was quenched by the addition of MeOH (3 mL) and then concentrated under reduced pressure. The crude was purified by flash chromatography Et Pet:AcOEt 2:8, rf 0,22 (0,07g – quant), m/z_{calc} 347.10 (100.0%), 348.11 (15.1%), 349.10 (4.5%), 349.11 (1.4%), 349.11 (1.1%).

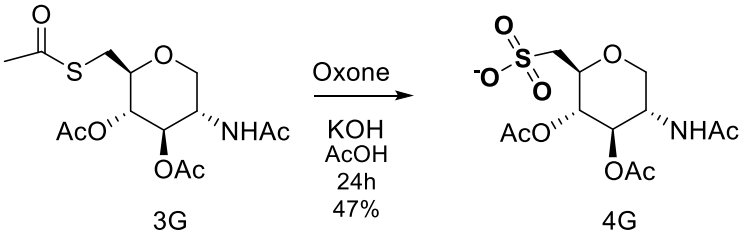
The ***new compound*** 3G has been characterized through ^{1}H and ^{13}C spectra.



[1]H NMR (400 MHz, Chloroform-d) δ 5.74 (d, *J* = 7.0 Hz, 1H, **NH**), 4.97 (t, *J* = 9.3 Hz, 1H, **H-3**), 4.90 (t, *J* = 9.5 Hz, 1H, **H-4**), 4.20 – 4.09 (m, 2H, **H-1a,2**), 3.51 – 3.42 (m, 1H, **H-5**), 3.22 (dd, *J* = 14.3, 2.6 Hz, 1H**, H-6a**), 3.09 (t, *J* = 12.4 Hz, 1H, **H-1b**), 2.99 (dd, *J* = 14.3, 6.7 Hz, 1H, **H-6b**), 2.33 (s, 3H, **SAc**), 2.09 (s, 2H, **Ac**), 2.05 (s, 2H, **Ac**), 1.92 (s, 1H, **Ac**).



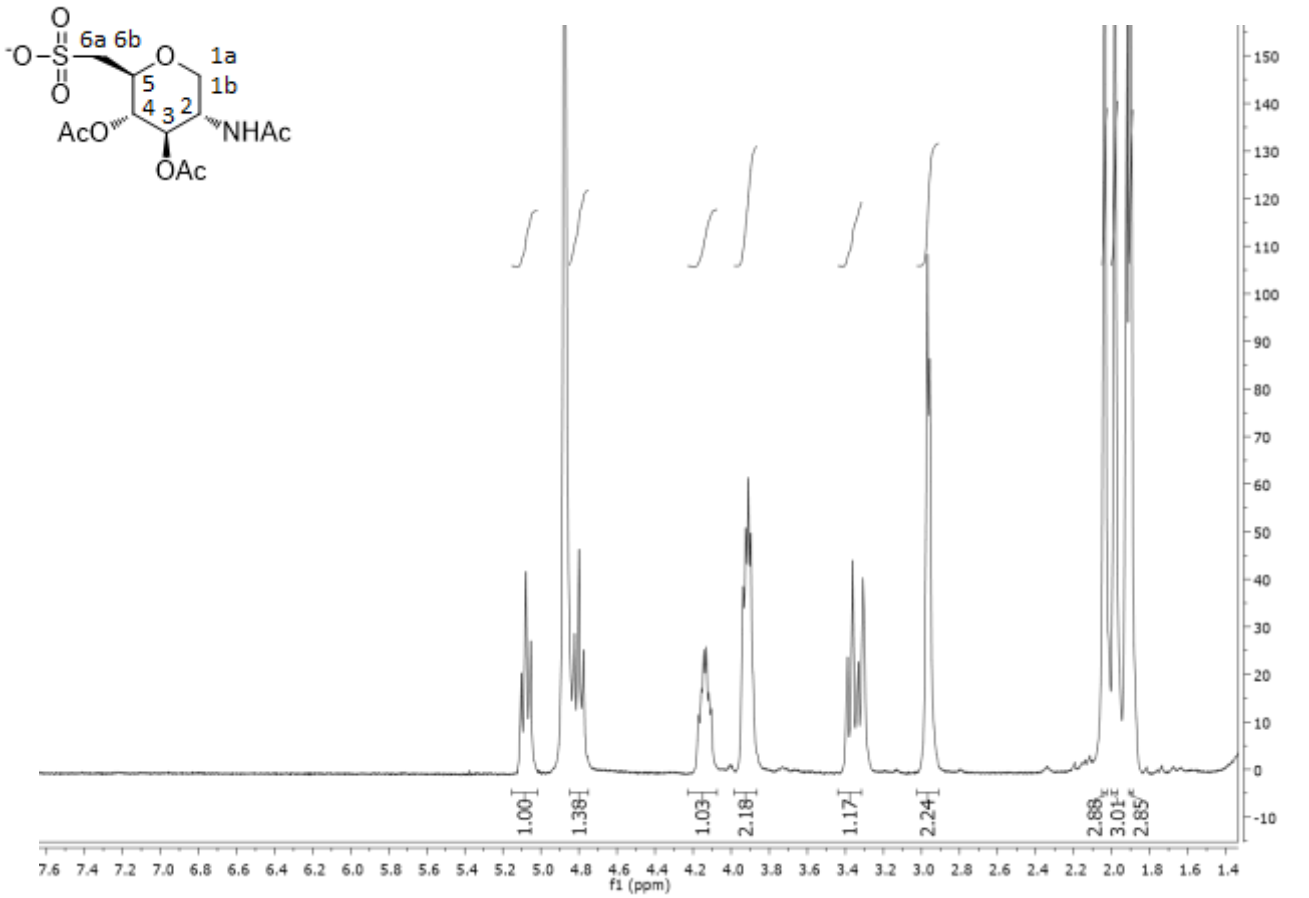
[13]C NMR (400 MHz, Chloroform-d) δ 194.96 (impurities), 172.22, 170.23, 169.60 (**C**O), 77.80, 74.11, 70.25, 68.09, 50.62, 50.61, 30.44, 30.32 (**C 5,3,4,1,2,6**), 23.20, 20.82, 20.74 (**CH_3**, tosyl-**CH_3**)



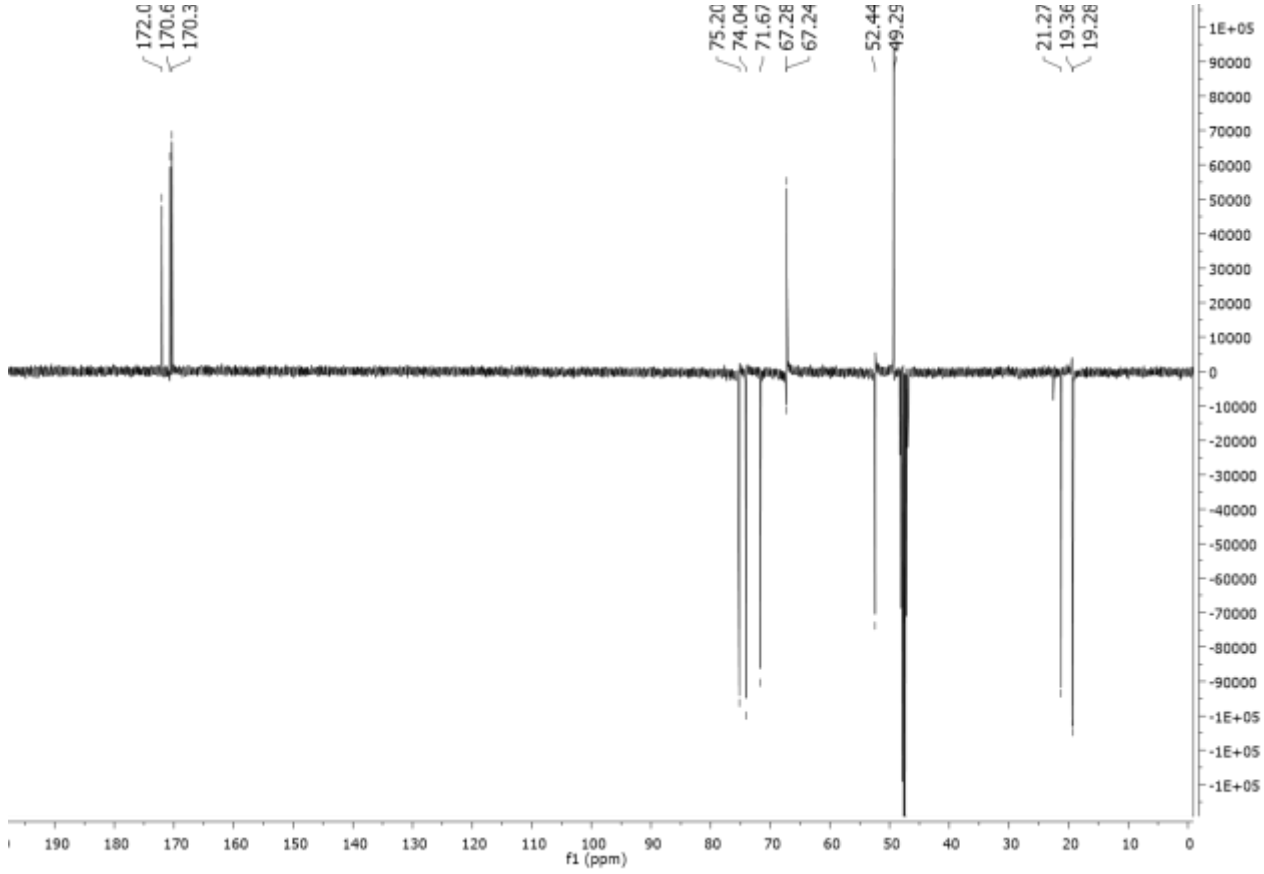
	g	MW	mmol	eq	d (g/mL)	V (mL)	[M]
3G	0,165	347,48	0,475	1			
Oxone	1,12	614,76	1,82	3,84			
KOAc	0,859	94,14	8,74	18,41			
AcOH						10	

The thioacetate 3G was dissolved in AcOH to which KOAc followed by Oxone was added, and the reaction was left stirring for 24h. After extraction with AcOEt the final product was purified AcOEt:MeOH 7:3, rf 0,32 (0,07g – 47% yield), m/z_{calc} 352.07 (100.0%), 353.07 (13.0%), 354.07 (4.5%), 354.08 (1.8%).

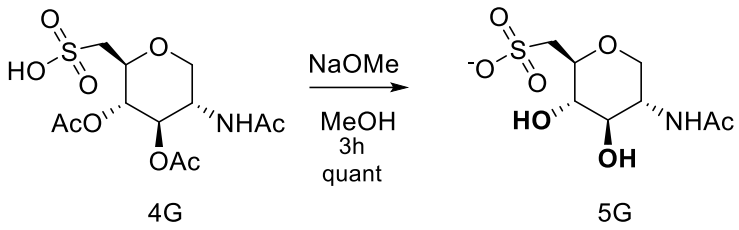
*The **final** and **new compound*** 4G, for in vivo test, has been characterized through ^{1}H and ^{13}C spectra.



^{1}H NMR (400 MHz, Methanol-d) δ 5.08 (t, *J* = 9.7 Hz, 1H, **H-4**), 4.80 (t, *J* = 9.5 Hz, 1H, **H-3**), 4.19 – 4.08 (m, 1H, **H-2**), 3.95-3.90 (m, 2H, **H-1a**,**5**), 3.36 (t, *J* = 11.4 Hz, 1H, **H-1b**), 2.98-2.95 (m, 2H, **H-6a**,**b**), 2.03 (s, 3H, **Ac**), 1.98 (s, 3H, **Ac**), 1.90 (s, 3H, **Ac**).



^{13}C NMR (400 MHz, Methanol-d) δ 172.02, 170.62, 170.35 (**C**O), 75.20, 74.04, 71.67 (**C**H), 67.28, 52.44 (**C**2), 49.29 (CH), 21.27, 19.36, 19.28 (CH_3**Ac**)

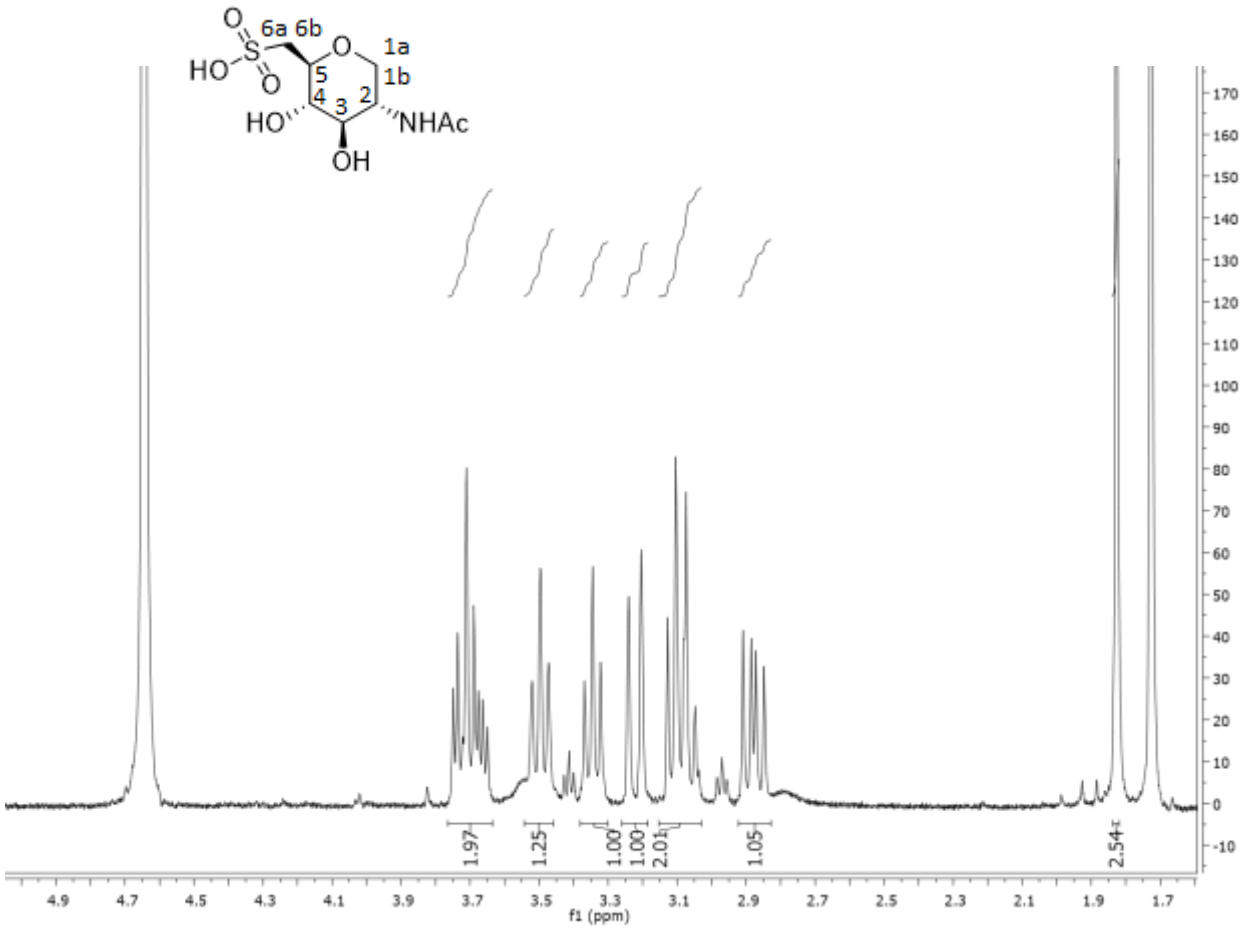


	g	MW	mmol	eq	d (g/mL)	V (mL)	[M]
4G	0,090	353,35	0,255	1			
NaOMe	0,082	54,02	1,52	6			
MeOH						1,5	0,1

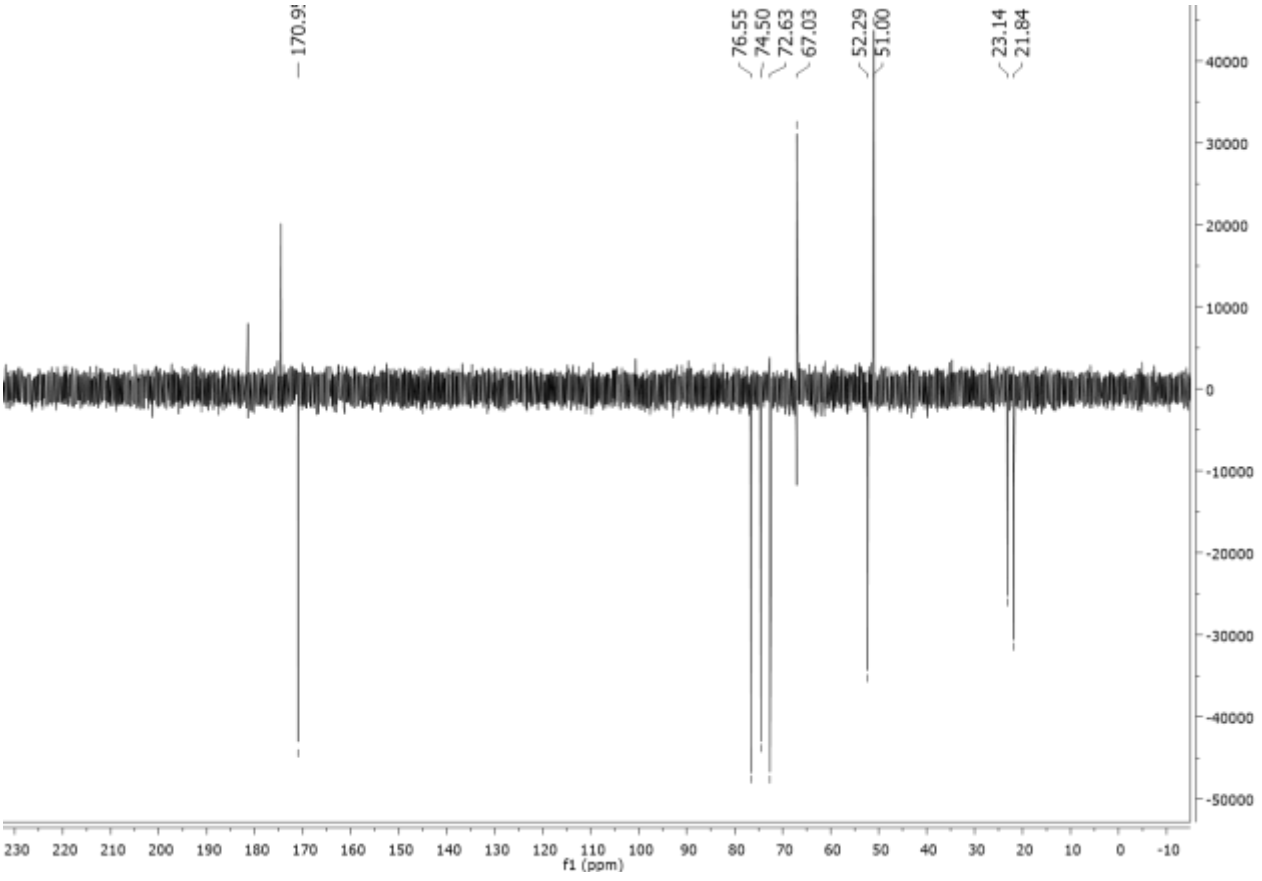
NaOMe was added to the product 4G dissolved in dry MeOH under inert atmosphere. After 3h the stirring the reaction was neutralized through the addition of Amberlite 5% HCl to pH 6. The resin was filtered and the solvent was evaporated.

The compound was purification RP-C18, gradient H_2O:MeOH (0,06g – quant), m/z_{calc} 268.05 (100.0%), 269.05 (8.7%), 270.05 (4.5%), 270.05 (1.4%).

*The **final** and **new compound*** 5G, for in in vitro test, has been characterized through ^{1}H and ^{13}C spectra.



^{1}H NMR (400 MHz, Deuterium Oxide) δ 3.77 – 3.64 (m, 2H, **H-1a, 2**), 3.50 (t, *J* = 9.6 Hz, 1H, **H-3**), 3.35 (t, *J* = 9.3 Hz, 1H, **H-4**), 3.22 (d, *J* = 14.8 Hz, 1H, **H-6a**), 3.10-3.06 (m, 2H, **H5, 1b**), 2.88 (dd, *J* = 14.9, 9.5 Hz, 1H, **H-6b**), 1.90 (s, 3H, **Ac**).



^{13}C NMR (400 MHz, Deuterium Oxide) δ 170.95 (**C**O), 76.55, 74.51, 72.73 (**C 3,4,5**), 67.04 (**C1**), 52.29 (**C6**), 51.00 (**C2**), 23.14, 21.84 (**Ac**).

A.4 HPLC: setting method

Buffer preparation[78]

Here are showed the buffer used for the analysis:

Buffer A	100 mM di buffer phosphate potassium pH 6.4 + 8mM di tetrabutyl ammonio hydrogen sulphate (TBAHS)
Buffer B	70% buffer A (without TBAHS) + 30% di ACN per HPLC.

Phosphate Buffer[79]

The phosphate buffet at pH 6.4 has been prepared weekly using KH_2PO_4 and K_2HPO_4 and preserved in freezer.

pH	Volume of 1 M K_2HPO_4 (ml)	Volume of 1 M KH_2PO_4 (ml)
5.8	8.5	91.5
6.0	13.2	86.8
6.2	19.2	80.8
6.4	27.8	72.2
6.6	38.1	61.9
6.8	49.7	50.3
7.0	61.5	38.5
7.2	71.7	28.3
7.4	80.2	19.8
7.6	86.6	13.4
7.8	90.8	9.2
8.0	94.0	6.0

Buffer phosphate 0.1 M at 25°C.

Experimental conditions

In order to get the analysis, we have used a gradient method; in table below are shown the elution time at flow rate of 0.8 ml/min, pression of 400 bar and λ 254 nm.

Elution time (min)	% buffer B
15.00	0
35.00	0-77
36.00	77-100
46.00	100
56.00	100-0

Eluent gradient for analysis

[78] Nakajima, Kazuki, et al. "Simultaneous determination of nucleotide sugars with ion-pair reversed-phase HPLC." *Glycobiology* 20.7 (2010): 865-871.

[79] http://www.unl.edu/cahoonlab/phosphate%20buffer.pdf

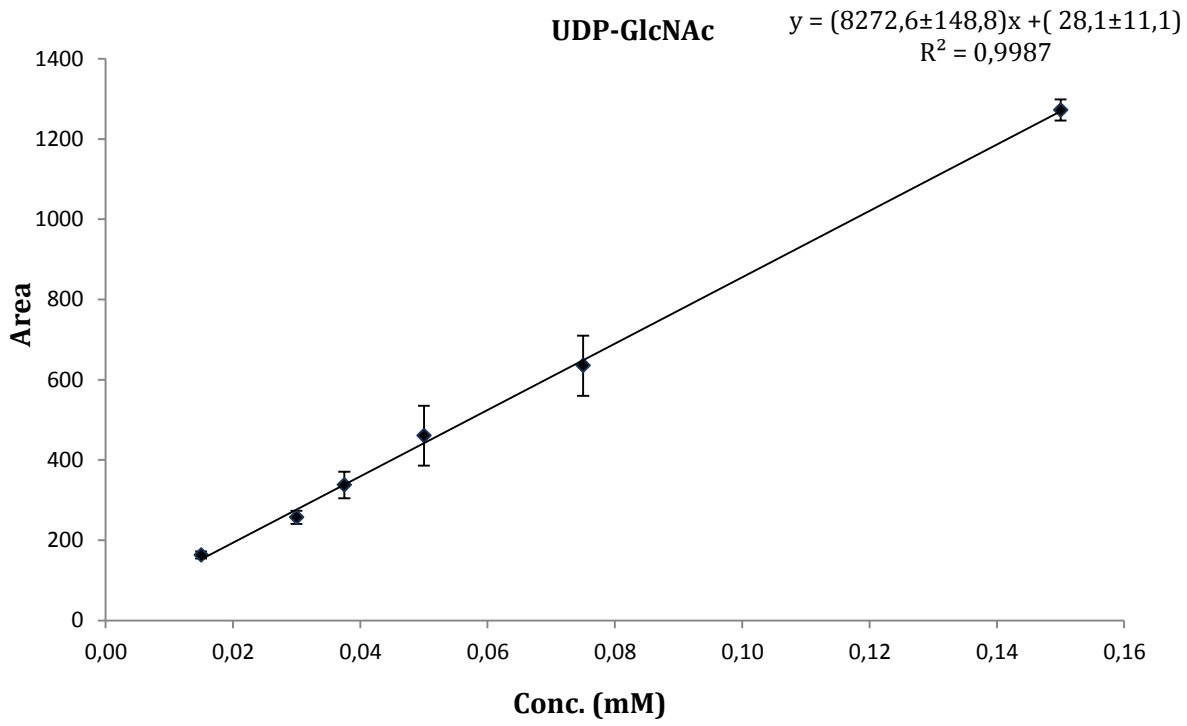
Calibrations lines

The UDP-GlcNAc 1.5 mM is the stock solution and, subsequently, standard calibration solutions were prepared by drawing the same volume of the two standards from the same dilution by performing cascade dilutions. (N.B . All dilutions were made using H_20 milliQ for the highest degree of purity).

Standard for the calibration lines

Diluition	Conc. UDP-GlcNAc (mM)	Mix
1:100	0.0150	100 mL UDP-GlcNAc 0.03 mM + 100 mL ATP 0.0394 mM
1:50	0.0300	50 mL UDP-GlcNAc 0.06 mM + 50 mL ATP 0.0788 mM
1:40	0.0375	50 mL UDP-GlcNAc 0.075 mM + 50 mL ATP 0.0985 mM
1:30	0.0500	50 mL UDP-GlcNAc 0.1 mM + 50 mL ATP 0.13 mM
1:20	0.0750	50 mL UDP-GlcNAc 0.15 mM + 50 mL ATP 0.197 mM
1:10	0.1500	50 mL UDP-GlcNAc 0.3 mM + 50 mL ATP 0.394 mM

Calibration lines



The validation of the method must be done through the evaluation of technical parameteres

The *linearity* is delivering results directly proportional to the concentration of the analyte at a given interval by using the calibration line. When the equation of the line is obtained, the quadratic correlation coefficient (R^2), in statistics also called Pearson's correlation index, is the index of the possible linearity relationship between two statistical variables. The value of R^2 can range from 0 to 1, respectively, in cases of absolute inadequacy of the model and absolute perfection of the model itself. Good correlation coefficient values of at least 0.99 are generally considered.

For *sensitivity* of the analytical method, we mean the variation of the Y signal (absorbance, electrical current, etc ...) as a function of the C concentration of the analyte. HPLC has good sensitivity because it can be assume that a significant variation in signal corresponds to an important variation of concentration.

Selectivity indicates the ability of an analytical method to not interfere with the presence of other components: UDP-GlcNAc is measured by UV detector at 254 nm wavelength (*Taniguchi, N. et al. , 2010*).

	UPD-GlcNAc
LOD (mM)	0,0059
LOQ (mM)	0,0196

The *precision* of the method is the agreement between repeated measurements under the same conditions (repeatability), depends on random errors that can not be eliminated, and is evaluated by a set of replicated measurements through standard deviation.

In order to get the best quantification of analyte, it's important to calculate the *leakage Limit* (*LOD*) the smallest amount of analyte in the sample that may be revealed but not necessarily quantified in the declared experimental conditions. On the other hand, the *quantification limit* (*LOQ*) represents the minimum quantity of analyte in a sample, quantitatively determined with adequate precision and accuracy.[**80**]

Analytical data of samples of the UDP GlcNAc nucleotide sugars were highly reproducible over three repetition of experiments.

[80] Shrivastava, Alankar, and Vipin Gupta. "Methods for the determination of limit of detection and limit of quantitation of the analytical methods." *Chronicles of Young Scientists* 2.1 (2011): 21-21.

Concentrations UDP			
Sample	Mean	Standard deviation	Coefficient of variance
1_100	163	8,5	0,05
1_50	257	16,5	0,06
1_40	338	33,2	0,10
1_30	461	74,9	0,16
1_20	635	74,7	0,12
1_10	1272	26,3	0,02
1_5	2870	90,3	0,03

Enzimatic reaction[81]

The enzymatic reaction has been done using cellular extract and adopting the following procedure:

- cellular extracts (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% NP40, 0.1 % SDS, phosphatase and protease inhibitors)
- 25 μL GlcNAc 6P (100 mM)
- 2,5 μL Glc 1,6 bis fosfato (3.6 μM)
- 25 μL UTP (100 mM)
- TRIS (TRIZMA base)

The enzymatic reaction is incubated at 37°C for 30', 1h, 1h 30' and 3h. Centrifuge 14 000 RCF for 20'

[81] Zhang, W., Jones, V. C., Scherman, M. S., Mahapatra, S., Crick, D., Bhamidi, S., ... & Ma, Y. (2008). Expression, essentiality, and a microtiter plate assay for mycobacterial GlmU, the bifunctional glucosamine-1-phosphate acetyltransferase and N-acetylglucosamine-1-phosphate uridyltransferase. *The international journal of biochemistry & cell biology*, *40*(11), 2560-2571.

Enzymatic reaction with inhibitors in cellular extract

To carry out the enzymatic reaction, lysed cell extracts were used, prepared by the following procedure:

10 µL	OD1	OD2	µg/µL	mg/aliquota	mg tot
I	0.355	-	5,29	1,589	9,5

30 µL	OD1	OD2	µg/µL	mg/aliquota	mg tot
I	0.276	0.278	4.7	-	1,300
II	0.355	-	5,29	1,589	9,5

A.5 Biological part

C Log*P*

cLog*P*, which measures the compound's hydrophilicity (log ([] octanol/[] water)), was predicted by using ChemDraw software.

Cell lines and shPGM3 (AGM1) clones[82]

Human breast cancer MDA-MB-231 cells were routinely cultured in Dulbecco's modified Eagle's (DMEM) medium containing 25 mM glucose, supplemented with 4 mM L-glutamine, 100 U/mL penicillin, 100 mg/mL streptomycin (complete medium) and 5% fetal bovine serum. The other breast cancer cell lines (SKBR-3, BT-474, Mcf7, T-47D, MDA-MB-361) and Retinal Pigmented Epithelial (hTERT-RPE) cells were culture in the same medium, with exception for serum that was 10%, while MDA-MB-468 were cultured in DMEM containing 25 mM glucose, supplemented with 2 mM L-glutamine, 100 U/mL penicillin, 100 mg/mL streptomycin and 10% fetal bovine serum.

Cells were grown and maintained according to standard cell culture protocols and kept at 37 °C with 5% CO2. The medium was replaced every 2-3 days and cells were splitted or seeded for experiment when reached the sub-confluence. All reagents for media were purchased from Thermo Fisher Scientific. Silencing of PGM3 was obtained transfecting MDA-MB-231 cells with plasmid from Sigma- Aldrich (MISSION shRNA for PGM3: NM_015599, clone ID: NM_01559.1-1309s21c1); also control plasmid (shCTR, non human or mouse shRNA) was purchased from Sigma-Aldrich (SHC002).

To transfect cells Polyfect transfection reagent (Qiagen) was used. Briefly, 2x105 cells were seeded in MW6 well, after 24h 1.5mg DNA was added to 100mL of medium (opti-mem, Thermo Fisher Scientific) and mixed; after that 10mL of transfection reagent was added to the mix, vortexed and incubated at room

[82] Thanks to Francesca Ricciardiello, Giuseppina Votta, Roberta Palorini - Department of Biotechnology and Biosciences, University of Milano-Bicocca, Milan, 20126, Italy

temperature for 10 minutes; then 0.6mL of cell medium was added and the transfection mixture distributed in the cell well containing 1mL of complete cell medium.
To select clones presenting stable transfection, cells were treated with 2mg/mL puromycin (Euroclone) for a week and then seeded in MW96 wells at a density of 1 cell/well under puromycin selection. After their proliferation and expansion, different clones were tested for PGM3 levels.

Cell treatment and cell viability evaluation

Where not differently specified, for experiments cells were seeded in complete growth medium and after 24 hours washed twice with Phosphate Buffer Saline (PBS) and incubated in complete medium with or without the inhibitors (tunicamycin, 3B, 2B). In some cases, a double treatment was performed and the molecules (FR054 with Glc*N*ac, thapsgargin, tunicamycin, NAC)
added to the cells as indicated in specific experimental schemes presented in the text or figures.
All chemicals and inhibitors were purchased from Sigma-Aldrich, except for Thapsgargin (Vinci- Biochem) and ERO1 inhibitor EN460 (Calbiochem, Merck Millipore). To measure cell proliferation, harvested cells were counted using the Burker chamber. Where indicated, cell viable count was performed using Trypan Blue Stain 0.4% (Thermo Fisher Scientific). GI50 calculation was performed with GraphPad Prism.
Cell viability was measured also with MTT test (Roche from Sigma-Aldrich) according to the manufacturer's protocol. Briefly, 5x103 cells were seeded into 100µL of medium in 96 flat bottom multiwell. At a specific time point 10µL of the labeling reagent were added to the cells. After 4 hours at 37ºC in humidified atmosphere, 100µL of the solubilization solution were added, and the 22 cells incubated overnight. After that, the absorbance at 620nm was measured with an ELISA reader; absorbance at 800nm was used as reference.
For the colony formation assay, after specific treatment the cells were collected and then re-seeded at low density in complete medium. After 12 days, cells were washed twice with PBS, fixed in PBS-formaldehyde 5%, and stained with 0.1% crystal violet for 5 minutes. After colorant dissolving by acetic acid 10% the absorbance was analyzed at spectrophotometer.

Adhesion and migration assays

For attachment assays, after a treatment of 24 hours with 250mM 3B or 2B, 1x105 cells/sample were washed, resuspended in DMEM 0.5% serum, seeded and allowed to adhere for 1 hour at 37°C in 12-well plates coated overnight at 4°C with 0.1% heat-denatured BSA. Non adherent cells were then removed by gentle washing with PBS conteining Ca^{2+} and Mg^{2+} (Euroclone) whereas adherent cells were trypsinized and counted.
Cell migration assays were performed in Boyden chambers (Neuro Probe), as previously described with minor modifications (Alfano et al., 2010). Briefly, 1x105 cells/chamber grown for 24 hours with 250mM FR054 or FR053 were detached by mild trypsinization, resuspended in serum free DMEM and inoculated into the upper compartments of Boyden chambers. 1% serum diluted in 26 DMEM with 0.1% bovine serum

albumin (BSA) was used as chemoattractant in the lower compartments. The two compartments are defined by 8-μm pore size PVPF filters (polyvinylpyrrolidone-free filters, 0.33 cm2, Sigma-Aldrich) coated overnight at 4°C with 50mg/ml collagen VI (Sigma-Aldrich) in order to accelerate cell attachment. After 4 hours of incubation at 37°C, the cells on the upper side of the filters were removed by scraping and cells on the lower side of the filters were stained with Mayer's hematoxylin (Sigma-Aldrich) and counted.

Mass spectrometry to detect intracellular molecules[83]

MS were recorded on a QSTAR elite LC/MS/MS system with a nanospray ion source.

1. **Sample preparation**

$7x10^6$ MDA-MB-231 cells (untreated and treated with 2B) were collected after trypsinization and washed twice in PBS. Cell pellet was collected in 1.5mL tubes and kept in ice. The dried pellet was resuspended using 70μL of solution D. The analyte calibrants were treated with the same procedure employed for the sample.

The calibration of the compounds 2B (de-acetylated and phoshorylated form) was obtained using a multiple chemiometric approach due to the lack of the analytical standard. The quantitation data of this analyte are so semiquantitative. The MRM transition 282 -> 97 has been selected to quantify the compounds. The selectivity of the MRM transition has been evaluated on the basis of a qualifier signal at m/z 242.9 and its structure has been defined through fragmentography tandem mass spectrometry (MS/MS) studies.

2. **Mass Spectrometry**

Acquisition parameters, unless stated otherwise, was as follows. The direct infusion flow rate,p rovided by means of a syringe pump, was 5μL/min with additional solvent provided at a flow rate of 100μL/min by a HPLC pump placed in-line. A Bruker Daltonics HCTultra Ion-Trap mass spectrometer (Bruker Daltonics, Breme, Germany) was used. The spectra were obtained in the positive and negative mode. An average of two microscans and a rolling average of four spectra were acquired. The nitrogen drying-gas temperature was 350°C, and its flow rate was 1L/min. The nitrogen nebulizing-gas flow rate was 12L/min. The internal capillary voltage was between –100 and –-1500V. The end-plate potential was –500 V. The N2 curtain gas flow rate was varied between 0.5 and 6L/min. The ionization sources parameters were fixed as follow: ESI ionization voltage (3000V) was applied via a home built external power supply to the spray needle, the APCI needle current was 3000nA, and the SACI surface potential was 47V used to obtain the in source ionization and the inter capillary voltage gradient was employed to produce the CIMS effect. HyStar software (Bruker Daltonics, Breme, Germany) was used for data acquisition. DataAnalysis(Bruker Daltonics) was used for data treatment.

[83] Thanks to H. De Vito and Ion Source & Biotechnologies (ISB)

Acknowledges

Il periodo di tesi di dottorato è stato un percorso da me tanto desiderato, nella consapevolezza di poter affrontare nuove sfide, poter migliorare e approfondire le conoscenze la livello scientifico. Esso è stato il risultato di un "corso di formazione" in cui scegli di essere coinvolto personalmente: mi ha permesso di affacciarmi al campo della sintesi e alla realtà delle applicazioni biologiche, fornendomi gli strumenti indispensabili per una ottima consapevolezza in entrambi i campi.

Nel mio cuore c'è la più grande motivazione che mi ha incoraggiato a perseguire il mio sogno e iniziare questo viaggio, pieno di difficoltà, sfide e fallimenti, cadute e scoraggiamento, ma anche soddisfazioni... un percorso che mi ha forgiato e temperato e mi ha aiutato a crescere moralmente e culturalmente.

Quindi grazie a TE e grazie a me stessa.

Ringrazio la mia famiglia e *colui che è e diventerà la mia famiglia*.

E ora passiamo ai ringraziamenti formali:

- alla Prof.ssa *Barbara La Ferla*, la mia tutor, per avermi dato l'opportunità di entrare in contatto con il campo di ricerca di sintesi durante questi 3 anni e a tutti gli studenti e collaboratori del Laboratorio 4014 – U3;
- al Prof. *Francesco Nicotra*, mio correlatore: esprimo la mia più sentita gratitudine per essere stata seguita, ascoltata e incoraggiata e, quando necessario, spronata a migliorare;
- Grazie all'Associazione Italiana per la Ricerca sul Cancro (*AIRC*), che ha finanziato la ricerca;
- un ringraziamento particolare al Prof. *De Gioia Luca* e alla Dott. *Giulia Filippi* per l'aiuto e l'insegnamento nella parte computazionale;
- Grazie al *laboratorio di biochimica tumorale 5048-U3* per i test biologici, al dott. H. De Vito e Ion Source & Biotechnologies (ISB) per l'analisi dei dati LC / MS;
- Grazie al Prof.re *Luca Zoia* per l'insegnamento dell'uso della strumentazione HPLC;
- Grazie al *CycloLab*, nella figura del dott. *Milo Malanga*, per il periodo di tirocinio Erasmus.

Ac	Acetyl
AcCl	Acetyl chloride
Ac_2O	Acetic anhydride
AcOEt	Ethyl acetate
AcOH	Acetic acid
AcSK	Potassium thioacetate
AGM1	N-Acetylglucosamine-phosphatemutase
AIBN	Azobisisobutyronitrile
ATP	Adenosine triphospate
Bn	Benzyl
BnBr	Benzyl bromide
Bu_3SnH	Tributyltin hydride
CD(s)	Cyclodextrin(s)
CH_2Cl_2/ DCM	Dichloromethane
DIPEA	*N*,*N*-Diisopropylethylamine
DMAP	Dimethylaminopyridine
DMF	*N*,*N*-Dimethylformamide
DNA	Deoxy-ribonucleic acid
ER	Endoplasmic Reticulum
EtOH	Ethanol
GFAT	Fructose 6-phosphate amidotransferase
GlcNAc	*N*-acetylglucosamine
GlcNAc-1-P	*N*-acetylglucosamine-1-phosphate
GlcNAc-6-P	*N*-acetylglucosamine-6-phosphate
GNK	Acetylglucosamine kinase
HBP	Hexosamine Biosynthetic Pathway
HBr	Bromide acid
HPLC	High-performance liquid chromatography

KRAS	V-Ki-ras2 Kirsten viral oncogen homologous sarcoma
MeCN/CH_3CN	Acetonitrile
MeOH	Methanol
MeONa	Sodium methylate
Me_3SiCl	Trimethylsilil chloride
NaN_3	Sodium azide
NaH	Sodium hydride
$NaHCO_3$	Sodium bicarbonate
NaI	Sodium iodide
NMR	Nuclear Magnetic Resonance
$Pd(OH)_2/C$	Palladium hydroxide on activated charcoal
PDAC	Adenocarcinoma of the Pancreatic Duct
Ph_3P	Triphenylphosphine
Py	Piridine
Py-SO_3	Sulfur trioxide pyridine complex
$SiMe_3N_3$	Trimethylsilil azide
TBAF	Tetra-*N*-butylammonium fluoride
TEA	Triethylamine
THF	Tetrahydrofuran
TLC	Thin layer chromatography
Ts	Tosyl
TsCl	Toluensulfonyl chloride
TMSOTf	Trimethylsilyl trifluoromethanesulfonate
UPR	Unfolded Protein Response
UDP-N-GlcNAc	Uridine diphosphate *N*-acetylglucosamine