UNIVERSITA' DEGLI STUDI DI MILANO - BICOCCA

SCUOLA DI DOTTORATO DI SCIENZE

Dipartimento di Biotecnologie e Bioscienze

Dottorato di Ricerca in Biotecnologie Industriali
XXV ciclo

Glycoderivatives: drug candidates and molecular tools

Giuseppe D'Orazio

Tutor: Dott.ssa Barbara La Ferla

To my family,

my friends

and my beloved Marta

Index

Preface 5

Riassunto 7

Chapter 1. Introduction 13

Carbohydrates – general overview 14

The therapeutic importance of carbohydrates: the development of glycomimetics 23
C-glycosides as glycomimetic scaffolds 27

Role of carbohydrates and glycomimetics in anti-inflammatory processes: inflammation and SGLT1 32

Inflammation, sepsis and septic shock 32
SGLT1: function and structure 43

Iminosugars: azaglycoderivatives and their widespread biological functions 58

Glycosidases inhibition: therapeutic role and biochemical features 58
Iminosugars as glycosidase inhibitors 62
Biological activity of iminosugars and their therapeutic applications 66

Multivalency and glycobiology 73

References 76

Chapter 2. Development of a library of dansyl-*C*-glycoderivatives as SGLT1 ligand tools: synthesis, biological evaluation and mechanism of action. 81

Abstract 82

Introduction 82

Results and discussion 84

Conclusions 97

Experimental Section 97

References 124

Chapter 3. Synthesis of a labeled SGLT1 ligand for *in vitro* and *in vivo* trafficking studies 127

Abstract 128
Introduction 128
Results and discussion 130
Conclusions 133
Experimental Section 133
References 142
Chapter 4. Generation of gold nanoparticles decorated with synthetic ligands of co-transporters SGLT-1 and B^0AT1 for the investigation of the multivalent-synergistic effect. 145
Abstract 146
Introduction 146
Results and Discussion 149
Conclusions 161
Experimental Section 162
References 185
Chapter 5. Iminosugar-decorated calix[4]arenes: chemical tools for the investigation of multivalent inhibition of glycosidases. 187
Abstract 188
Introduction 188
Results and Discussion 190
Conclusions 199
Experimental Section 200
References 216
Chapter 6. Antiproliferative activity of Arsenical C-glucoside derivative on neuroblastoma cell line SN-K-BE 219
Abstract 220
Introduction 220

Results and Discussion 222
Experimental Section 224
References 227
Chapter 7. Conclusions 229
Pubblications and communications 233
Papers 234
Oral communications 234
Other communications 235
List of abbreviations 237

Preface

The research work presented in this manuscript is a result of collaborations between the research group in which I worked, supervised by Dr. Barbara La Ferla, PhD and Prof. Francesco Nicotra, and several Italian and European coworkers.

Chapter 2 and **chapter 3** describe a research project in collaboration with the research group of Prof. Cristiano Rumio (University of Milan – Humanitas Clinical Institute (Milan)). An acknowledgment goes to Dr. Diego Cardani which performed the biological experiments.

In **chapter 4** is described a research work carried out in collaboration with the research group of Prof. Soledad Penàdes (Laboratory of Biofunctional Nanomaterials – CicbiomaGUNE, Donostia – San Sebastiàn, Basque Contry, Spain). In particular, the preparation of the gold nanoparticles was carried out during an one month research period in this research center, under the supervision of Dr. Marco Marradi, PhD.. My research period abroad was carried out within the COST Network project (COST Action CM1102).

Chapter 5 is a result of a collaboration with the research group of Prof. Alessandro Dondoni (University of Ferrara, Italy) and Prof. Alberto Marra (University of Montpellier 2, France). In particular, their work was focused on the preparation of the calixarene structures and their functionalization with the prepared iminosugars.

Chapter 6 presents a work performed in collaboration with Prof. Marco Salvetti research group (University of Rome – La Sapienza) and Dr. Michele Pitaro (Xenuspharma s.r.l, Rome).

Riassunto

I carboidrati rappresentano la classe di macromolecole biologiche più abbondante in Natura. La complessità chimica e strutturale che li caratterizza è alla base delle svariate funzioni svolte nel mondo vivente. Tale variabilità è dovuta sia dal numero di unità che li compongono, dal numero di atomi di carbonio che costituiscono ciascuna unità e dal tipo di funzionalità carbonilica presente. Queste distinzioni comportano una notevole variabilità in termini di possibili strutture generabili; a ciò va però aggiunta l'enorme diversità dovuta alle varie combinazioni generate dalla presenza di più carboni stereogenici in ogni singolo monosaccaride.
La Natura si avvantaggia di tutte queste peculiarità mostrate dai carboidrati, che sono utilizzati dagli esseri viventi come molecole di riserva e fonte di energia, come elementi strutturali di piante o come *carrier* di informazioni. È senza dubbio quest'ultima funzione quella di maggiore interesse da un punto di vista biologico e clinico: le cellule del sistema immunitario, ad esempio, comunicano tra loro grazie all'interazione tra specifiche proteine di membrana espresse sulla superficie di talune cellule, e complesse strutture zuccherine costituenti glicoproteine di membrana di cellule partner. La ricerca e lo studio di queste complesse catene oligosaccaridiche o di loro analoghi è alla base, per esempio, dello sviluppo di vaccini, per bloccare o scardinare lo stesso riconoscimento che avviene tra particella virale e cellula. Oltre a ciò, i carboidrati rappresentano anche substrati per reazioni enzimatiche che causano la loro trasformazione, come le idrolisi di oligo- e polisaccaridi e coniugazione a proteine, processi alla base di fondamentali eventi quali la digestione o il corretto ripiegamento delle proteine. A tali processi biologici sono inevitabilmente associate numerose patologie, aspetto

che rende ancora più importante il ruolo dei carboidrati. Tutto ciò ha spinto la ricerca scientifica verso lo sviluppo di nuove molecole, glicomimetici o glicoderivati, in grado di mimare la funzione bioattiva di queste macromolecole. Inoltre, la generazione di tali derivati permette di superare alcune intrinseche lacune dei carboidrati in quanto tali, come le scarse proprietà *drug-like*. I carboidrati possono anche essere utilizzati come scheletro per il design e la generazione di potenziali farmaci, potendo sfruttare le numerose funzionalità ossidriliche per la coniugazione di opportuni farmacofori, avvalendosi inoltre delle diverse combinazioni stereochimiche dei monosaccaridi. Questi aspetti hanno portato, nel corso dei decenni, ad un incremento nello studio del ruolo di queste macromolecole nella vita; la glicobiologia è ormai una disciplina il cui studio risulta cruciale nella comprensione degli eventi cellulari, fisiologici e patologici: la presenza in commercio di farmaci a base saccaridica è ormai una realtà consolidata, così come la generazione di innumerevoli glicoderivati e glicoconiugati come strumenti di studio dei processi in cui i carboidrati sono coinvolti.

Le potenzialità sopra menzionati rappresentano la base del lavoro svolto durante il dottorato di ricerca, esposto in questo manoscritto. Abbiamo posto la nostra attenzione verso l'utilizzo dei carboidrati per la generazione e la sintesi di molecole bioattive e potenziali NCE (nuove entità chimiche) e lo sviluppo di glicoderivati e coniugati come *tools* per lo studio di importanti processi biologici.

Una parte del lavoro riguarda la generazione di potenziali molecole antiinfiammatorie a base saccaridica, ligandi di un trasportatore localizzato a livello intestinale, SGLT1 (*Sodium-Glucose co-Transporter 1*, co-trasportatore sodio glucosio 1), per il quale recentemente un importante ruolo immunologico nella protezione da eventi infiammatori (sepsi, shock endotossico, infiammazioni croniche e mucositi) è stato

individuato. Inizialmente, diversi lavori di letteratura hanno dimostrato che alte concentrazioni di glucosio *in vitro* ed *in vivo* causano ad una sorta di attivazione di questo trasportatore intestinale, e ciò porta ad una serie di eventi intracellulari che si traducono nel blocco della produzione di molecole pro-infiammatorie, quali citochine e chemochine. La nostra successiva ricerca di analoghi di glucosio non metabolizzabili, in grado di agire a concentrazioni farmacologiche, ha portato alla generazione di un primo *lead compound*, di natura C-glicosidica (Figura 1). Negli ultimi anni, al fine di ampliare la nostra conoscenza sul meccanismo di azione - ancora non completamente noto - di questa molecola è stata inizialmente sviluppata una piccola libreria di analoghi (Cap. 2), per studi di relazione struttura attività, e in secondo luogo è stata messa a punto una via di sintesi per la generazione di un derivato radiomarcato del *lead compound* (Cap. 3), per studiarne il destino fisiologico.

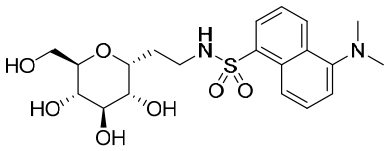


Figura 1 - Struttura del C-glicoside con potente attività antiinfiammatoria.

Ulteriori lavori scientifici hanno dimostrato come sia SGLT1 sia un co-trasportatore sodio-glutammina (B^0AT1), localizzato anch'esso sulla membrana delle cellule epiteliali intestinali e il cui ligando naturale, la L-glutammina, mostra attività antiinfiammatoria, svolgano la loro funzione sotto forma di oligomeri o organizzandosi sulla membrana come clusters. Al fine di sfruttare tale comportamento e ipotizzando una possibile azione antinfiammatoria sinergica tra il C-glicoside e L-glutammina, sono state generate nanoparticelle d'oro decorate con derivati di tali molecole (Figura 2), quindi strutture polivalenti, come strumenti di indagine e

studio di possibili fenomeni di multivalenza associati a tali trasportatori (Cap. 4).

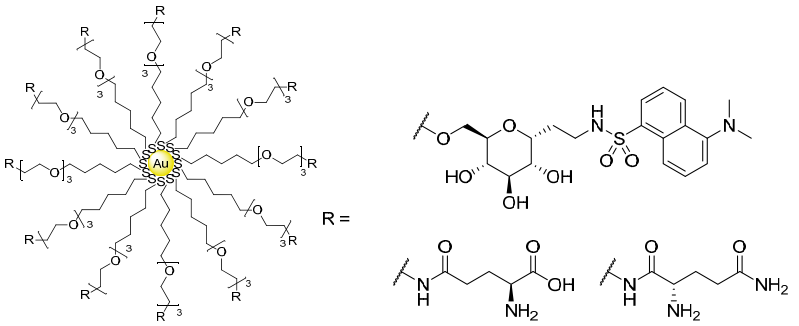


Figura 2 - Schema generale delle nanoparticelle d'oro funzionalizzate con derivati del C-glicoside e L-glutammina.

Il concetto di multivalenza è alla base di una seconda parte del lavoro di ricerca, in cui sono stati sintetizzati strutture polivalenti (calix[4]areni) rivestiti con copie multiple di derivati di imminozuccheri, analoghi saccaridici estremamente importanti come potenziali agenti contro malattie di tipo metabolico (diabete e obesità) o causate da disordini nel processamento di glicolipidi e glicoproteine (malattie da accumulo lisosomiale). I target molecolari di tali composti (inibitori o chaperoni chimici), sono enzimi chiamati glicosidasi, diffuse e convolte in moltissimi processi biologici; recentemente, sono state avanzate ipotesi di comportamento multivalente per gli imminozuccheri verso tali enzimi. Sono quindi stati preparati calixareni polidecorati con derivati imminosaccaridici (Figura 3) (Cap. 5), sfruttando approcci di *click chemistry*, al fine di generare strumenti di studio della inibizione multivalente recentemente riscontrata, associata a questi enzimi.

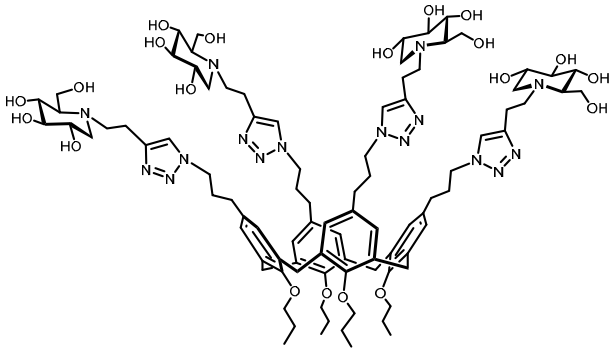


Figura 3 - Struttura di un calix[4]arene polisostituito con derivati di imminozuccheri.

Ad ultimo, abbiamo rivolto la nostra attenzione allo sviluppo di glicoderivati di tipo C-glicosidico, quindi metabolicamente inerti, modificati con gruppi contenenti un atomo di arsenico, al fine di sviluppare potenziali agenti antiproliferativi e citotossici (Cap. 7). L'ipotesi su cui si fonda questo lavoro è stata quella di coniugare una entità chimica contenente arsenico, elemento noto fin dall'antichità per la sua tossicità ma tuttora usato in clinica come agente antileucemico (come triossido di arsenico) e una unità di glucosio, il cui trasporto e catabolismo è fortemente incrementato nelle cellule tumorali a causa del loro fenotipo glicolitico. Il maggior assorbimento porterebbe ad un accumulo di arsenico intracellulare con conseguente aumento della citotossicità. Il C-glucoside recante arsenico come ditioarsenale (As^{III}) (Figura 4) mostra un promettente effetto citotossico verso una linea cellulare di neuroblastoma, e rappresenta un lead compond per la generazione di agenti antitumorali.

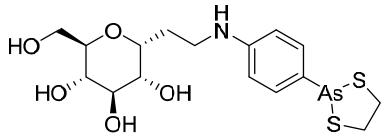


Figura 4 - Struttura dell'arseno(III)-C-glucoside con promettente attività antiproliferativa verso la linea cellulare tumorale di neuroblastoma.

Chapter 1. Introduction

Carbohydrates – general overview

Carbohydrates are defined currently as polyhydroxyaldehydes and polyhydroxyketones, or compounds that through acidic hydrolysis can generate these substances[1]. Among biomacromolecules, carbohydrates are the most complex and diverse class of biopolymers, compared to nucleic acid and proteins. A wide array of available monosaccharides, building blocks of more complex oligo- and polysaccharides, as well as the different stereochemical connections between carbohydrates result in a huge complexity. Moreover, the chain length of the oligosaccharides can also vary widely from monosaccharides up to branched oligosaccharides with more than thirty building blocks, or in the case of polysaccharides to several thousand building blocks[2]. The nine more common monosaccharides which can be found in mammalian cells (Figure 1[2]) diverge for their structural and stereochemical diversity, and can be combined to generate a huge number of linear or branched structures, with a higher grade of diversity than the structures constituted by nucleotides or aminoacids.

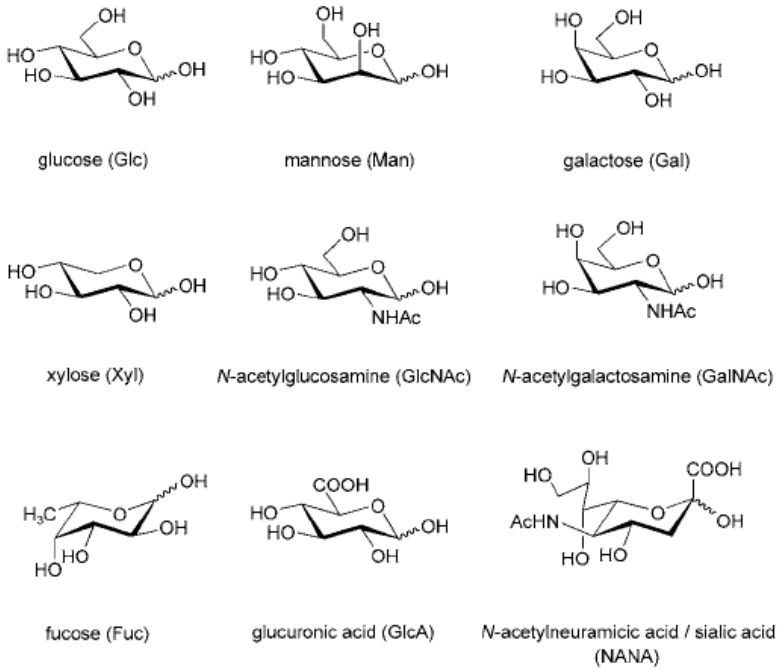


Figure 1 - Most common mammalian monosaccharides.

For their variable and complex nature, carbohydrates show different roles in living organisms. The major part of the carbohydrates exists in Nature as polysaccharides, constituted by a high number of monosaccharides connected with glycosidic bridges. The types of monosaccharidic entities, the chain length, the nature of the glycosidic bond and the branch degree allow to have polysaccharides with a high grade of diversity.

These macromolecules exert in Nature several roles. Some of them are used as tool to store chemical energy, and are involved in energetic metabolism, some others hold a structural function. The principal storage polysaccharides are starch, present in plant cells, and glycogen, in animals. Starch is constituted by D-glucose linked together by α-1,4-*O*-

glycosidic bonds, which form two types of chain, amylose (a linear polymer) and amylopectin (a branched D-glucose polymer with both α-1,4 and α-1,6-*O*-glycosidic bonds). Starch is the principal source of sugar of human diet. As amylopectin, glycogen is a branched polymer of D-glucose, and is the principal storage polysaccharide in animals. Cellulose, in which D-glucose is linked by β-1,4-*O*-glycosidic bonds, is a polysaccharide with a structural function: is the constituent of plant cell wall.

Oligosaccharides are constituted by short chains of different sugar monomers. Oligosaccharides are often present in Nature not as free entities but connected with other biomacromolecules, constituting hybrid structures, called glycoconjugates. Glycoproteins, glycolipds, peptidoglycans and lipopolisaccharides belong to this class of biomolecules. Glycoconjugates, mainly located on the cell surface, exert extremely important functions in living organisms, since they are involved in many biological events and processes, like inflammation, cell-cell recognition, immunological response and metastasis formation (Figure 2[3])[2].

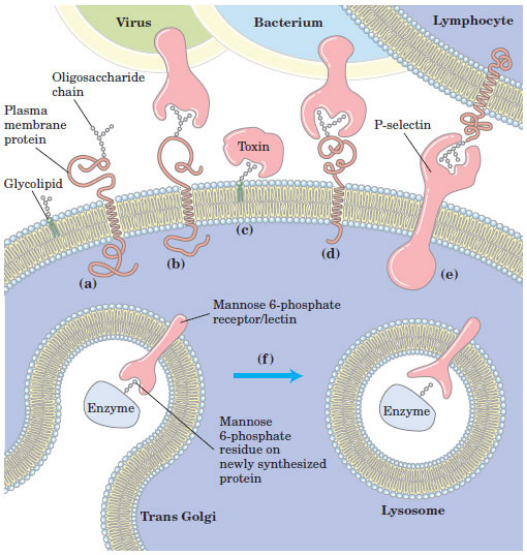


Figure 2 - Role of oligosaccharides in adhesion and cell recognition.

Glycoproteins. Secreted proteins or proteins bound to cell membrane are often glycosylated. The covalent attachment of an oligo- or monosaccharide on a protein occurs on two possible sites: on the amidic group of an asparagine (Asn) residue, in the case of *N*-glycosilation, or on the hydroxyl group of serine or threonine sites, for *O*-glycosilated proteins (Figure 3[3]).

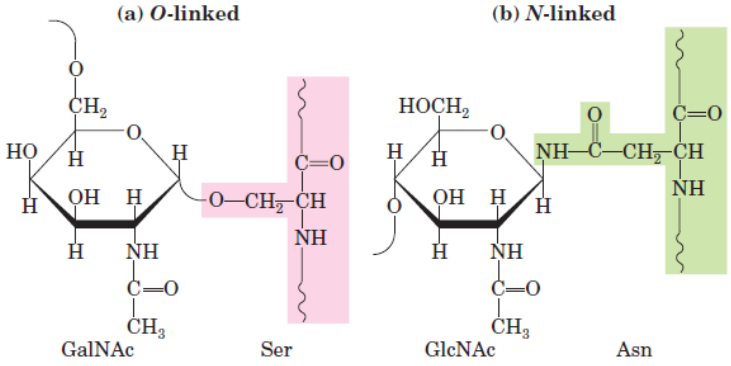


Figure 3 - Oligosaccharide linkages in glycoproteins.

The glycosylation is a finely-regulated cellular event, which involves many different carbohydrate-processing enzymes, located mainly in *Endoplasmic Reticulum* (ER) and Golgi apparatus of eukaryotic cells.
N-linked oligosaccharides are different among proteins, but their biosynthesis is common in the first steps. An oligosaccharidic core made of fourteen monosaccharide residues is synthesized on a isoprenoid carrier molecule called dolichol phosphate, which transfers the oligosaccharide inside the ER lumen, allowing its conjugation on a Asn residue of a putative peptide. Once in the ER, the 14-mer undergoes the so called "trimming" mediated by processing enzymes, which modify the carbohydrate structure to build different oligosaccharides depending on the final destination of the protein ("targeting"). In Golgi apparatus, the *O*-glycosilation on Ser or Thr residues, mediated by other glycosidases and glycosyltransferases, occurs. The conjugation of oligosaccharides on proteins in ER is related to another important biological function: the protein folding quality control (QC) (Figure 4)[4]. The ER environment allows proteins to acquire the correct tridimensional structure by the action, in concert, of several proteins; this process leads to the elimination of unfolded peptides while the proteins with a correct structure can continue their path in Golgi apparatus. The proteins involved in the QC are glycosidases and glycosyltransferases, that add or remove saccharidic residues of the *N*-linked oligosaccharide, and lectins, which are carbohydrate-binding proteins, able to recognize if a protein is correctly folded or not.

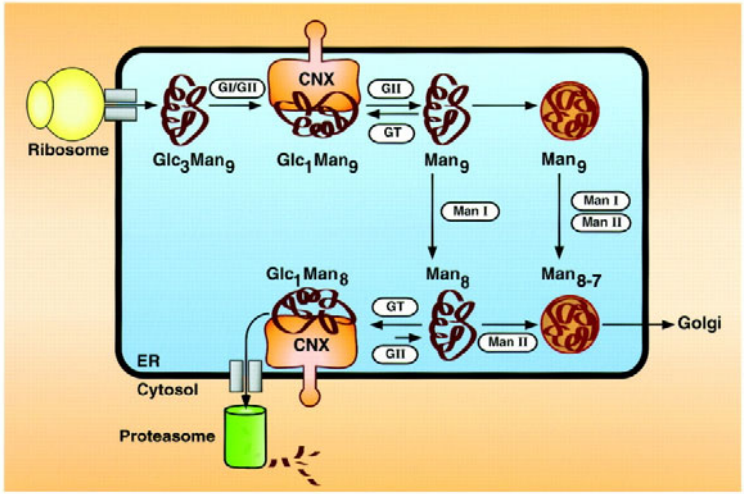


Figure 4 - Scheme of ER quality control.

The oligosaccharide structure has an important function not only for the biological activity, the protein folding and the localization of the peptide to which it is linked, but is also crucial for adhesion phenomena. The correct construction of the oligosaccharide by the cell machinery is therefore fundamental for biological processes such as viral infections, inflammation, immune response and metastasis formation. Moreover, the assembly of the oligosaccharides for protein glycosylation is an accurate process, finely regulated and a malfunctioning of even one of the enzymatic activities involved in the process can cause severe pathologies[5].

Although the biological function of protein glycosylation is not still completely clear, it is known that oligosaccharides constituting glycoprotein, play a pivotal role in immune response, viral replication and infections, cell-cell adhesion and inflammation processes[6]. Furthermore, alterations in oligosaccharide sequence and structure of glycoproteins are associated with several pathological states, as

rheumatoid arthritis and cancer; in fact, changes in oligosaccharidic chain on the surface of cells can represent the molecular bases of the abnormal behavior of tumor cells, allowing the tissue invasion and metastasis formation[7].

The biological importance of glycoproteins is also proven by the their role in the inflammation process, in which the extravasation of leukocytes through body tissues is mediated by the interaction between a complex oligosaccharide located on the surface of this immune cells, called Sialyl-Lewis X, and some proteins able to bind this saccharidic structure (Lectins), as P- and E-selectins, localized on the membrane of endothelial cells[8] (Figure 5[9]). In healthy states, this process is finely regulated and correctly carried out, but when it is out of control, it can trigger excessive and chronic inflammation states, which generate severe pathologies, such as arthritis or endotoxic and septic shocks.

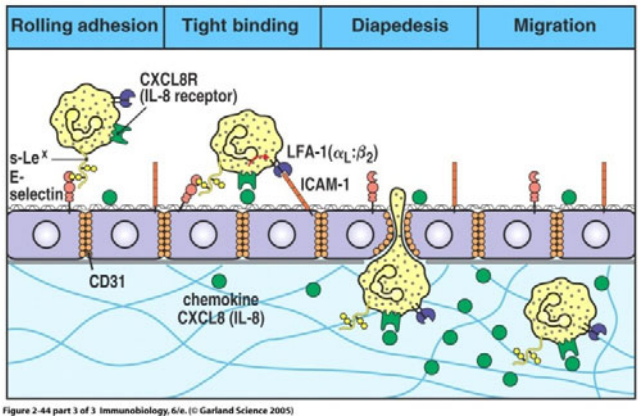


Figure 5 - Extravasation process of leukocytes mediated by carbohydrates-lectines interactions.

Selectins belong to a big class of proteins called lectins, able to interact specifically with oligosaccharides located on the surface of almost all eukaryotic cells; the specific interaction cause several intra- and

extracellular events, and for these reasons lectins are involved in many biological events, like cell adhesion, immune response and apoptosis.

Beside the simpler glycoproteins described above, it is possible to find more complex structures, made of longer and bigger polysaccharides, anchored to proteins, on the surface of cells. These conjugates are called proteoglycans (PGs), and are made of a core protein, and several polysaccharidic chains *O*- and/or *N*-linked. The polysaccharides constituting proteoglycans are called glycosaminoglycans (GAGs). GAGs are repetitions of a disaccharide unit, made of a hexouronic acid and a hexosamine. Based on the backbone structure, it is possible to divide GAGs into four classes: HSGAGs (heparin/heparan), CSGAGs (chondroitin/dermatan), keratan and hyaluronic acid. The biosynthesis of these polysaccharides takes place in Golgi apparatus (for HSGAGs) where even the *O*- or *N*-glycosilation of core proteins of PGs occurs, or directly by an integral plasma membrane proteins, for hyaluronic acid. Once the polysaccharide structure is prepared, many successive modifications on the single monomers constituting the chain occur: events of epimerization, N-deacetylation, *O*- and *N*-sulphation allow to have a high particular and modified structure, which is crucial for the biological activity associated to these biomolecules. The extensive presence of these glycoconjugates at the cell-extracellular matrix (ECM) interface is a proof of their crucial biological function. GAGs exert regulatory functions in development, angiogenesis, cancer progression, microbial pathogenesis and is are key factors in anticoagulation process[10].

Glycolipids. Oligosaccharides can be found also linked to lipids; among glycolipids, the most important are glycosphingolipids, in which an oligosaccharide (or a monosaccharide) is connected with a hydrophobic

part, called ceramide, constituted by a long-chain aminoalcohol (sphingosine) and a fatty acid chain attached on the amino group. Important glycosphingolipids are gangliosides, cerebrosides and globosides, in which monosaccharides such as D-glucose, D-galactose or N-Acetyl-D-gluco/galactosamine, or oligosaccharides made of sialic acid are linked to the ceramide. These glycolipids are localized especially in brain and nervous tissue. These glycolipids mediate cellular recognition mechanisms and are exploited by pathogens to enter the cells or to target their toxins; for example, choleric toxin recognizes and binds ganglioside G_{M1} provoking its entrance in the cell. Some other glycosphingolipids are expressed in particular types of tumors, like G_{M2} and G_{D2} gangliosides expressed in breast cancer.
As for glycoproteins, the incorrect trimming of oligosaccharides constituting the glycolipids can cause their accumulation in cellular lysosomes, due to defects of the functioning of lysosomal glycosidases, devoted to the degradation of this class of glycolipids. This represents the principal cause of Lysosomal Storage Diseases (LSDs), severe pathologies in which mutations in the aminoacidic sequences of the carbohydrate-processing enzymes cause defects in the structure and function of the protein; the accumulation of glycosphingolipids (GSLs) in lysosomes lead to important and lethal transformations in the cells. The degradation pathway of GLSs includes diverse enzymatic activities that, if altered, lead to different pathologies, like Fabry, Gaucher of Krabbe disease.

The big number of structural variations which is possible to generate with carbohydrates, considering both the different stereochemical and regiochemical combinations, makes carbohydrates “ideal tools” to share information between cells in living systems. Cells use carbohydrates

located on their surface to transmit and transfer biological information, allowing communication between different actors of immune systems. Carbohydrates of glycolipids or glycoproteins represent epitopes recognized by immune system cells; specific glycosylation patterns located on the membrane of both organism's cells and of pathogens represent the criteria at the basis of the "self" and "non-self" recognition processes mediated by immune system. For example, A and B antigens which represent the blood group determinants are the terminal portion of glycosphingolipids located on the erythrocytes and vascular endothelial cells membrane.

The therapeutic importance of carbohydrates: the development of glycomimetics

The key role of carbohydrates in the life and therefore in the processes at the basis of relevant physiopathological states, has stimulated the scientific community towards the research of molecules and entities with a saccharidic structure that can be used in the treatment of several diseases, in which the role of carbohydrates is crucial. Compounds which mimic the carbohydrate shapes and biological functions are called glycomimetics; these bioactive class of compounds overcomes the drawbacks of carbohydrates as such (easy degradation in physiological conditions, difficult synthesis, preparation and purification, low activity and drug-like properties)[11]. On the contrary, the advantage in using carbohydrate scaffolds for the design of potential drugs is due to the numerous hydroxyl functionalities, that can be exploited for the conjugation of suitable pharmacophores, taking also advantage of the different stereochemical combinations.

Examples of carbohydrate-based drugs present in the market are low molecular weight derivatives of heparin, polysaccharide mainly involved in the interaction with antythrombin III; their anticoagulant function allow the usage in the treatment of heart pathologies[2]. These derivatives are oligosaccharides constituted by units of GAGs in the form of sodium sulphates (Figure 6)[11].

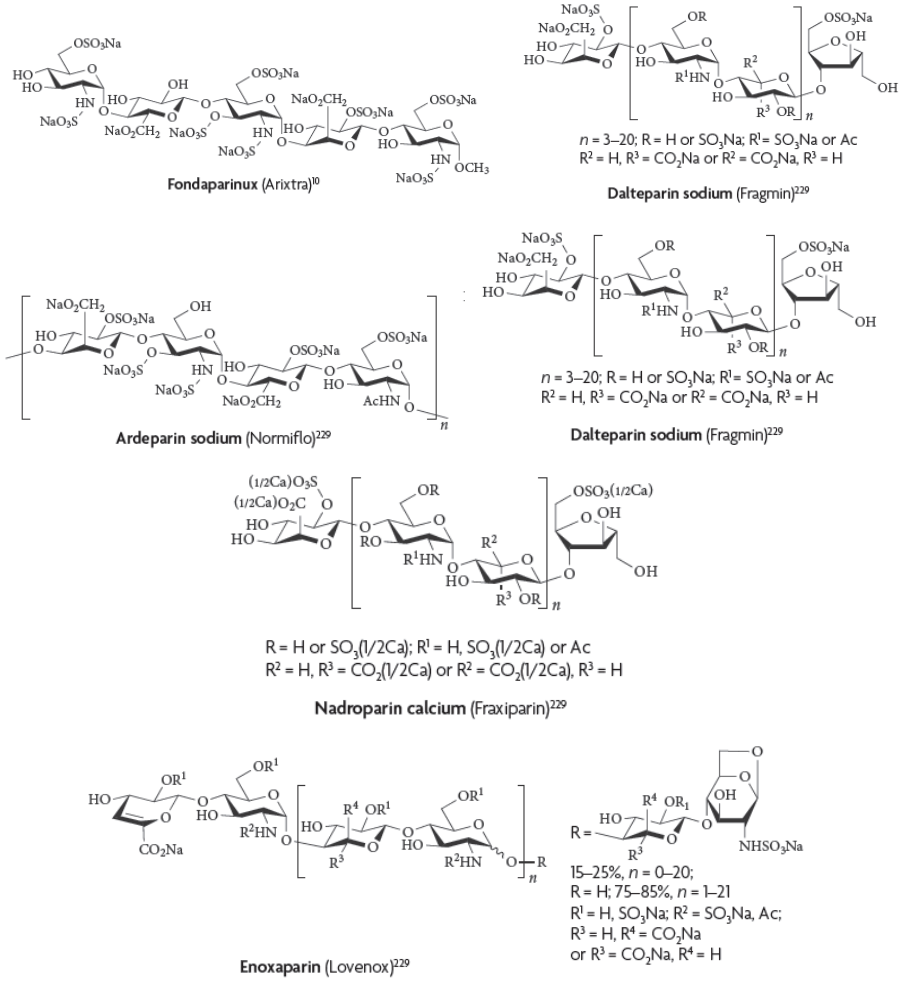


Figure 6 - Low molecular weight derivatives of heparin currently in the market[11].

Another class of glycomimetics is represented by inhibitors of digestive α-glycosidases located on the surface of intestinal epithelium, involved in the digestion of poly- and oligosaccharides, and of glycosidases involved in the intracellular trimming and targeting of glycolipids and glycoproteins. Inhibition of digestive α-glycosidases has an important role in the treatment of diabetes (as for Voglibose, Miglitol and Acarbose), while the ability of some "inhibitors" to bind to the enzyme active site is transformed into a chaperon activity able to rescue misfolded enzymes in LDSs pathologies such as Gaucher's disease (Miglustat), a pathology related to a lysosomal accumulation of glycosylceramides, caused by a defective form of the β-glucosidase enzyme (Figure 7[11]).

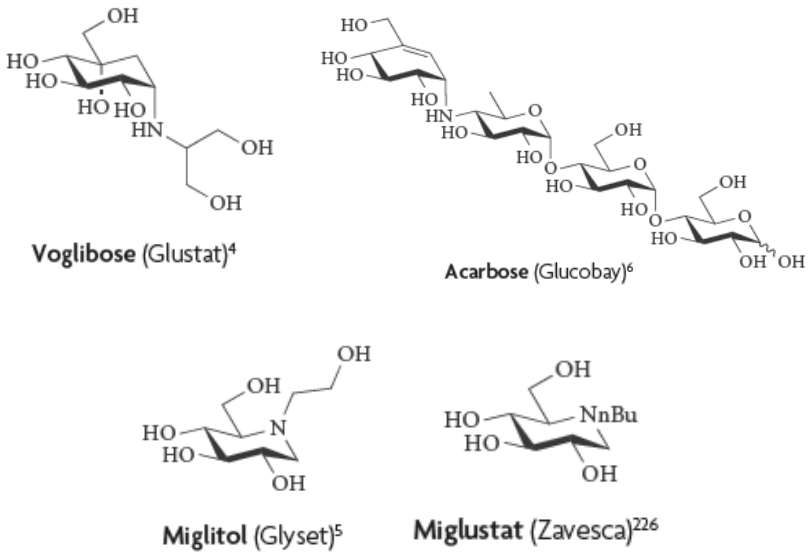


Figure 7 - Some commercially available glycosidases inhibitors for diabetes and Gaucher's disease.

Other carbohydrate-based drugs available in the market are inhibitors of viral glycosidases (sialidases), such as Zanamivir (Relenza®) and

Oseltamivir (Tamiflu®), respectively by GlaxoSmithKline and Roche, used as antivirals for influenza treatment. These drugs are sialidase or neuraminidase inhibitors, viruses use these enzymes to hydrolyze units of sialic acid (N-Acetyl Neuramminic Acid, NANA) located on the surface of host cells, in order to facilitate the release of new viral nanoparticles. Zanamivir mimics the transition state of the hydrolytic process and, as consequence, its structure is very similar to that of sialic acid. Oseltamivir represents the attempt to make more bioavailable and more stable this inhibitor, eliminating polar and metabolically susceptible groups (Figure 8[12]).

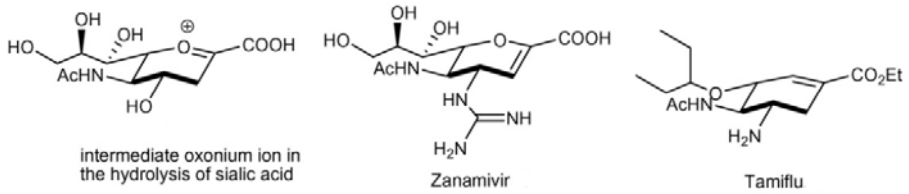


Figure 8 - The oxocarbenium ion intermediate of the neuraminidase reaction and two anti-flu drugs Zanamivir and Oseltamivir.

Topiramate (Topamax®) and Sodium Hyaluronate (Orthovisc®) represent carbohydrate-base drugs acting against epilepsy and osteoarthritis, respectively (Figure 9[12]).

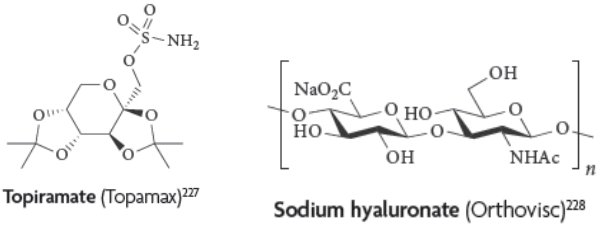


Figure 9 - Structure of Topiramate and Sodium Hyaluronate.

C-glycosides as glycomimetic scaffolds

The research of novel carbohydrate-based drugs has prompted the attention towards the development of different saccharidic scaffolds, with drug-like properties, that can be suitably functionalized, decorated and modified in order to generate bioactive compounds. There is a high necessity and interest in the synthesis of new glycoderivatives, due to an explosion of the glycobiology field in the last decades, potentially useful for biological, biochemical and pharmacological studies. The field of organic synthesis on carbohydrates is focused since long time on new synthetic methods, and in particular on new approaches for the formation of the glycosidic bond[13]. Despite the numerous elegant strategies developed for the efficient formation of glycosidic linkages, the stereoselective synthesis of α- and β-glycosides remains a challenge. Beside the more famous and well-known *O*-glycosides, a special attention has been directed on *C*-glycosides, whose definition is well summarized in Figure 10[14]:

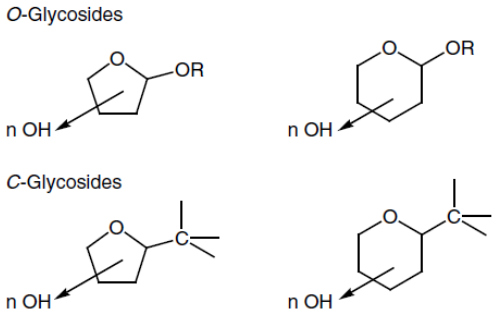


Figure 10 - C-glycoside definition.

C-glycosides are important biological active compounds. Their significance is due to the intrinsic metabolic stability, making them ideal structures for drug candidates development[13]. Synthetic strategies and

approaches for the construction of C-glycosides are fewer than for the *O*-glycosidic bond; however, current methods provide the preparation of carbon-carbon bond *via* generation of nucleophile, electrophile or radical species at the anomeric carbon of the sugar ring[13].

By a structural point of view, *C*-glycosides do not differ substantially from *O*-glycosides; the substitution of exocyclic oxygen of a glycosidic bond with a carbon atom has a marked effect on chemical character and on stability, summarized and listed in Table 1[14]:

Table 1 – Comparison between physical and chemical properties of O- and C-glycosides.

	***O*-Glycosides**	***C*-Glycosides**
Bond length	O–C = 1.43 Å	O–C = 1.54 Å
Van der Waals radius	O = 1.52 Å	O = 2.0 Å
Electronegativity	O = 3.51	C = 2.35
Dipole moment	C–O = 0.74 D	C–C = 0.3 D
Bond rotational barrier	CH_3–O–CH_3 = 2.7 kcal/mol	CH_3–CH_3 = 2.88 kcal/mol
H-Bonding	Two	None
Anomeric effect	Yes	No
Exoanomeric effect	Yes	No
Stability	Cleaved by acids and enzymes	Stable to acids and enzymes
Conformation	$C_{1'}$–$C_{2'}$ antiperiplanar to O_1–C_1	$C_{1'}$–$C_{2'}$ antiperiplanar to C_1–C_2

Some physical features (bond length, Van Der Waals radius, bond rotational barrier) are quite similar between the two class of glycosides. The most important difference is associated to the chemical reactivity: *C*-glycosides, do not present anomeric effect which, on the contrary, characterize *O*-glycosides, and furthermore *C*-glycosides present a better chemical stability both in acidic environment and towards the action of enzymes, like glycosidases.

These reasons make *C*-glycosides an interesting carbohydrate class, especially by a chemical point of view, although these molecules are abundantly present in Nature. It is the case of the *C*-glucosidic derivative of resorcinol or of anthraquinone (Figure 11[14-15]).

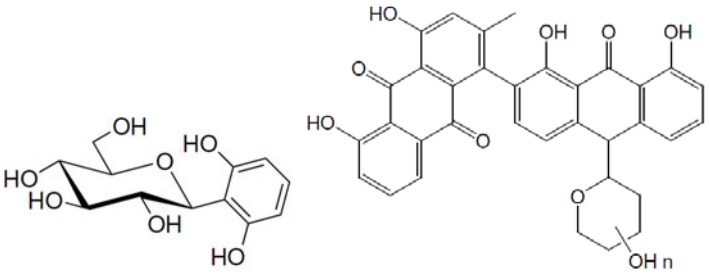


Figure 11 - Natural C-glycosides: on the left, the resorcinol (1,3-dihydroxybenzene) derivative, on the right a C-glycoside derivative of anthraquinone.

Other *C*-glycosides found in Nature possess particular and important biological activities: examples are pirazomycin, an antiviral, and showdomycin, an antibacterial and antitumor drug (Figure 12[14]).

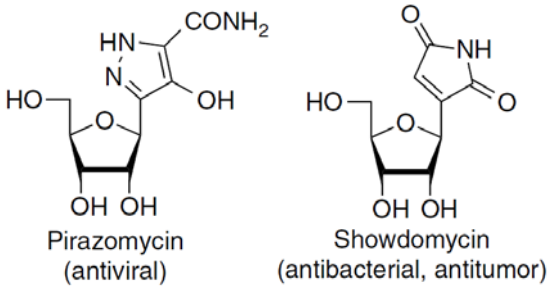


Figure 12 - Natural C-glycosides with biological activity.

The discovery of natural *C*-glycosides with associated biological activity and their intrinsic stability has prompted to a high interest towards the synthesis of new analogues of this class of carbohydrates, able to interact and inhibit specific enzymes, to act as agonist or antagonist in receptor-ligands interaction phenomena and to induce an immunogenic response[14]. Among them, we find Urdamycinone B, a carbohydrate-anthraquinone hybrid, able to bind to DNA (and thus cytotoxic), *C*-glycosides inhibiting specifically some glycosidases, and glycomimetics of oligosaccharide Sialil-Lewis X, whose interaction with Selectins mediates

the starting point of inflammation processes and the generation of tumor metastasis (Figure 13[14]).

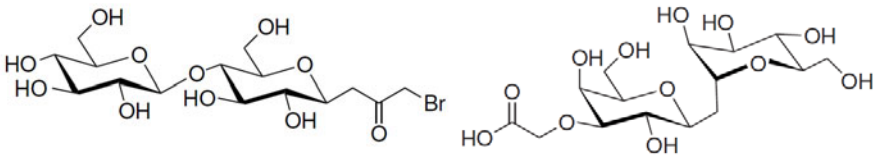
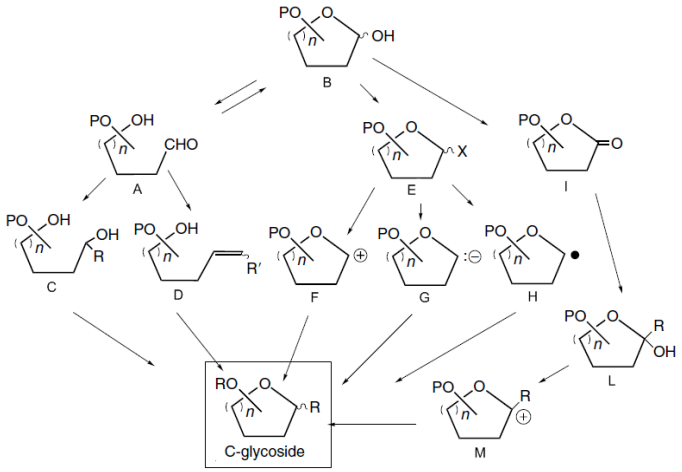


Figure 13 - On the left, a synthetic C-glycoside, irreversible inhibitors of glycosidases; on the right, the C-disaccharide mimetic of Sialil-Lewis X.

Nowadays, some procedures and methods for the generation of *C*-glycosides exist; Scheme 1 summarizes the strategies and synthons used for the preparation of desired *C*-glycosides, starting from aldoses or ketoses. The carbonyl group of the open monosaccharide (A), which is in equilibrium with the emiacetal/emiketal form (B), can undergo transformation by species bearing nucleophilic carbon atom, or alternatively to olefination reactions. In the first case, an open reduced form of the sugar is obtained (C); after, this can be cyclized with the loss of a water molecule, to get the final C-glycoside. In the second option, the olefination step generates an structure bearing an activated double bond (i.e. a double bond in α-position to an electron withdrawing group), which is appropriate for a conjugate 1,4 addition by the free OH group of D. The emiacetal/emiketal form can be converted into a *C*-glycoside by two strategies. The anomeric position can be activated with the conversion of its hydroxyl group in a better leaving group (an halogen, an imidate of a thioderivative), to generate cationic (oxocarbenium ion, F) , anionic (G) or radical (H) intermediates, depending on the nature of the leaving group. The reaction of these intermediates with suitable species bearing nucleophile, electrophile or radical carbon atoms, will generate the C-glycoside. The activation of the anomeric position can be also achieved by using chemical reagents such as Lewis acids, which cause

the direct activation of the hydroxyl anomeric group, to give the oxonium intermediate F. The last strategy is represented by the oxidation of the anomeric portion to the respective lactone I, which reacting with organometallic species is transformed in the lactol L, easily transformed in the final C-glycosides, via the oxocarbenium ion intermediate M, with reducing reactants.



Scheme 1 – General scheme for C-glycoside synthesis.

Role of carbohydrates and glycomimetics in anti-inflammatory processes: inflammation and SGLT1

Inflammation, sepsis and septic shock

Sepsis describe a complex clinical syndrome, caused by a harmful response that a host activates as a result of an infection. This response is characterized by a hyperactivation and deregulation of the mechanisms which lead the immune system to react against an invasion caused by an external pathogen[16]. From a clinical point of view, sepsis appears with fever, mental confusion, transient hypotension, thrombocytopenia and diminished urine output. If not treated, the patient can undergo to irreversible damages to the respiratory and renal systems, abnormal blood coagulation and deep hypotension. One of the main features of this pathology is the high rate of death, about 30% of the patients, and is close to 50% in patients which possess the more severe and acute septic shock. Sepsis is defined as the systemic response to an infections[17], and should be distinct for other pathological states, as:

- SIRS (Systemic Inflammatory Response Syndrome), which involves the whole body; non necessary it derives from a microbial infections, but presents similar syndromes to sepsis (body temperature lower than 36°C or higher 38°C, heart frequency of 90 bpm, increase respiratory frequency and profound alteration of blood white cells);
- Bacteremia and septicemia: presence of bacteria in the blood.
- Sepsis and septic shock: are the most severe forms of sepsis, characterized by a accentuated hypotension, vascular hypoperfusion, and several body organ failures.
- Multi-organ failure: the most critical state, with seriously altered body functions which do not permit an internal homeostasis maintenance without external actions[17].

The most common sites of infection that lead to such types of response are the respiratory apparatus, the abdominal cavity, the urinary tract and blood. The diagnosis of microbiological type characterized a half of the patient cases and in particular is due to Gram negative infections.

For this reason, the study of the structural components of bacteria which cause sepsis in an important scientific field, to understand the mechanism responsible of the disease and to identify potential therapeutic targets. These components, recognized by immune systems, belong in the so called Pathogen-Associated Molecular Patterns (PAMPs)[16]. In Gram negative bacteria, Lipopolysaccharides (LPSs) have a predominant role. The external membrane of Gram negative bacteria is structured in an asymmetric manner, and consist of phospholipids in the internal section, and of glycolipid anchors of LPS in the external side (Figure 14[18] [19]).

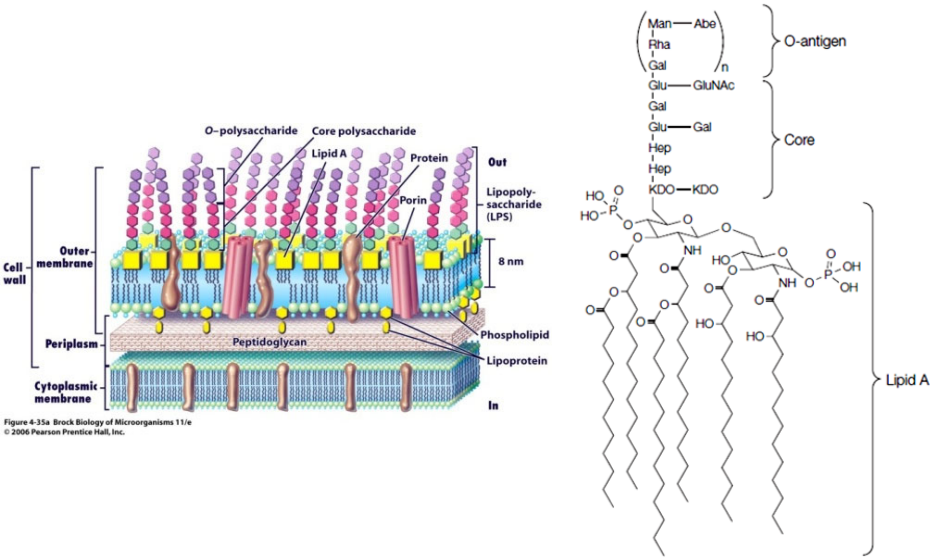


Figure 14 - Structure of Gram negative bacteria envelope (left) and LPS structure (right). Bacteria LPS are constituted by a glycolipids portion, called Lipid A, an oligosaccharidic core and a more external section, the O-antigen.

In Gram positive bacteria, the structural determinants for sepsis and in general inflammation states are structures contained in the cell wall, like peptidoglycans and lipoteichoic acid. Molecules causing inflammation states which remains anchored to the cell are called endotoxins, while compounds that are produced and secreted on the external side of cells are exotoxins, produced especially by Gram positive bacteria. These type of toxins have a great interest because are able to bind the Class II Major Histocompatibility Complex (MHC-II) and to some T-lymphocyte receptors, triggering a massive activation of these cells and a strong release of pro-inflammatory cytokines[20]. Not only bacterial glycoproteins or glycolipids show pro-inflammatory activities: flagellin, a protein located on the cell wall of flagellated bacteria, or non-methylated CpG sequence of naked bacterial DNA, are able to bind specific host receptors, belonging the Toll-like receptor family (TLR), as in the case of LPSs. TLRs play a key role in the recognition of structures of microbial or pathogenic origin; the interaction between these receptors and the respective pro-inflammatory ligands is at the basis both of the recognition of pathogens which come in contact with host defense barriers, to produce an adequate immune response ("non-self"), and of the recognition of all commensal bacteria agents which constitute the intestinal flora and belong to the "self"(Figure 15)[19].

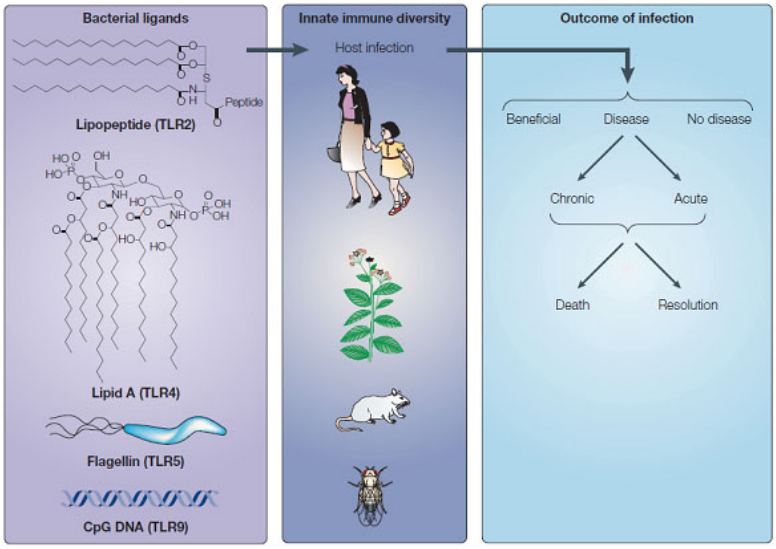


Figure 15 - The complexity of interaction between bacteria determinants and immune system.

TLRs family is constituted by several members, each one responsible for the recognition of diverse bacterial, viruses and fungi structural elements. As depicted in Figure 15, the receptor involved in LPSs recognition is TLR4; TLR2 recognizes structure of Gram positive cells, TLR5 is the receptor for flagellin and TLR9 binds CpG elements of bacteria DNA[21]. In these years, several studies allowed a deep comprehension of the mechanism at the basis of these specific inflammatory responses, and nowadays it is clear that the process is to be ascribed not only to TLRs but to a more extended protein complex, which can be defined as "LPS Receptor". Several protein actors are involved, as LPS binding protein, that binds LPS and presents it to other proteins located on the cell membrane: TLR4, CD14 and MD2. The association of all these proteins with LPS forms a protein complex which allows the activation of signal cascades in the cell and the beginning of

the inflammatory response. Cells which express this complex are members of immune system, like macrophages, dendritic cells but also intestinal epithelial cells (IECs), which are always involved in the external environment recognition. The activation of a particular signal cascade transduction mediated by several intracellular proteins leads definitively to the nuclear translocation of the transcriptional factor NFkB, event that increases the expression of pro-inflammatory cytokines, like IL-1 or TNFα[16]. These cytokines are the principal pro-inflammatory molecules secreted after the signal cascade transduction, released within 30-90 minutes after LPS exposure. Successively, they mediate a second level of inflammatory cascade, in which new cytokines, lipid mediators and Reactive Oxygen Species (ROS) are synthesized. The intense pro-inflammatory response determining the sepsis state is balanced by a complex of counter-regulatory molecules which maintain an immunological equilibrium[16]. These anti-inflammatory modulators are antagonist like the TNF soluble receptor, some complement inactivator proteins or anti-inflammatory cytokines like IL-10.

Role of intestinal mucosal cells. Intestinal mucosa represents a body district where several function are intertwined: digestion of food ingested with the diet and substances absorption are the activities which characterize this organ, but a fundamental role is represented by the first real barrier that the body has against external environment. Intestinal epithelium is not only a physical barrier against pathogens, but plays actively in immune and inflammatory response[22]. Intestinal mucosa has to recognize immediately eventual pathogenic threats present in the intestinal lumen, and, as a consequence, has to control the immune response against these agents. But intestine has to maintain also the hyporesponsivity towards the pool of commensal

microorganisms living in the intestinal lumen. Intestinal epithelial cells are able to discriminate between the "non-self" (the pathogens) and "self" (the own bacterial flora). For intestinal cells, as well as for immune system actors, the recognition of pathogen structure (PAMPs) is mediated by the Pattern Recognition Receptors (PRRs), in particular by TLRs. For this reasons, it is possible to define the intestinal epithelial cells as "sensors" of external environment, and, if activated, can address the development of the inflammatory response, not only recruiting the competent immune cells, but directly influencing their action[22]. TLRs are widely expressed in different cellular types of gastrointestinal mucosa (stomach epithelial cells, small intestine, colon, macrophages, monocytes, intestinal dendritic cells). The discrimination between self and non-self, realized by TLRs, probably is due to expression level of these proteins on the membrane of IECs: in healthy tissues, a low expression level of TLR2 and 4 is present, in order to minimize the intestinal commensal bacteria recognition. On the contrary, TLR4 is highly expressed in the IECs when pathological condition occurs, determined by chronic inflammation states, like Chron disease or ulcerative colitis. Several molecular mechanisms allow the tolerance of IECs towards bacterial flora: a reduced expression of TLRs, an increase level of TLR4 signal transduction cascade suppressors or production of external regulators which suppress the TLR4-mediated pathway. With such mechanism, enterocytes maintain a hyporesponsivity towards commensal bacteria, due to a constant basal level of activation of pro-inflammatory pathways. The action of the different TLRs expressed on the surface of intestinal cells is the same shown by other innate immune system cells; in particular, TLR4 is responsible, also in this apparatus, for the recognition of LPS associated to Gram negative bacteria present in the intestinal lumen. The subcellular localization of the different TLRs, in

particular TLR4, has a key role in the creation of the responsiveness of IECs towards pathogens; in fact, TLRs are strategically localized both on the cellular surface and in intracellular compartments. TLR2 and TLR4 are located mainly on apical membrane of enterocytes (that constitutes the so called brush border membrane, made of numerous evaginations of villi, fundamental to increase the absorption surface of nutrients)[23]. In this manner, IECs can monitor bacteria presence in gut lumen. Other TLRs are localized at intracellular level (as TLR9, which recognize CpG DNA sequences) or on basolateral membrane of enterocytes (the portion of plasma membrane faced towards the internal side of intestinal villus, like TLR5[23]).

When genetic dysfunctions associated with the action of TLRs occur, inflammatory process can become chronic, leading to inflammatory intestinal diseases, called IBDs (Inflammatory Bowel Diseases)[24]. These pathologies are characterized by a constant and deregulated response towards commensal bacteria. The etiological causes are multiple, but genetic predisposition is fundamental[24]. In pathological situations, like IBDs, the surface expression of TLRs is increased, leading to a hypersensitivity of IECs even towards the bacterial flora.

One of the most important effect which is possible to observe in inflammation and sepsis states is the LPS-TLR4 mediated activation of signal transduction cascades that lead to cell apoptosis of enterocytes. Apoptosis is a common and physiological phenomenon in the intestine, and is associated with the high turnover of the IECs, to maintain a correct balance of cell proliferation and death and then to a correct homeostasis. The physiological extrusion of enterocytes from intestinal villi do not compromise the protective function of intestinal mucosa, which is lower in some circumstances (excessive exposure to pathogens and cytotoxic agents)[25]. When this exposure occurs, the apoptosis

phenomena increase, thus leading to a diminished mucosal integrity, with negative consequences on the protective function of this apparatus. Cells which undergo this process suffer from sensible morphological changes, as dimension reduction, mitochondrial swelling, plasma membrane alterations and nuclear fragmentation[26].

Role of glucose in the protection of inflammation events. Recent works[25-26] [22] have demonstrated that high exogenous glucose concentrations, administered *in vitro* to IECs, are able to protect these cells from damages caused by pathogens expressing LPSs. These works have proposed that the mechanism of glucose-mediated cytoprotection depends on an increased glucose uptake, by enterocytes, mediated by Sodium-Glucose Co-Transporter 1 (SGLT1). This protein represents the principal way with which the intestinal epithelium absorbs glucose, and it's localized mainly on the brush border membrane. SGLT1 mediates the unidirectional transport of glucose from intestinal lumen to the internal side of epithelial cell. Further information and details on the structure and function of this transporter are described afterwards.

Previously, it has been mentioned that enteric bacteria can be indirectly considered as responsible for the pathogenesis of the IBDs. However, a high number of cells of *Enterobacteriaceae*, *Proteobacteria* and *Bacteroids* genera are often identified in ileal mucose of patients with Chron disease[26]. An augmented intestinal bacterial load leads inevitably to a higher exposure of intestinal mucosa towards bacterial pathogenic structures, in particular LPSs. The consequent enterocyte-mediated inflammatory response produce a vast damage to the mucose, characterized by apoptotic phenomena and injury on thight juctions, present between mucosa cells, fundamental for the homeostatic equilibrium and immunological protection. In particular, thight junction

damages lead to a higher mucose permeability, since frequently in patients affected by Chron disease, LPS or anti-endotoxin antibodies are found in blood plasma[26]. For the capacity to generate apoptotic event in epithelial cells, LPSs are used in studies in which the cytoprotective function, SGLT1-mediated, of high glucose concentration is evaluated[25-26] [22]. Preliminary studies performed by Buret et al.[26] have allowed to verify the mechanisms which triggers the apoptotic phenomena in IECs, is mediated by the interaction of bacterial LPS and TLR4. The results have demonstrated a participation of the mitochondrial pathway in the apoptotic mechanism, and that a higher glucose uptake can modify the apoptotic level cause by LPS. The increased uptake is due to a high glucose level located on the external of these cells, in which the higher glucose absorption is mediated by SGLT1. The higher SGLT1-mediated glucose uptake provokes profound changes of apoptotic intracellular signaling, thus leading to a cytoprotective effect against LPSs exposure. SGLT1 role was also confirmed testing the cytoprotective ability of glucose on epithelial cells in the presence of phlorizin, a SGLT1 specific inhibitor[27]. This molecule inhibits the glucose uptake of almost 60%, thus suppressing the glucose-mediated cytoprotection. Successive studies[25] allowed to understand that the change of glucose uptake in protective effect phenomena leads to changes in the production of proinflammatory cyto- and chemokines. Furthermore, the higher uptake is directly associated to an increase of specific activity of SGLT1. The augmented transport activity depends exclusively on V_{max} increase and not on the affinity constant for glucose. This means higher expression membrane level for this transporter. The higher expression levels are due to higher translocation level from cytoplasm to the membrane, mediated by cytoskeleton structures, and not due to increasing intracellular concentration of SGLT1. The translocation of SGLT1 on

apical membrane of enterocytes is a finely regulated process, in which growth factors as EGF (epidermal growth factor) and IGF (insulin growth factor) are involved[28] [29].

Further studies were performed by the research group of Prof. C. Rumio[22], regarding the protective effect that SGLT1 activation has on IECs, to which bacterial LPS were administered. In this case the attention was focused to anti-inflammatory and not anti-apoptotic ability of high glucose concentrations, able to activate SGLT1, blocking the intracellular signaling that leads to pro-inflammatory mediators production. The activation of apoptotic pathway requires higher LPS concentrations (> 10 µg/mL) than those necessary for the pro-inflammatory response (from 200 ng/mL to 10 µg/mL)[26]. ELISA test performed on HT29 (human adenocarcinoma cells) insulted with LPS or CpG-ODN (TRL9 ligand), revealed that high glucose concentrations (5 g/L) can reduce the IL-8 production, a marker of inflammatory states (Figure 16[22]).

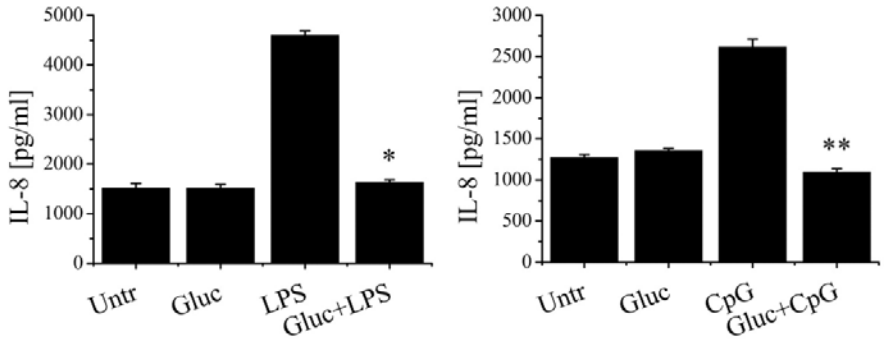


Figure 16 - IL-8 production levels of HT29 cells treated with LPS or CpG-ODN and/or glucose (5 g/L).

To prove the role of SGLT1 in this phenomenon, its action was blocked both silencing its gene (with the siRNA technique) or inhibiting (with phlorizin) the transporter. In both cases, HT29 cells treated with pro-

inflammatory molecules and exposed to glucose, produce IL-8 with level comparable to that secreted in with normal glucose concentrations (Figure 17[22]).

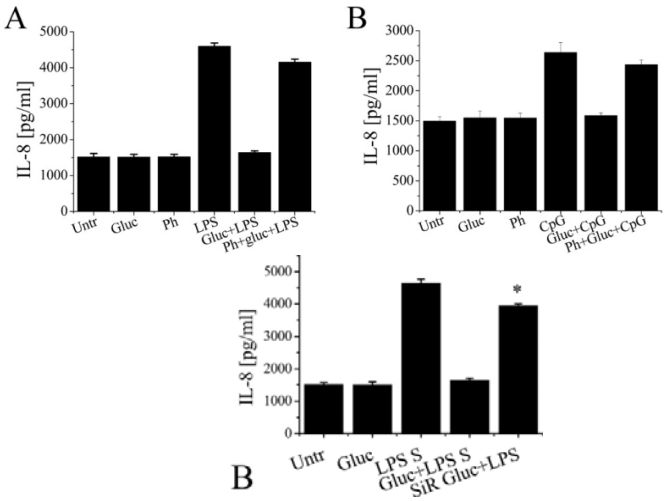


Figure 17 - Levels of IL-8 produced by HT29 cells, treated with LPS or Cpg-ODN, in presence of phlorizin (left) and silencing SGLT1 (right).

Similar experiments and results were obtained using a non-metabolizable glucose derivative, 3-*O*-Me-D-glucopyranose (3OMG), in order to exclude any possible metabolic causes of the higher glucose uptake. The anti-inflammatory effect of glucose was observed also *in vivo*, treating mice orally with 2,5 g/kg of D-glucose or 3OMG, and insulting with LPS or Cpg-ODN. Serum levels of different chemokines were lower in the presence of glucose or its derivative, and mice treated with LPS, D-Galactosamine (murine model of septick shock) and D-glucose showed a complete level of survival, compared to mice which were not treated with high glucose. Also with *in vivo* test the SGLT1 role was proven: only the oral administration of glucose and phlorizin (hence the presence of these two molecules in the gut lumen) led to the same results obtained *in vitro*. With intraperitoneal administration of glucose,

no protective effect was observed, and no blocking of anti-inflammatory response was noticed with phlorizin administration. In conclusion, experimental data clearly indicate that glucose, at high concentration, administered *in vitro* or orally *in vivo* is able to:

- guarantee a cytoprotection on IECs from possible apoptotic events provoked, in patients affected by IBDs, by hypersensitivity to commensal bacterial agents located normally in the gut, which cause is principally due to altered level of expression of TLRs on the enterocyte surface;
- reduce or stop the systemic inflammatory response triggered by LPS or PAMPs which, interacting with TLRs, lead to the expression of genes codifying for pro-inflammatory modulators.

In both cases, a pivotal function is carried out by SGLT1, whose activation due to the high glucose concentration allows the beginning of the protective phenomena against apoptosis, inflammation and septic shock. SGLT1 is therefore not only devoted to the nutrient absorption, and in particular sugars, but it shows also a key role under an immunological point of view, actively participating to processes at the base of the protective function of intestinal epithelium against external environment.

SGLT1: function and structure

Digestion and absorption of carbohydrates assumed with the diet is a process that involves all the gastrointestinal system of mammalian, starting from the mouth, the first body district, where saliva amylases begin the digestion of amylopectin and amylose chains of starch. Then the digestion, stopped in stomach by gastric juices, starts again in duodenum by pancreatic amylases which break down starch chains into maltose and isomaltose, glucose dimers. The last phase of the digestion

is localized at the intestinal level, where maltases secreted by epithelial cells of microvilli hydrolize the dimeric units in single glucose molecules. Other disaccharides, like lactose or sucrose, are hydrolyzed in single monosaccharides (glucose, fructose and galactose) by the action of other glycosidases (β-galactosidase o lactase and saccharase or invertase). Glucose plays a central role in energetic metabolism: an adult man uses about 250 g of glucose every day to perform all the body functions (half from the diet, half from glycogen or gluconeogenesis)[30]. The most difficult task is the maintenance of the sugar concentration in the blood (80 – 120 mg/dL) to ensure a constant and continue supply of glucose for the brain[31]. But surprisingly, glucose is not an essential component of the human diet: this observation is confirmed by cases of individuals intolerant to this monosaccharide, as for patients affected by the glucose-galactose malabsorption disease (GGM), which can live normally without glucose ingestion.

The carbohydrate absorption occurs in the intestine after a long phase of digestion; the polarity of these molecules avoids their free diffusion across the plasma membrane. The absorption of sugars requires the existence of specific transport proteins localized both on the apical membrane of enterocytes, which constitute the brush border membrane (BBM) and on basolateral membrane, structure that separates these cells from blood flux. The two localization reflects the two steps required for sugar absorption: at first, there is an accumulation of sugars in the BBM, then the translocation in the blood occurs[30, 32]. In both processes specific proteins are required; these transporters are the GLUT proteins (gene family SLC2) and the sodium-glucose co-transporters, SGLTs (gene family SLC5). These proteins are located mainly in the intestine, but are also involved in the sugar reuptake in kidneys from glomerular filtrate and in the glucose uptake across the Blood Brain Barrier (BBB)[33]. GLUT

transporters are devoted to the passive translocation of sugars, without energy consumption; examples are GLUT1 (in BBB) and GLUT4 (for insulin-dependent glucose uptake)[30]. SGLTs protein carry out an active transport of sugars across the membrane, i.e. with energy consumption. SGLTs constitute a protein family of 11 members in human, whose genes are identified by the genes SLC5. SGLTs belong to a bigger class of transporters, the SSSFs (Sodium Substrate Symporter Family). These proteins, both in eukaryotic and prokaryotic cells, carry out the transport of different solutes coupling it with the transport, according to their electrochemical gradient, of different ions, such as sodium or protons. The flux of ions is produced by the presence of an electrochemical gradient across the two sides of the membrane (SMF, Sodium Motive Force). This allows the accumulation of energy to pump the solutes inside the cell, acting against their gradient. The sodium gradient is generated by sodium/potassium-ATPase pump (Na^+/K^+-ATPases) that continuously translocates sodium at the external side of the cell, across the basolateral membrane of the cell. For this reason, member of SSSF family are called active secondary transporters because the translocation of solutes is not matched directly with the energy consumption (ATP), as on the contrary happens for Na^+/K^+-ATPase pumps. In the co-transport, the energy necessary for the solute translocation is produced with other processes, in the form of an electrochemical gradient[33]. The direction and the rate of the glucose transport is a function of the direction and intensity of the sodium gradient across the apical membrane[30].

SSSF transporters are grouped in different subfamilies, according to sequence similarity[34]. Over two hundred members of this family exist, both of prokaryotic and eukaryotic origin. From a functional point of view, we can distinguish some subclasses: transporters of sodium-glucose (human SGLTs), sodium-aminoacids (like PutP of *E. coli*, a

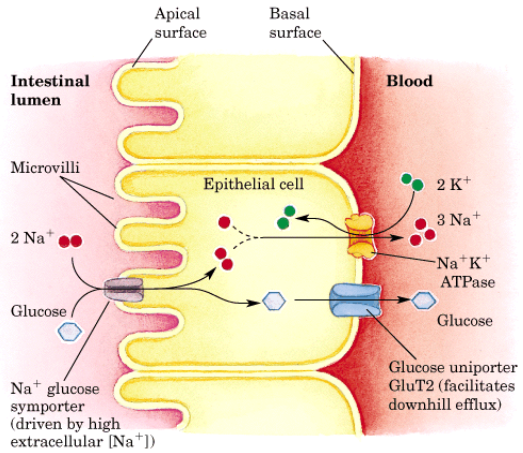
sodium-proline transporter), sodium-vitamins (SMVT of *H. sapiens*) and sodium-ion (NIS, sodium-iodide contransporter of *E. coli*).

In human genome, 11 members of the subfamily of SLC have been identified (Table 2); among them, the most important is certainly SGLT1 (SLC5A1), mainly expressed in BBM of IECs, in a smaller part in proximal tubules of kidneys, where it is involved in the sugars reuptake from the glomerular filtrate, work principally carried out by SGLT2. SGLT1 is the most studied member of this family, because mutations in its gene are associated with a pathology called Glucose-Galactose Malabsorption (GGM)[35].

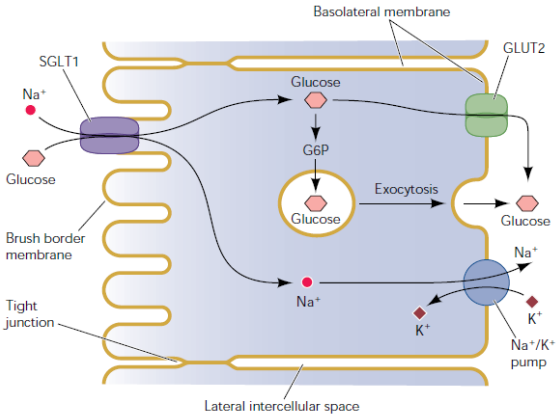
Table 2 - Members of human genes family SLC5.

Human gene	Protein name	Main substrates	Tissue distribution	Gene locus
SLC5A1	SGLT1	Glucose, Galactose	Small intestine>>Kidneys, heart	22q13.1
SLC5A2	SGLT2	Glucose	Kidneys, Heart, Liver, Thyroid, Muscles, Brain	16p12-p11
SLC5A3	SMIT	Myo-inositol>> Glucose	Brain, heart, kidneys e lungs	21q22.12
SLC5A4	SGLT3	Sodium (H^+)	Small intestine, skeletric muscles, kydneys	21q22.2-q12.3
SLC5A5	NIS	I^-(ClO_4^-, SCN^-, NO_3^-, Br^-)	Thyroid, colon, ovaries	19p13.2-p12
SLC5A6	SMVT	Biotin, lipoate, pantothenate	Brain, Heart, Kidneys, Lungs	2p23
SLC5A7	CHT	Choline	Spinal cord	2q12
SLC5A8	SGLT4	Glucose, Mannose	Small intestine, Kidneys, Liver, Lungs and brain	1p32
SLC5A9	SGLT5	?	Kidneys	17p11.2
SLC5A10	SGLT6	Myo-inositol, glucose	Small intestine, Brain, Kidneys, Liver, Heart and Lungs	16p12.1
SLC5A11	AIT	Iodine	Thyroid	12q23.1

SGLT1 and sugar absorption. SGLT1 mediates the glucose (Glc) and galactose (Gal) absorption in the intestine, with their translocation across the brush border membrane of intestinal villi; the required energy is given by the sodium electrochemical gradient generated by the Na^+/K^+-ATPase pump of the basolateral membrane of the same cells. A gradient of sodium between the intestinal lumen (external side) and cytoplasm (internal side) is created and this allows the symport of sodium and glucose or galactose inside the cell (Figure 18a[3]). This process shows a precise stoichiometry; for every molecule of Glc or Gal, two sodium ions are translocated inside the IEC. Experimental evidences indicate that the co-transport seems to be completely reversible, depending exclusively on the direction of the sodium electrochemical gradient and of the monosaccharides[36]. In Figure 18 the successive passage of the two sugars into the blood stream, mediated by GLUT transporters, is depicted. However, some studies suggest an alternative pathway for this second step, in which glucose is translocated across the basolateral membrane by an exocytosis mechanism (Figure 18b[33]).



a)



b)

Figure 18 - Glucose absorption in intestinal epithelial cells.

It is well known that glucose entry, mediated by enterocytes, leads to a high rate of water absorption in the intestine. This is due by the action of SGLT1, responsible for sodium absorption, which inevitably allows the translocation of other anions (Cl^- and HCO_3^-) inside the cell. By this point of view, SGLT1 can be seen as an important vehicle of water absorption at intestinal level[33]. For each glucose or galactose, two sodium and 260 water molecules enter in the cell. This feature is at the basis of an important therapy of the 20th century, the Oral Rehydration Therapy (ORT), fundamental for the treatment of acute diarrhea[37]. The introduction of this therapy have considerably lowered the number of death among children affected by acute diarrhea (60% from 1980 to 2000). ORT is simple, cheap and often associated with cholera therapies, because this pathology leads to a strong loss of liquids, that can be balanced with the oral administration of salts and glucose solutions. In this way, the glucose assumption stimulates the SGLT1-mediated sodium and water uptake by IECs (for each glucose mole, 2 moles of NaCl and 4 – 6 L of water are co-transported)[38].

On the contrary, SGLT1 is also the cause of a pathology called GGM (glucose-galactose malabsorption, OMIM 182380). A defective form of SGLT1, or its absence in the intestinal epithelium, due to mutations in the respective gene, prevents the absorption of the two monosaccharides. Diarrhea is the principal symptom, that ceases when the assumption of the two sugars is interrupted. No defects are present in patients with this pathology in the absorption of fructose, since its transport in the intestine is mediated by GLUT5, localized in the apical membrane of enterocytes[33].

Furthermore, SGLT1 represents a target for type II diabetes, since several studies indicate that its expression levels (together to GLUT5) are three-four fold higher in such patients compared to healthy individuals[39]. This leads inevitably to an increased capacity of glucose and fructose absorption by intestinal epithelial cells. A plethora of scientific works aim to the synthesis and development of SGLT1 inhibitors, able to reduce glucose absorption; in Nature, a SGLT1 inhibitor still exists, phlorizin, a β-glucosides present in a variety of fruits (cherries, apples, service tree). Phlorizin is a member of the class called chalcones, and is constituted by a hydrophobic unit of two aromatic rings, linked to the glucose by a β-glucosidic bond. This molecule has a strong affinity and inhibition constant for SGLT1[40].

Structural features of SGLT1. The structure of SGLT1 has been the subject of numerous studies within the last fifteen years, and the knowledge about it are due in particular to the works of Rolf K. H. Kinne and Ernst M. Wright. The human isoform of SGLT1 (hSGLT1) (UniProt entry P13866) is made of 664 aminoacids, with a total molecular weight of 75 KDa[32] [41]. The tridimensional structure of the protein has not been resolved yet, but recently the tertiary structure, from X-ray diffraction

data, of the variant of *Vibrio parahemoliticus* is available[42]. To date, the knowledge on the secondary structure of the transporter and on the presence and localization of ligands bindig sites (monosaccharides, sodium, or inhibitors like phlorizin) are discordant, since there are differences between the experimental methods used and, moreover, an intrinsic difficulty in the structural studies of transmembrane protein exists. SGLT1 is structured in 14 α-helix transmembrane domains, as for all the SSSF members[33, 43]. The N-terminal hydrophilic terminus is localized in the extracellular side, whereas for the C-terminus the localization is still uncertain, although many studies confirm the extracellular hypothesis[43]. In the N-terminus some *N*-glycosilation sites are present, like Asn248[44], probably necessary for the membrane targeting of the protein; however, the glycosylation is not strictly required for the transport activity[43]. In the figure below (Figure 19[33]), a topological scheme of the structure of SGLT1 is present.

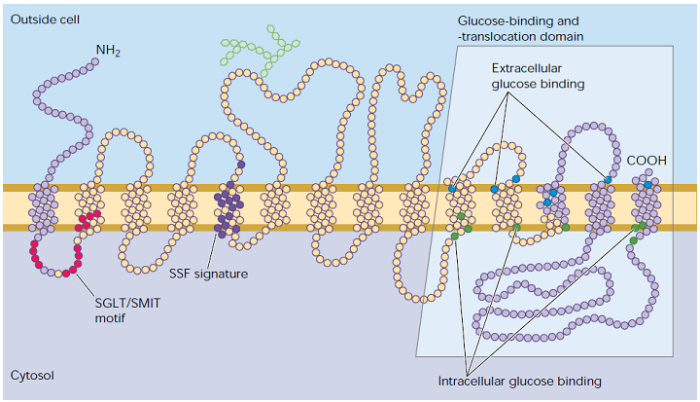


Figure 19 - Secondary structure model for SGLT1. The 664 aminacids constituting the transporter are disposed on 14 transmembrane α-helices. In this model, both N- and C-terminus are located extracellularly. In green the oligosaccharide which represents the glycosylation, on N248 residue. In yellow, residues which sharing the same architecture among SSSF members, which also show a consensus sequence, in violet. In red the consensus sequence between SGLT1 and SMIT (Sodium-MyoInositol coTransporter) is highlighted.

The presented model is the most accredited, although a first model, based on the gene structure of SGLT1, provided a cytoplasmic localization of the C-terminal extremity, and even a topology of 12 transmembrane α-helix[45].

Concerning the ligands binding sites (for glucose/galactose, sodium, phlorizin), it is now clear that the N-terminus is involved in sodium binding, where the role of A166 residue seems to be fundamental[46]. The C-terminal extremity participates in the binding and transport of sugars and/or inhibitors[47] [48]. The role, structure and localization of the C-terminus are the aspect towards which most of the studies have been focused in the last years. One of the most critical and intriguing aspect is the localization of the site/s of interactions of the two monosaccharides and inhibitors. Initially, the hypothesis provided a unique binding site; recently, several studies clarified the existence of at least two regions, one involved in initial binding (first step), directed extracellularly, and the second responsible for the real translocation inside the cell (intracellular site)[49]. A recent work has allowed to identify a precise localization for the two binding sites on the C-terminal portion[50]. Previously, it was possible to reduce the area related to the glucose binding site: several works established an important role for loop 13, an unfolded region between the α-helix TM13 and TM14 (aa 541-638), which is involved in the binding with inhibitor phlorizin[27] and alkylglucosides[51]. It has been hypothesized a direct role for this loop in the binding with glucose, too; this portion was isolated, immobilized on a lipid bilayer, in order to study its interaction features[50]. Exploiting SPR (Surface Plasmon Resonance) analysis, it was understood that loop 13 contains both a non-specific sugar interaction site, in its early portion (aa 548-571), probably located intracellularly, and a stereospecific region, in terminal part (aa 622 – 633), extracellular. This study has proposed

indirectly a double localization for the loop 13, which seems to be free to cross the plasma membrane. According to author's hypothesis, the two regions of loop 13 allow to have a multistep recognition and binding process; the initial stereospecific interaction, mediated by late loop 13, discriminates between D-Glc and other non-transported monosaccharides (as L-Glc), while the intracellular interaction is a second binding step, which allows the translocation of the sugar inside the cell. A further work indicates the existence of a sodium-independent glucose translocation[52]. This uniport has a low affinity and is not inhibited by phlorizin, which interaction was confirmed to occur extracellularly[53]. Is therefore possible to conclude that two binding sites for glucose or galactose exist: one with high affinity (K_D 0,5 mM) and one with low affinity (K_D 3 – 4 mM). Fuorescence studies indicate that sugars bind to this site with a stacking interaction between their pyranose ring and indol ring of Trp residue in position 561. This residue is located in the early portion of loop 13 which localizes intracellularly, associated with the non-stereospecific interaction. Therefore, accordingly with all these results, it is possible to hypothesize that the high affinity binding site is involved in stereospecific recognition of sugars, whereas the low affinity site is intracellular and mediates the sodium-dependent translocation of Glc or Gal. Recently, new insights for loop 13 localization have been proposed: it could behave as a re-entrant mobile loop, during binding and translocation events[54]. This idea could lead to solve many disputes about its nature.

Structural determinants for ligands binding. Many works led to understand which SGLT1 residues interact with functional groups on transported monosaccharides or inhibitors, like phlorizin. Figure 20[55] shows a model for the interactions between hydroxyl groups of glucose

and side chains of aminoacids which constitute the binding domain for sugars. This region is localized, as mentioned, between α-helix TM X and XIV. After the binding with two sodium ions, in the N-terminal region of the protein, conformational changes occur, thus allowing the exposure of the glucose binding site towards the extracellular side. A hydrophilic pocket, constituted by Q457, T460, Q445 and R499 is formed[55]. These residues form dipolar interactions and hydrogen bonds with OH groups of sugars.

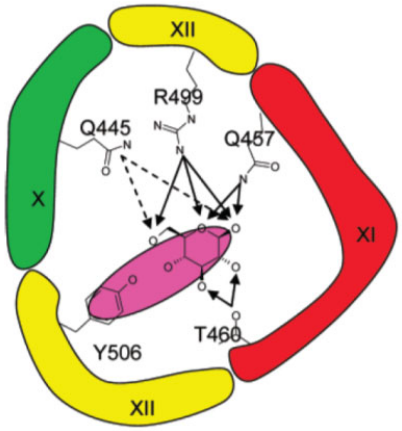


Figure 20 - Model of sugar binding domain of SGLT1.

Among the residue involved in glucose/galactose binding, the most important one is certainly Gln457: several mutation experiments of this aminoacid show a dramatic loss of affinity for glucose binding[56]. Moreover, Q457R is one of the more relevant genetic determinants of the glucose galactose malabsorption, because this mutation avoids completely the sugar absorption in the intestine.

Studies carried out with other ligands, as the inhibitor phlorizin, allowed to have a more precise interaction model of SGLT1. Phlorizin binds to the same region involved in glucose interaction, with a two-step process: the first one represents the rapid formation of a complex, while in the

second step a slow isomerization of the inhibitor occurs, causing the occlusion of the glucoside inside the binding site[53]. The aromatic portion of phlorizin (aglycone) binds to a region of loop 13 comprises between aminoacid 604 and 610, where, accordingly to this model, stacking interactions between the aromatic rings of the inhibitor and Phe602 and Phe609 residues occurs (Figure 21)[53].

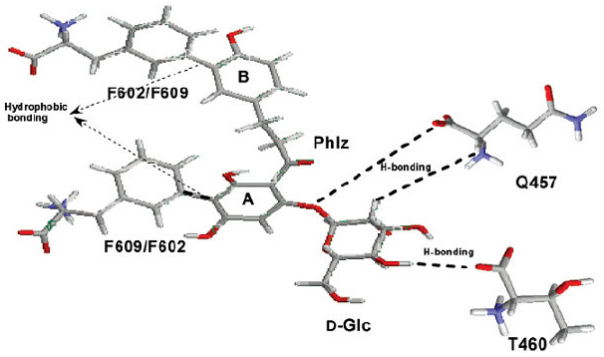


Figure 21 - Principal interaction between phlorizin and hSGLT1.

Thanks to these studies, it has been possible to understand the structural determinants for ligand interaction with SGLT1, and in particular which functional groups of ligands are critical for the binding with the C-terminal domain of SGLT1[53] [55]. First, the hydroxyl group in position C-2 of the two monosaccharides is crucial for the binding, with a compulsory equatorial configuration (D-mannose, epimer for C-2 position, in neither recognized nor transported by SGLT1). C-3 hydroxyl group is important, since SGLT1 has a low affinity for the monosaccharide with the opposite configuration on this carbon, D-allose, but shows a hydrogen bond acceptor rather than donor, since 3-OMe-D-Glc, a non-metabolizable analogue, is efficiently transported[22]. Hydroxyl group in position C-4 has an important role but its configuration is non-critical, since both Glc and Gal, epimers for this position, are translocated, with a very similar affinity constant. Substituting this OH

group with a fluoride atom is possible to increase the affinity for the protein: this means that the group in position C-4 should be a hydrogen bond acceptor. Primary OH in C-6 position is not fundamental, but the CH_2 group to which it is linked seems involved in hydrophobic interactions. Endocyclic oxygen and anomeric OH group are involved in binding, interacting by hydrogen bonds with Gln457 and Arg499 residues. The substitution of the endocyclic oxygen with nitrogen (as in the case of iminosugars), carbon (inositol) or sulphur (thiasugars) leads to a lower binding affinity with the transporter. Recently, a big increase in the knowledge of SGLT1 has been achieved with the resolution, at atomic level, of the tridimensional structure of the *Vibrio parahemoliticus* variant (vSGLT1) of this transporter[42]. This variant possesses a sequence identity with the human homologue of 32% (and a 60% of similarity), which makes the bacterial variant a good crystallographic model. The structure was obtained in the co-crystallized form with D-Galactose. The study confirms the 14 TM α-helices structure, and the extracellular localization of both the N- and C-termini (Figure 22).

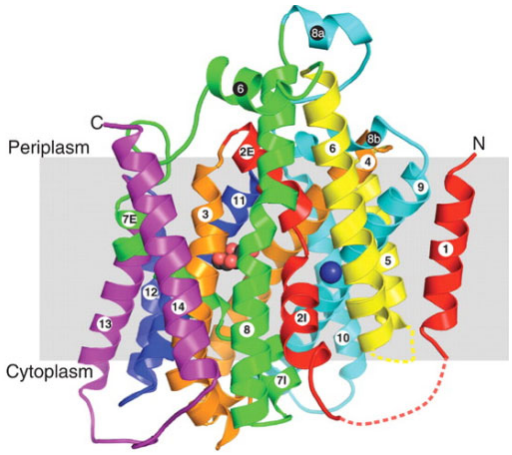


Figure 22 - Tridimensional structure of vSGLT1.

The galactose binding site is comprised between two groups of hydrophobic residues which form two barriers, one intracellular and one extracellular. These portions control the ligand transport. The extracellular barrier is made of tyrosine, metionine and phenylalanine residues which block the ligand, avoiding its leakage in the extracellular space. The intracellular gate is represented by a tyrosine residue in position 263, involved in stacking interactions with the pyranose ring of the sugar. Analyzing the residues which surround the sugar (Figure 23), it is possible to observe that there is a high similarity between them and the groups described for the human variant, according to the available models in literature.

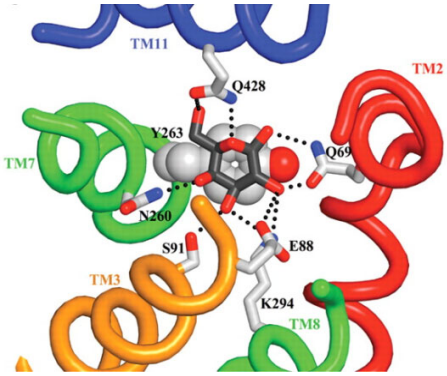


Figure 23 - Galactose binding site, from the extracellular side, obtained removing the residues constituting the extracellular gate.

The identification of the sodium binding site has required a comparison of the vSGLT1 with the model of LeuT, a bacterial transporter for proline. This site has been localized between α-helices TM 2 and 9, 10 Å far from glucose binding site. Furthermore, a functional model for this protein has been proposed. According with this work, the transporter alternates between two conformation, one extracellular-oriented, which allows the sugar binding, and one directed intracellularly, for the Glc/Gal release.

Transport kinetics of Glucose and Galactose. Transport mechanism of glucose mediated by SGLT1 has been deeply studied and modeled in the last years, thanks to the cloning and the expression of the transporter in *Xenopus* oocytes and exploiting electrophysiological techniques, which allow the measurement of the currents associated to the sugar transport. In fact, SGLT1 is an electrogenic protein, since the sugar transport is sodium-mediates in physiological conditions[32]. The co-transport is achieved by alternating access mechanism, which provides conformational changes induced by ligands. The overall mechanism was simplified in a six-state model, depicted in Figure 24[31].

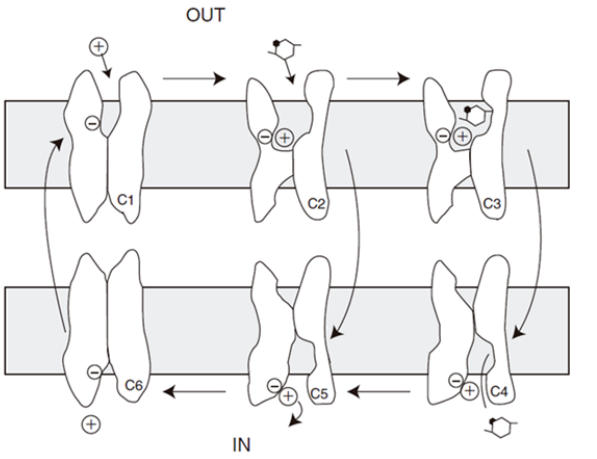


Figure 24 - Six-state kinetic model for SGLT1. First, the protein is negatively charged, with a net valence of -2 (C1). The binding of two sedum ions causes a drastic conformational change which allows the extracellular exposure of the sugar binding site (C2). C3 and C4 states represent the translocation process of the sugar inside the cell; once the monosaccharide is released, the SGLT1 affinity for sodium ion decreases and the ion is released (C5). This passage brings back the protein to its default state (C6).

Iminosugars: azaglycoderivatives and their widespread biological functions

Glycosidases inhibition: therapeutic role and biochemical features

Glycosidase inhibition has been subjected to extensive interest in the past decades. The cleavage of the glycosidic bond is a biologically widespread process and thus inhibitors of this event have many therapeutic applications[57]. The processes in which glycosidases are involved are both catabolic and anabolic, as digestion, lysosomal degradation of glycoconjugates, biosynthesis, trimming and degradation of glycoproteins. Each mentioned process has prompted towards the research of new inhibitors: for example, the inhibition of intestinal α-glycosidases can be used for diabetes treatment, lowering blood glucose levels; inhibitors of glycosidases involved in the synthesis and trimming of oligosaccharides, located on the surface of cells, can be exploited as possible treatments against influenza virus infections, as in the case of the neuraminidase inhibitors Oseltamivir and Zanamivir; inhibitors or in general ligands of glycosidases involved in the degradation of glycosphingolipids can be used as therapy for lysosomal storage disease (Gaucher and Fabry disease). Furthermore, glycosidases inhibition can be used against fungi and insects, and also be used as therapeutic agents for diabetes, obesity, genetic disorders and tumors[58].

The glycoside bond can undergo both to acidic and enzymatic hydrolysis. In acidic media, the protonation of the exocyclic oxygen in the glycosidic bond converts it into a leaving group, so the formation of an oxocarbenium ion intermediate occurs[57] (Figure 26, A).

The enzymatic hydrolysis proceeds similarly but some differences are present. There are two possible mechanism for this enzymatic transformation; in the first case, the alternated action of two different acid residues located on the active sites of these enzymes leads to a product in which the anomeric configuration is the same of the starting compound. These enzymes are called retaining glycosidases (Figure 26, B). The catalysis occurs by protonation of the exocyclic oxygen by one acidic group, thus activating the aglycone moiety as a good leaving group. When it is eliminated, a cationic intermediate, the oxocarbenium ion, is formed (Figure 25[57]).

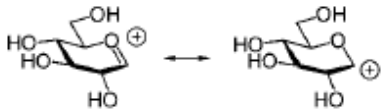


Figure 25 - Resonance limit structure of the oxocarbenium ion.

This species reacts with the other catalytic residue of the active site, which bears an acidic group in form of nucleophile, and a glycosilester intermediate is formed. The aglycone part is then substituted in the catalytic pocket by a water molecule, which is in a partial deprotonated form (due to the proximity to the conjugate base of an acid residue)[59]; its action leads to the displacement of the ester intermediate, with the formation of the hemiacetal with retention of configuration. Therefore, the mechanism of the retaining glycosidases provides two separate steps and two transition states[57]. In the inverting glycosidases, one acid residue acts at the same manner, while the other directs a water molecule to attack the activated anomeric carbon. Hence, an inversion of configuration takes place, since the water attack comes from the opposite site of the anomeric oxygen activation (Figure 26, C).

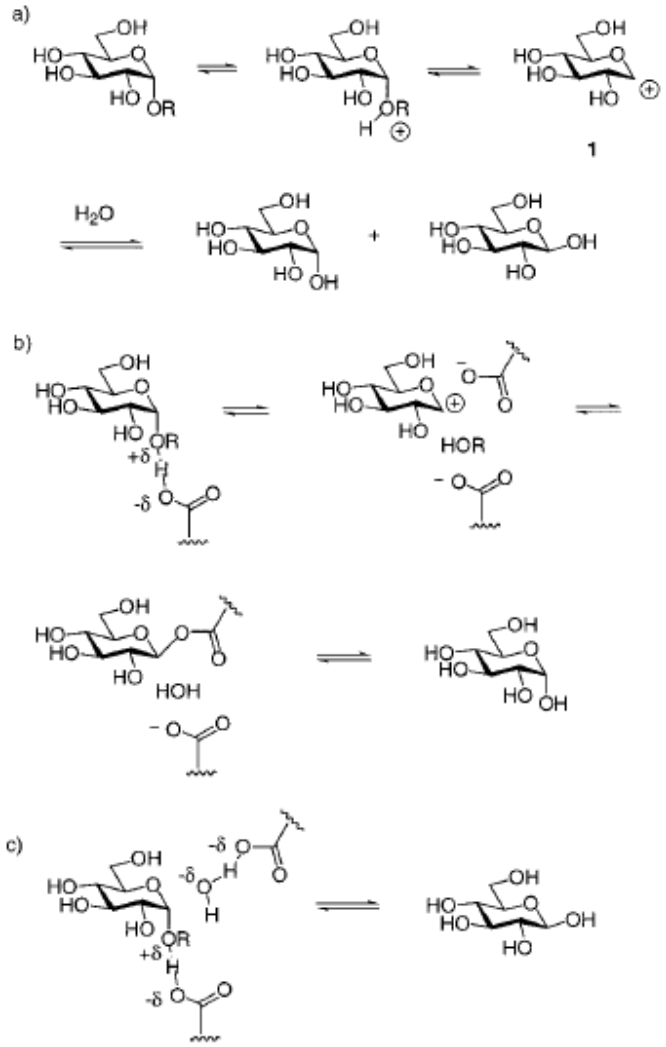


Figure 26 - Acidic and enzymatic reactions for glycosides hydrolysis.

Since the catalytic mechanism for these enzymes has been elucidated, an attractive approach for the development of inhibitors is the creation of compounds which structures resemble the transition state (TS) intermediate of the enzyme-catalyzed reaction. The rationale of this

strategy is the belief that the TS represent the step of the reaction with the highest degree of enzymatic stabilization. For glycosidases, the hypothesis is that the stabilization is very high, with a dissociation constant between the TS and the enzyme almost of 10^{-20} M[57].

Some glycosidase inhibitors are still present on the market, for the treatment of several disease in which important glycosidases are involved. Previously, some of these molecule have been mentioned. Glycosidase inhibitors used for diabetes treatment are Voglibose, Miglitol and Acarbose. Acarbose, a potent α-glucosidase inhibitors, was isolated by a strains of *Actinoplanes*, present on market since 1990 in Germany as Glucobay® and in USA since 1996[58]. This molecule is able to mimic the TS of the enzymatic hydrolysis reaction. Instead, Miglustat is a glycosidase inhibitors used for the therapy of the lysosomal storage pathology called Gaucher disease. Oseltamivir and Zanamivir represent antiviral drugs which inhibit viral glycosidases, used as anti-influenza agents. It is known that the first phases of viral infections are based on the recognition by virus of glycoconjugates on the membrane of host cells. Therefore it is frequent that antibiotics or antimicrobial drugs, both of natural and chemical origin, contains or are constituted by structures which inhibit these interactions[60].

Among antiviral drugs, the most important agents are inhibitors of neuraminidases, viral enzymes whose scope is the cleavage of glycosidic bond between sialic acid residues of some glycoconjugates located on the membrane of infected host cells. In particular two inhibitors, Zanamivir and Oseltamivir, are in market since 1999 and are active against several virus strains.

Iminosugars as glycosidase inhibitors

As mentioned, molecules mimicking the oxocarbenium ion TS of the glycoside cleavage can be potentially exploited as glycosidase inhibitors, and have been widely investigated[57]. Among these inhibitor molecules, the most important and studied are the iminosugars or azasugars. Iminosugars, whether of natural or synthetic origin, are small organic molecules that mimic carbohydrates or their hydrolysis TS, containing a nitrogen atom instead of oxygen in the ring system[61]. The first discovered molecule belonging to this family was nojirimycin, in 1966[62]. It was described as an antibiotic produced in the fermentation culture by *Streptomyces roseochromogenes* R-468 and *S. lavendulae* SF-425[62]. It shows a high inhibitory potency towards α- and β-glucosidases. In the last decades, there has been an increasing interest towards this class of molecule, and iminosugars have become versatile tools for medicinal chemists, especially for their valence as potential therapeutic agents[63]. The research has been focused on iminosugar supply both from natural sources and by chemical synthesis, providing a plethora of different nojirimycin analogues.

By a systematic and structural point of view, iminosugars can be divided into five different classes: piperidines, pyrrolidines, indolizidines, pyrrolizidines, azepanes and nortropanes[62] (Figure 27[61]).

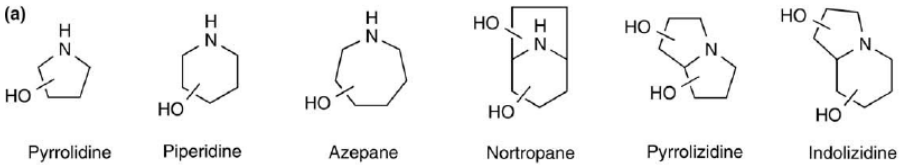


Figure 27 - Stylized representations of the iminosugar families.

Piperidines have a six-atoms ring system. Among this class, nojirimycin was the first discovered, and it can be considered as a glucose analogue. Succesively, nojirimycin B (manno-NJ) from *S. lavendulae* and galactostatine (galacto-NJ) from *S. lydicus* were also isolated (Figure 28[62]).

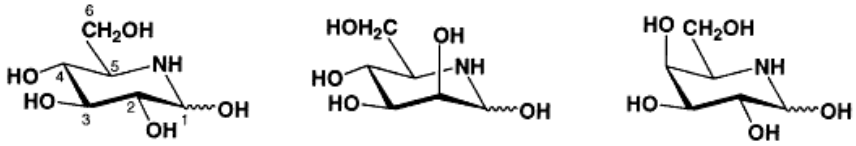


Figure 28 - Structure of nojirimycin (left), manno-NJ (centre) and galactostatine (right), three piperidine-based iminosugars.

Since these compounds are unstable for the presence of a hemiaminal group, the research has been oriented towards the more stable 1-deoxyderivative of nojirimycin analogues. Examples are 1-deoxynojirimycin (DNJ), first chemically synthesized from L-sorbose and then isolated in Nature, 1-deoxymannojirimycin and fagomine (Figure 29).

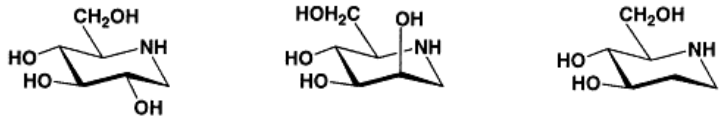


Figure 29 - Some 1-deoxy iminosugars, from left to right: 1-DNJ, 1-deoxymanno-DNJ and fagomine.

Pyrrolidines have a five carbon-membered ring which mimics the furanosidic ring of sugars. The first isolated molecule of this class was (2R,5R)-bis(dihydroxymethyl)-(3R,4R)-dihydroxypyrrolidine (DMDP, 2,5-dideoxy-2,5-imino-D-mannitol), in 1976 in leaves of *Derris elliptica*[62]. Many other similar molecules have been discovered, like (6-deoxy-DMDP, 2,5-imino-1,2,5-trideoxy-D-mannitol), 1,4-dideoxy-1,4-imino-D-

arabinitol (D-AB1), the 2-epimer of D-AB1 (1,4-dideoxy-1,4-imino-D-ribitol), the polyhydroxylpyrrolidine nectrisine and 2,5-dideoxy-2,5-imino-DL-glycero-D-manno-heptitol (homoDMDP) (Figure 30).

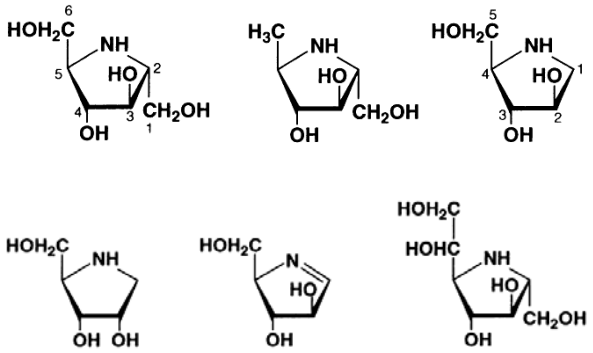


Figure 30 - Examples of pyrrolidines, from left to right: DMDP, 6-deoxy-DMDP, D-AB1, 2-*epi*-D-AB1, nectrisine and homo-DMDP.

Indolizidines (Figure 31) are constituted by a six and a five carbon-membered rings fused together. These bicyclic alkaloids show a less structural relationship to monosaccharides but the hydroxyl configuration of each ring is similar to those of sugars. The first discovered molecule are swainsonine, drug of the legumens *Swainsona canescens*, and castanospermine, from *Castanospermum australe*, in 1979 and 1981, respectively. From literature data results that bicylic compounds are more potent and more selective inhibitors than piperidines towards glycosidases. These iminosugars could be seen as the bicycle derivative of DNJ, since it has an ethylene bridge between the hydroxymethyl group and the ring nitrogen. 6-*epi*-castanospermine represent the manno- analogue, and it shows a good inhibition against human neutral α-mannosidase[62].

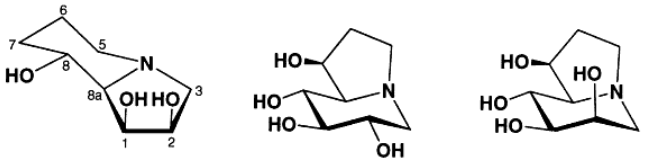


Figure 31 - Indolizidines: swainsonine, castanospermine and 6-*epi*-castanospermine.

Pyrrolizidines are other iminosugars with two fused five-membered cycles (Figure 32). Alexine was the first molecule of this family discovered, isolated from the legume *Alexa leiopetala*. Also an epimer in position C-7a was found, called 7a-*epi*-alexine or australine, from the seeds of the plant *Castanospermum australe*. This molecule can be considered as the contracted variant of the castanospermine or as a DMDP derivative, with an ethylene bridge between the hydroxymethyl group and the endocyclic nitrogen. An interesting analogue is the 7a-epi-alexaflorine, which bears a carboxylic acid in C-3 position of a pyrrolizidine ring.

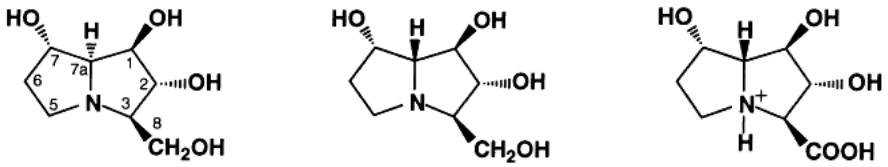


Figure 32 - Pyrrolizidine structures: alexine (on the left), 7a-*epi*-alexine or australine (centre) and 7a-*epi*-alexaflorine (on the right).

The last discovered family of natural iminosugars is represented by nortropanes (tropanes which lack of a N-methyl group), characterized by hydroxyl groups in different position and stereochemistry and a new aminoketal functionality, which generates a tertiary hydroxyl group at the bicyclic ring bridgehead[62]. Among these molecules, the most famous are the calystegines (Figure 33).

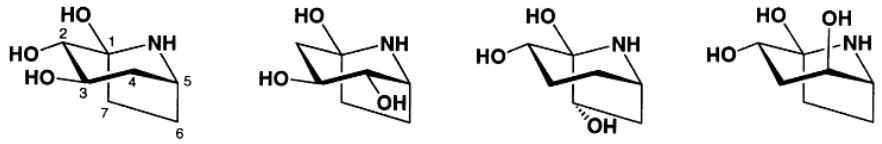


Figure 33 - Examples of calystegines, belonging to the nortropane iminosugar family.

Biological activity of iminosugars and their therapeutic applications

Scientific literature is full of works which describe the exploitation of natural and synthetic iminosugars as glycosidase inhibitors, since, as mentioned, these molecules are strongly involved in a wide range of biological events. Among the early cited biological targets, the actions of iminosugars range over psoriasis, metalloproteinases, protozoan infections and osteoarthritis (See Ref. 19-22 in Le, et al. [60]).

Iminosugars structure fit very well with the TS resemblance criterion mentioned before. In fact, these molecules are recognized by glycosidases mimicking the oxocarbenium ion transition state, the key intermediate in the glycoside bond hydrolysis. The presence of the nitrogen atom in the ring makes them metabolically inert and facilitates the protonation, thus the creation of a positive charge, at physiological pH (Figure 34[57]).

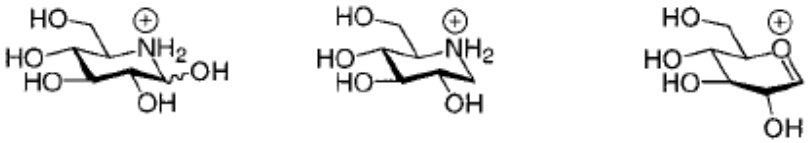


Figure 34 - Comparison between the protonated form of NH, 1-DNJ and the oxocarbenium ion intermediate occuring during glycoside hydrolysis.

Furthermore, the recognition is allowed by the structure similarity of iminosugars with monosaccharides. All these features suggest that iminosugars are perfect glycosidases inhibitors; actually, the presence of the protonated form of the nitrogen atom is not a sufficient requisite for an ideal oxocarbenium ion TS, since these molecules have chair conformations instead of the half chair conformation of the TS (due to the partial sp^2 character of anomeric carbon). For the oxocarbenium ion TS is impossible to have a charge on the ring oxygen and a chair conformation[57]; thus, NJ and DNJ, for example, are not perfect TS analogue and this could explain the good, but not surprising, inhibition constant of these two compounds.

As mentioned, one of the most important direction toward which the development of iminosugars has been focused is represented by the inhibition the glycoside hydrolase activity, in particular those regarding the digestion of dietary carbohydrates. As known, intestinal di- and oligosaccharidases are fixed component of the membrane of cell constituting the brush border membrane of the small intestine. These enzymes are involved in the digestion and hydrolysis of dietary carbohydrates, allowing their successive absorption. Due to the important physiological function, it is clear that their inhibition can be exploited for the regulation of carbohydrate absorption, in order to control several pathological states, as diabetes and obesity[64]. Acarbose, a commercially available drug (GLYCOBAY™), shows a potent inhibition activity towards pig intestinal sucrase, with an IC_{50} of 0,5 μM. It reduces postprandial blood glucose and increases insulin secretion[64]. Also the natural iminosugar DNJ showed an inhibitory effect against mammalian α-glucosidases *in vitro*, but the scarce *in vivo* effect led to the development of more effective DNJ inhibitors, as Miglitol or Emiglitate

(Figure 35[62]), reported as good agents against the postprandial elevation of blood glucose[62].

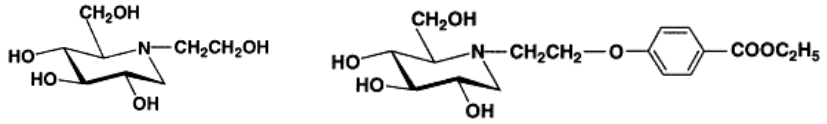


Figure 35 - Structures of miglitol (on the left) and emiglitate (on right).

The biological activity of iminosugars was also detected on processing glycosidases involved in the trimming of oligosaccharidic chain linked on Asn residues of glycoproteins. These enzymes are responsible of the diversity of the sugar chain connected with a N-linkage to glycoproteins, which are biosynthesized and correctly modified in the ER. In fact, a newly synthesized polypeptide is glycosylated on a particular Asn residue present in a consensus sequence, and successively this sugar chain is modified by a series of reaction both in ER and Golgi apparatus, by specific glycosidases and mannosidases (trimming) and glycosil transferases, which act elongating the chain with fucose, galactose, N-acetyl-glucosamine and sialic acid residues. Iminosugars as castanospermine and DNJ are potent inhibitors of glucosidase I and II localized in the ER. Mannosidases I and II are instead blocked by other iminosugars, such as deoxymannojirimycin, kifunensine (a two-fused-ring iminosugar), 1,4-deoxy-1,4-imino-D-mannitol. The inhibitory activity of iminosugars against processing glycosidases leads to a different glycosylation pattern of the N-linked oligosaccharide, thus altering the correct targeting of the nascent glycoprotein, delaying their translocation in Golgi, their secretion outside the cell or otherwise causing their degradation mediated by other cellular systems. However, the importance of iminosugars in this case is given by their potential ability to prevent an unfair retention of some glycoproteins in ER, due to

their unfolded state, although they maintain a fairly functional activity. It represents the molecular explanation of several pathologies, like cystic fibrosis, familiar hypercholesterolemia or heritable forms of pulmonary emphysema[64].

Many scientific works are focused on the use of iminosugars as therapeutic agents against the so called lysosomal storage disease (LSDs), caused by an accumulation of glycosphingolipids (GSLs) in lysosomes. These disorders are due to defective enzymes responsible for the degradation of glycolipids in lysosomes. A glycosphingolipid accumulation in these organelles occurs, causing lysosomal swelling and definitively the disease. The defect is represented by mutations in the genes coding for these enzymes, which are less or not active, thus leading to partial or absent glycosphingolipid degradation[64]. The degradation pathway include eight enzymatic transformation, and each one, if hampered, leads to a different pathology (Figure 36[62]).

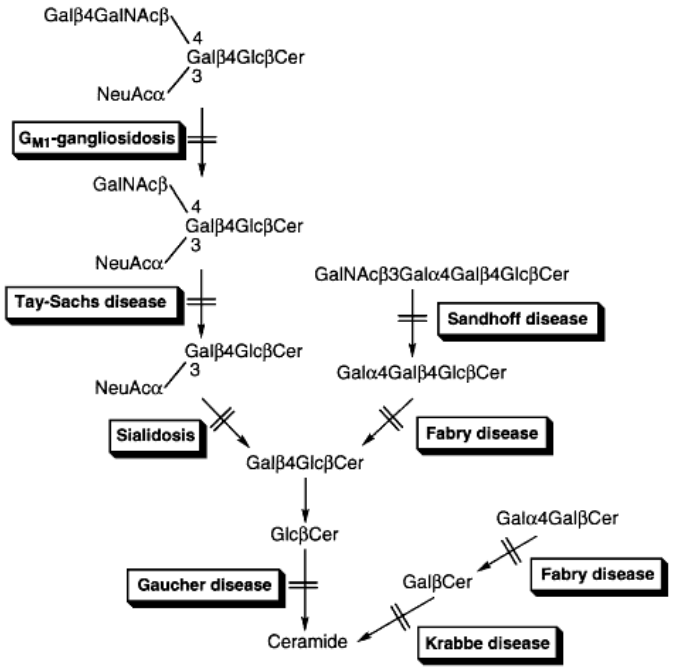


Figure 36 - Lysosomal degradation pathway for glycosphingolipids. On each blocked passage the associated disease is indicated.

Several therapies have been developed to treat these pathologies. In the enzyme replacement therapy (ERT), a periodic substitution of the defective enzyme is performed, to ensure a control of GSLs level in lysosomes. This treatment is not free of heavy side effects; in one of the most important LSD, Gaucher disease, ERT consists of injection of recombinant glucocerebrosidases, which helps the regression of the disease. SRT (Substrate Reduction Therapy) consist of a chemiotherapic treatment with drugs as Miglustat (N-butyl-DNJ), able to inhibit the glucosylceramide synthase, a glycosyl transferase responsible for the synthesis of many GSLs. Miglustat is used for the treatment of type I Gaucher disease, caused by a dysfunction of glucocerebrosidase enzyme that provokes an intracellular accumulation of glucosylceramide causing

liver and spleen enlargement[61]. The reduction of substrate (GSL) levels allows a less accumulation of glycolipids in lysosomes. In the last decades, a new approach has been developed for the treatment of these pathology, and it consists on the development of molecules able to restore the correct protein folding of the target defective enzymes. Since iminosugars are good ligands for these enzymes, some of them have been developed and tested for this ability. Iminosugars would act as real "molecular chaperones", since the interactions that occur between iminosugar ligands and the binding site of the target enzyme can stabilize the active and functional conformation of the enzyme. The correct conformation leads not only to a restored lysosomal activity, but also to a proper protein trafficking, since in some cases the diminished enzymatic activity is correlated to a less lysosomal glycohydrolase content.

The best known LSDs are Fabry and Gaucher diseases. The first pathology is caused a deficiency of human α-galactosidase A (α-GalA)[64]. The substrate of this enzyme, globotriosylceramide, progressively accumulates within visceral tissues and body fluids[64]. After a first therapeutic approach based on ERT, recently the use of 1-deoxygalactonojirimycin A as molecular chaperone for this enzyme has been developed[58].

Since, as mentioned, glycosidases are highly involved in the oligosaccharide biosynthesis pathway and since tumor cells are decorated with a different glycosylation pattern then healthy cells, iminosugars show important biological activities against tumor development. Inhibiting the glycohydrolase activities, it could be possible to alter the glycosylation structure of tumor cells, thus contrasting the tumor growth. As example, swainsonine stops the tumor growth and stimulates the immune response; castanospermine and Me-DNJ showed

anti-metastasis ability, inhibiting cell aggregation and cell adhesion of cancer cell to vascular endothelium. Other glucosidases are often present at high level in the interstitial fluids of tumor because they degrade glycoconjugates of extracellular matrix (ECM), allowing the tumor cell invasion. Iminosugars such as 2-acetamido-1,5-imino-1,2,5-trideoxy-D-glucitol and an iminosugar-like glucuronic analogue are able to inhibit these glycosidases.

Iminosugars have been developed also as antiviral agents. α-Glucosidases I and II, involved in the biosynthesis of glycoprotein of viral capside during a viral cell cycle of an infection are molecular targets for the iminosugar N-butyl-DNJ and its peracylated prodrug derivative, which were developed as anti-HIV drugs.

Multivalency and glycobiology

The surface of a cell is a complex and dynamic environment[65], made of a huge number of proteins and sugars of different types and shape. The cell calyx acts both as a sensor of the external environment and as information carrier, because it permits the communication between cells (also pathogens). The cell-cell contact or between cells and biomolecules is based on specific ligand-receptor binding events. The rationalization of these processes, i.e. studying how single ligands bind their specific receptors and what happens in the cell after these events, allows the development of strategies, approaches and molecules as potential tools able to stop or block these events. But the cell-cell interactions cannot be fully rationalized in this manner, since, as mentioned, the cell calix is full of molecules that interact with their receptors on the other cell, and the biological intracellular effects are multiple. Hence, it is clear that another way to study this process must be taken into account. In general, on the cell surface it is possible to find oligosaccharides of glycoproteins and glycolipids, displayed as multiple copies, and these structures are called multivalent. Multivalent structures of carbohydrates of a cell surface are involved in many cell-cell recognition events, because they interact with receptors of the opposite cell. For examples, chemokines interact with glycosaminoglycans displayed in a multivalent fashion and this recognition leads to intracellular events that definitively causes leukocytes recruitment[65]. The knowledge about glycans-glycan binding proteins/lectins interactions is well structured. Lectins are carbohydrate-binding proteins which bind specific moieties of glycan chains, although this interaction exerts a very low affinity. However, this feature is compensated in biological systems by multivalency interaction, both from lectin structure, containing more

than one carbohydrate binding site, and/or by the presentation of the glycan moieties with a multivalent fashion on the cell surface[66]. Multivalent ligands often have increasing affinity for their target receptors compared to their monovalent form[65]. The reasons for the stronger biological effect of presenting and displaying a carbohydrate, or in general a molecule, in a multivalent manner, i.e. the rationale models for the multivalency, are multiple (Figure 37[65]).

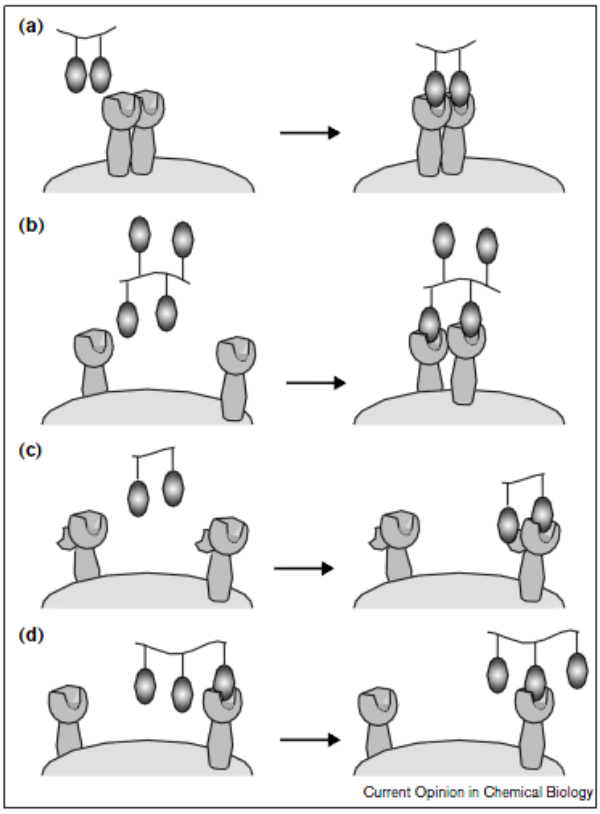


Figure 37 - Mechanism of cell-surface receptors – multivalent ligands interactions. A) Chelate effect. B) Cluster effect. C) Subsite binding. D) Increasing of ligand local concentration.

For example, multivalent ligands can bind and be recognized by an oligomeric receptor on the cell surface of the interacting cell. This hypothesis is called chelate effect. The high affinity of the multivalent structure for the oligomeric receptor is due to a facilitated binding of the monomeric ligands after the first receptor-multivalent ligand interaction. This is possible also for receptor not structured as oligomers, since the membrane bilayer shows a high fluidity, therefore allowing the receptor clusterization (cluster effect). The clusterization leads to the activation of intracellular signaling pathways, often associated with an amplification of the signal, thus explaining the biological effect of multivalency. Alternatively, with a multivalent-displayed ligand it is possible to exploit the subsites often possessed by some membrane proteins, as for example allosteric sites. Even in the case of only one receptor, it is possible to rationalize a multivalent effect, because a polyvalent-presented ligand can display a higher local concentration of the ligands. In fact, for multivalency-based events, statistical effects due to effective concentrations and rebinding occurrences should be taking into account, although the multivalency effect was also described as "the enhancement in the activity of a multivalent ligand beyond what would be expected due to the increase in sugar local concentration alone"[67].

The study of multivalent interactions has prompted towards the development and the preparation of multivalent structures and scaffold decorated with biological interesting ligands, such as carbohydrates or peptides. The chemical synthesis of multivalent systems requires to take into account many critical aspects, such as the shape and the nature of the multivalent structure, the selection of the spacer which divides the scaffold and the ligand, the stability of the whole structure, its solubility in aqueous solutions and the toxicity. The types of scaffold used for multivalent ligand generation are several: glycoliposomes,

glycocalixarenes, fullerenes, glycopolymers, nanoparticles, carbon nanotubes have been developed for the study of a plethora of biological events in which the multivalency shows a critical role. The multivalent-functionalized structures can exert both inhibitory or agonistic effects[65].

References

[1] G. H. Schmid, *Organic chemistry*, WCB/MacGraw-Hill, Dubuque, IA, **1997**.

[2] D. B. Werz, P. H. Seeberger, *Chemistry-a European Journal* **2005**, *11*, 3194-3206.

[3] D. L. Nelson, M. M. Cox, *Lehninger principles of biochemistry*, W.H. Freeman and Co., New York City, NY, **2008**.

[4] L. Ellgaard, M. Molinari, A. Helenius, *Science* **1999**, *286*, 1882-1888.

[5] R. G. Spiro, *Glycobiology* **2002**, *12*, 43R-56R.

[6] R. A. Dwek, *Chemical Reviews* **1996**, *96*, 683-720.

[7] S. Hakomori, *Cancer Research* **1985**, *45*, 2405-2414.

[8] J. W. Dennis, M. Granovsky, C. E. Warren, *Bioessays* **1999**, *21*, 412-421.

[9] C. Janeway, *Immunobiology : the immune system in health and disease*, Garland Science, New York, **2005**.

[10] R. Sasisekharan, R. Raman, V. Prabhakar, in *Annual Review of Biomedical Engineering, Vol. 8*, **2006**, pp. 181-231.

[11] B. Ernst, J. L. Magnani, *Nat Rev Drug Discov* **2009**, *8*, 661-677.

[12] F. Nicotra, L. Cipolla, B. La Ferla, C. Airoldi, C. Zona, A. Orsato, N. Shaikh, L. Russo, *Journal of Biotechnology*, *In Press, Corrected Proof*.

[13] M. J. McKay, H. M. Nguyen, *Acs Catalysis* **2012**, *2*, 1563-1595.

[14] F. Nicotra, C. Airoldi, F. Cardona, P. K. Johannis, in *Comprehensive Glycoscience*, Elsevier, Oxford, **2007**, pp. 647-683.

[15] D. E. Levy, P. Fügedi, *The organic chemistry of sugars*, Taylor & Francis, Boca Raton, **2006**.

[16] J. Cohen, *Nature* **2002**, *420*, 885-891.

[17] R. C. Bone, R. A. Balk, F. B. Cerra, R. P. Dellinger, A. M. Fein, W. A. Knaus, R. M. H. Schein, W. J. Sibbald, *Chest* **1992**, *101*, 1644-1655.

[18] J. A. Yethon, C. Whitfield, *Current Drug Targets - Infectious Disorders* **2001**, *1*, 91-106.

[19] S. I. Miller, R. K. Ernst, M. W. Bader, *Nat. Rev. Microbiol.* **2005**, *3*, 36-46.

[20] P. M. Lavoie, J. Thibodeau, F. Erard, R. P. Sekaly, *Immunol. Rev.* **1999**, *168*, 257-269.

[21] S. Bauer, C. J. Kirschning, H. Hacker, V. Redecke, S. Hausmann, S. Akira, H. Wagner, G. B. Lipford, *Proceedings of the National Academy of Sciences of the United States of America* **2001**, *98*, 9237-9242.

[22] M. Palazzo, S. Gariboldi, L. Zanobbio, S. Selleri, G. F. Dusio, V. Mauro, A. Rossini, A. Balsari, C. Rumio, *Journal of Immunology* **2008**, *181*, 3126-3136.

[23] E. Cario, *Gut* **2005**, *54*, 1182-1193.

[24] T. Wemer, D. Haller, *Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis* **2007**, *622*, 42-57.

[25] L. C. H. Yu, J. R. Turner, A. G. Buret, *Experimental Cell Research* **2006**, *312*, 3276-3286.

[26] L. C. H. Yu, A. N. Flynn, J. R. Turner, A. G. Buret, *Faseb Journal* **2005**, *19*, 1822-1835.

[27] M. M. Raja, N. K. Tyagi, R. K. H. Kinne, *Journal of Biological Chemistry* **2003**, *278*, 49154-49163.

[28] B. M. Chung, L. E. Wallace, J. A. Hardin, D. G. Gall, *Canadian Journal of Physiology and Pharmacology* **2002**, *80*, 872-878.

[29] A. N. Alexander, H. V. Carey, *American Journal of Physiology-Gastrointestinal and Liver Physiology* **2001**, *280*, G222-G228.

[30] E. M. Wright, B. A. Hirayama, D. F. Loo, *Journal of Internal Medicine* **2007**, *261*, 32-43.

[31] E. M. Wright, D. F. Loo, B. A. Hirayama, E. Turk, *Chapter 64. Sugar Absorpition, In: Johnson LR et al. Physiology of the Gastrointestinal Tract, 4th Ed., San Diego : Elsevier/Academic Press* **2006**.

[32] E. M. Wright, *Annu. Rev. Physiol.* **1993**, *55*, 575-589.

[33] E. M. Wright, D. D. F. Loo, B. A. Hirayama, E. Turk, *Physiology* **2004**, *19*, 370-376.

[34] H. Jung, *Febs Letters* **2002**, *529*, 73-77.

[35] E. M. Wright, E. Turk, M. G. Martin, **2002**, pp. 115-121.

[36] M. Quick, J. Tomasevic, E. M. Wright, *Biochemistry* **2003**, *42*, 9147-9152.

[37] N. Hirschhorn, W. B. Greenough, *Scientific American* **1991**, *264*, 50-56.

[38] aT. Zeuthen, B. Belhage, E. Zeuthen, *Journal of Physiology-London* **2006**, *570*, 485-499; bD. D. F. Loo, E. M. Wright, T. Zeuthen, *Journal of Physiology-London* **2002**, *542*, 53-60.

[39] J. Dyer, I. S. Wood, A. Palejwala, A. Ellis, S. P. Shirazi-Beechey, *American Journal of Physiology-Gastrointestinal and Liver Physiology* **2002**, *282*, G241-G248.

[40] D. F. Diedrich, *Archives of Biochemistry and Biophysics* **1966**, *117*, 248-&.

[41] **UniProtKB P13866 (SC5A1_HUMAN) http://www.uniprot.org/uniprot/P13866**.

[42] S. Faham, A. Watanabe, G. M. Besserer, D. Cascio, A. Specht, B. A. Hirayama, E. M. Wright, J. Abramson, *Science* **2008**, *321*, 810-814.

[43] E. M. Wright, E. Turk, *Pflugers Archiv-European Journal of Physiology* **2004**, *447*, 510-518.

[44] E. Turk, C. J. Kerner, M. P. Lostao, E. M. Wright, *Journal of Biological Chemistry* **1996**, *271*, 1925-1934.

[45] E. Turk, M. G. Martin, E. M. Wright, *Journal of Biological Chemistry* **1994**, *269*, 15204-15209.

[46] A. K. Meinild, D. D. F. Loo, B. A. Hirayama, E. Gallardo, E. M. Wright, *Biochemistry* **2001**, *40*, 11897-11904.

[47] M. PanayotovaHeiermann, S. Eskandari, E. Turk, G. A. Zampighi, E. M. Wright, *Journal of Biological Chemistry* **1997**, *272*, 20324-20327.

[48] M. PanayotovaHeiermann, D. D. F. Loo, C. T. Kong, J. E. Lever, E. M. Wright, *Journal of Biological Chemistry* **1996**, *271*, 10029-10034.

[49] T. Puntheeranurak, B. Wimmer, F. Castaneda, H. J. Gruber, P. Hinterdorfer, R. K. H. Kinne, *Biochemistry* **2007**, *46*, 2797-2804.

[50] B. Wimmer, M. Raja, P. Hinterdorfer, H. J. Gruber, R. K. H. Kinne, *Journal of Biological Chemistry* **2009**, *284*, 983-991.

[51] M. M. Raja, H. Kipp, R. K. H. Kinne, *Biochemistry* **2004**, *43*, 10944-10951.

[52] A. Kumar, N. K. Tyagi, P. Goyal, D. Pandey, W. Siess, R. K. H. Kinne, *Biochemistry* **2007**, *46*, 2758-2766.

[53] N. K. Tyagi, A. Kumar, P. Goyal, D. Pandey, W. Siess, R. R. H. Kinne, *Biochemistry* **2007**, *46*, 13616-13628.

[54] T. Puntheeranurak, M. Kasch, X. B. Xia, P. Hinterdorfer, R. K. H. Kinne, *Journal of Biological Chemistry* **2007**, *282*, 25222-25230.

[55] B. A. Hirayama, D. D. F. Loo, A. Diez-Sampedro, D. W. Leung, A. K. Meinild, M. Lai-Bing, E. Turk, E. M. Wright, *Biochemistry* **2007**, *46*, 13391-13406.

[56] A. Diez-Sampedro, E. M. Wright, B. A. Hirayama, *Journal of Biological Chemistry* **2001**, *276*, 49188-49194.

[57] V. H. Lillelund, H. H. Jensen, X. F. Liang, M. Bols, *Chemical Reviews* **2002**, *102*, 515-553.

[58] N. Asano, *Glycobiology* **2003**, *13*, 93R-104R.

[59] T. D. Heightman, A. T. Vasella, *Angewandte Chemie-International Edition* **1999**, *38*, 750-770.

[60] G. T. Le, G. Abbenante, B. Becker, M. Grathwohl, J. Halliday, G. Tometzki, J. Zuegg, W. Meutermans, *Drug Discovery Today* **2003**, *8*, 701-709.

[61] G. Horne, F. X. Wilson, J. Tinsley, D. H. Williams, R. Storer, *Drug discovery today* **2011**, *16*, 107-118.

[62] N. Asano, R. J. Nash, R. J. Molyneux, G. W. J. Fleet, *Tetrahedron-Asymmetry* **2000**, *11*, 1645-1680.

[63] D. D'Alonzo, A. Guaragna, G. Palumbo, *Current medicinal chemistry* **2009**, *16*, 473-505.

[64] N. Asano, *Curr. Top. Med. Chem.* **2003**, *3*, 471-484.

[65] L. L. Kiessling, J. E. Gestwicki, L. E. Strong, *Current Opinion in Chemical Biology* **2000**, *4*, 696-703.

[66] M. Mammen, S. K. Choi, G. M. Whitesides, *Angewandte Chemie-International Edition* **1998**, *37*, 2755-2794.

[67] Y. C. Lee, R. T. Lee, *Neoglycoconjugates : preparation and applications*, Academic Press, San Diego, **1994**.

Chapter 2. Development of a library of dansyl-*C*-glycoderivatives as SGLT1 ligand tools: synthesis, biological evaluation and mechanism of action.

Abstract

In this paper we report the generation of a library of analogues of *C*-glucoside **1**, an anti-inflammatory glycomimetic recently developed in our laboratory, in order to achieve additional information about its unclear mechanism of action. Preliminary biological experiments performed on an *in vitro* model of doxorubicin-induced mucositis, a severe intestinal inflammatory state, indicate that the presence of the aromatic moiety which characterize all the compounds of the library is important for the biological activity, while less influence seems to be associated to the sugar part. The results will be exploited for the generation of new and more potent antiinflammatory compounds and for the comprehension and rationalization of the antiinflammatory mechanism of action.

Introduction

The sodium glucose co-transporter 1 (SGLT1) is a high affinity/low capacity glucose and galactose transporter, mainly expressed on the apical membrane of intestinal epithelium cells and of cells of S3 segment of proximal tubule of kidneys[1] [2]. This transporter is responsible for glucose and galactose absorption in intestinal tissue and for sugar reuptake in kidneys. Its fundamental physiological function has made this protein an important molecular target for the development of compounds able to inhibit the sugar absorption, to be exploited as potential antidiabetic drugs. Many efforts have been made to clarify the physiological, functional and structural features of this transporter; nowadays the knowledge of its functional behavior is profound, although the overall tridimensional structure of the human variant remains still

unknown. Recently, alongside the known physiological function, a new immunological function role for this protein has been discovered Several works indicate its involvement in the D-glucose-mediated protection of the intestinal mucosa against a series of insults and injury, as lipopolysaccharides (LPSs). High glucose concentrations, tested both *in vitro* and orally on mice, are able to protect the intestinal epithelium from LPS-induced inflammatory injury in murine models of septic shock, and data suggest a SGLT1 engagement in this phenomenon. The high glucose dosage required represents a significant drawback for the usage of this compound as potential drug, due to the inevitable impact on metabolism, so we developed a library of glycomimetics which could act at the same manner but with pharmacological concentration. Among the synthesized compounds, the most active was the *C*-glycoside **1** which exerts a high protective effect at concentration of several orders of magnitude lower than D-glucose. This compound is a *C*-glycoside, thus a non-metabolizable glycoderivative, important feature to take into account in the development of carbohydrate-base drugs.

In these years our efforts have been focused on the study of the SGLT1-mediated protective role of **1** on several inflammation states and injuries, as Chron disease, ulcerative colitis and on doxorubicin and 5-fluorouracil-induced mucositis. The latest pathological states are serious side effect of doxorubicin and 5-fluorouracil-based tumor chemotherapy. Mucositis are common in patient treated by a chemotherapy regimen for head and neck cancer, and derive from the direct cytotoxic effect of the drugs on the mucosal cells. The heavy anticancer agents induce apoptosis and reduced proliferation on epithelial cells, thus leading to enhanced permeability and ulceration of the mucosae. For these reasons, it is necessary to develop new supporting therapies for chemotherapy-induced mucositis.

Furthermore, our research has been oriented to the comprehension at the basis of the protective role exerted by glucose and synthesized glycomimetics. Alongside to the studies on the possible intracellular processes, with the construction of the putative signal transduction pathway which could connect the SGLT1 activation and the blockage of the inflammatory response, our efforts are directed on the study of how D-glucose and *C*-glycosides can activate SGLT1, involving it in the cellular protection events.

After a first library of potential SGLT1 ligands (see La Ferla, et al. [3]), constituted by glycomimetics which bear mainly differences in the aromatic moiety, we decided to build a new small library of *C*-glucoside **1** analogues, in order to clarify and rationalize the importance of the sugar region in the activity of these compounds. The developed molecules have been preliminarily tested for their protective activity on a doxorubicin-induced mucositis in *in vitro* models, to gain preliminary information on structure-activity relashionship.

Results and discussion

Compound **1** (Figure 38) is a *C*-glucoside with a dansyl residue connected to a glycopyranose ring through an ethyleneamine spacer. The synthetic approaches developed for this compound were described previously, in La Ferla, et al. [3]. Biological evaluation of the small glycomimetic library prepared in that work, in order to identify some structural requirements correlated with the anti-inflammatory activity on HT29 cell line treated with LPS, indicated that *C*-galactose analogue is still effective as anti-inflammatory agents, at the same manner of compound **1**, that the aromatic entity is crucial for the activity and that the absence of the dimethylamino group on the naphtyl moiety causes a decrease in the

anti-inflammatory effect. With the present work we moved our attention to the biological importance of the saccharidic portion of the molecules, and we designed and synthesized a new small library of *C*-glucoside **1** analogues (Figure 39).

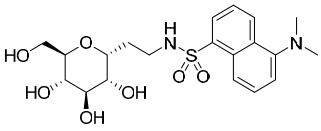


Figure 38 - Structure of *C*-glycoside 1.

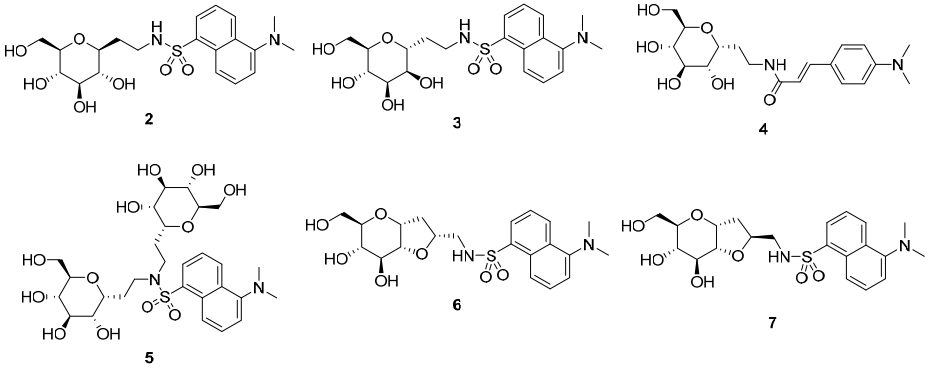
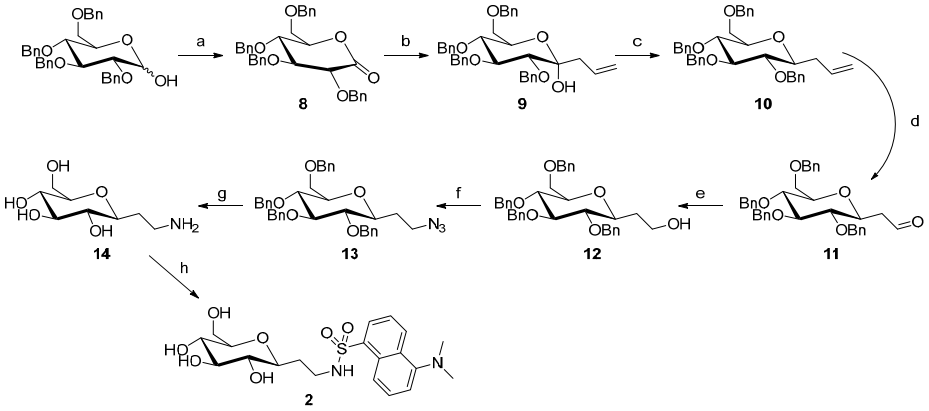


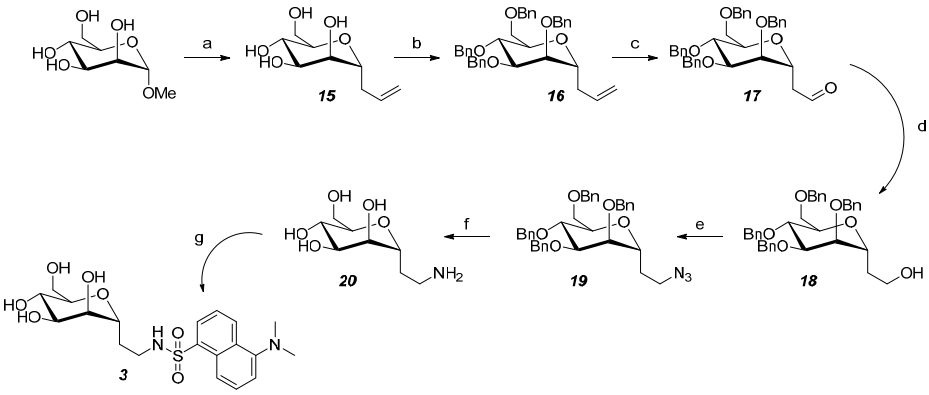
Figure 39 - Library of compound 1 analogues.

Compound **2** represent the dansyl-β-glucoside analogue of **1**. We planned to design and synthesize this compound since, according to our hypotesis, β-glucosides should not have anti-inflammatory activity: this class of compounds acts as inhibitors of SGLT1, rather than activators/agonist of this transporter/receptor, (as we suppose compound **1** is). One example is represented by the most important SGLT1 inhibitors, phlorizin, which binds to SGLT1 avoiding glucose absorption mediated by the transporter. Furthermore, phlorizin is able to stop the glucose-mediated protective effect against inflammatory and apoptotic insults[4] [5] [6]. A similar choice was done for compound **3**, the manno- analogue: it is well established that D-mannose is neither

recognized nor transported by SGLT1[7]; hence, hypothesizing that our bioactive *C*-glycosides exert the anti-inflammatory activity, as SGLT1 agonists, through a binding event to the same interaction site of sugars, compound **3** should not exert the protective effect. The synthetic pathways for compound **2** and **3** are depicted in Scheme 2 and Scheme 3. For compound **2**, 2,3,4,6-tetra-*O*-benzyl-D-glucopyranose was chosen as starting material, and converted to the β-*C*-allyl glucoside **10** according to procedure already described[8]. Briefly 2,3,4,6-tetra-*O*-benzyl-D-glucopyranose was oxidized to protected gluconolactone **8**, followed by *C*-allylation upon reaction with suitable Grignard reagents to afford lactol **9,** which was then subjected to reductive deoxygenation affording *C*-Allyl glucoside **10**. Its ozone-mediated oxidation and successive hydride reduction afforded alcohol **12**; the obtained primary alcohol was converted into an azido group (compound **13**) upon Mitzunobu reaction. The hydrogenation/hydrogenolysis on this compound afforded unprotected *C*-ethyleneamine derivative 14, which was converted in the final product **2**, by reaction with dansyl chloride. For compound **3**, the *C*-allylation was performed directly on the unprotected α-*O*-Me-D-mannospyranoside, exploting the higher stereoselection associated to the anomeric functionalization reaction on this monosaccharide. The exhaustive benzylation gave intermediate **16**, and the same pathway used for compound **2** was carried out, affording the final compound **3**.

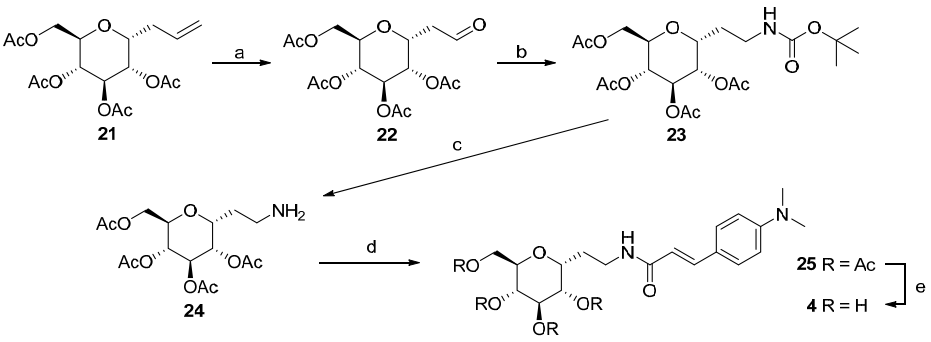


Scheme 2 - Synthesis of compound 2. Reagents and conditions: a) Pyridinium chlorochromate, CH_2Cl_2 dry, molecular sieve 4 Å, r.t., 12 h[8]; b) AllylMgBr, Et_2O dry, -78°C, 3 h[9]; c) Et_3SiH, $BF_3{\cdot}OEt_2$, CH_3CN, -78°C, 18 h[9]; d) O_3, CH_2Cl_2, -78°C, 1h, then Ph_3P, -78°C to r.t., 12h, 73%; d) $NaBH_4$, CH_2Cl_2/EtOH, r.t., 3 h, 90%; e) Ph_3P, DIAD, $(PhO)_2PON_3$, THF dry, 0°C to r.t., 1h, 93%; f) H_2, $Pd(OH)_2$/C, AcOH, AcOEt/MeOH, r.t., 24 h, quant.; g) dansylCl, Et_3N, H_2O/THF, r.t., 1h, 35%.



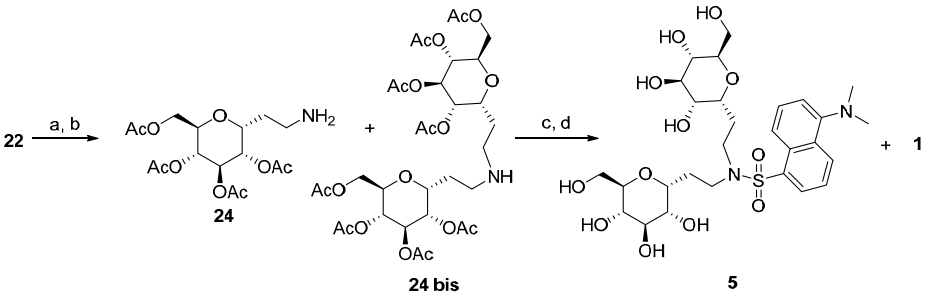
Scheme 3 - Synthesis of compound 3. Reagents and conditions: a) BTSTFA, CH_3CN dry, reflux, 3 h, than TMSOTF, AllylTMS, 0°C to r.t., 16 h; b) BnBr, NaH 60%, DMF dry, r.t., 5h, 75% (two steps); c) O_3, CH_2Cl_2, -78°C, 1h, then Ph_3P, -78°C to r.t., 12h; d) $NaBH_4$, CH_2Cl_2/EtOH, r.t., 3 h, 60% (two steps); e) Ph_3P, DIAD, $(PhO)_2PON_3$, THF dry, 0°C to r.t., 1h, 80%; f) H_2, $Pd(OH)_2$/C, AcOH, AcOEt/MeOH, r.t., 24 h, 64%; g) dansylCl, Et_3N, H_2O/THF, r.t., 1h, 25%.

C-glycoside **4** is derived from a coupling reaction of the *C*-glucosyl ethanamine **24** with a derivative of cinnamic acid, commercially available. Compound **4** was synthesized due to the high similarity with **1**, bearing a dimethylamino group attached to a vinyl-functionalized aromatic nucleus, structure present in the dansyl entity. The synthesis (Scheme 4) involves the obtainment of the amine intermediate **24**, prepared by a protecting group strategy, in which the peracetylated *C*-allyl glucopyranose **21** (McGarvey, et al. [9]) is first converted to the corresponding aldehyde **22**; the obtained carbonyl group is transformed into a tert-butyl carbamate (**23**), according to a procedure of Dubé and Scholte [10]. The carbamate and acetyl esters deprotection afforded amine **24,** which was finally coupled with *N,N*-dimethylamino cinnamic acid and deprotected to afford the desired product.



Scheme 4 - Synthesis of compound 4. Reagents and conditions: a) OsO_4, $NaIO_4$, tBuOH/Acetone/H_2O, r.t., 5 h[11]; b) tert-butylcarbamate, Et_3SiH, TFA, CH_3CN, r.t., 3h, 71%; c) TFA 50% in CH_2Cl_2, r.t., 2 h, quant.; d) N,N-dimethylaminocinnamic acid, HOBt, DCC, iPr_2EtN, DMF, r.t., 12 h, 47%; e) MeONa, MeOH, r.t., 2 h, 24%.

Compound **5** (Scheme 5) represents a side product isolated during the synthesis of *C*-glucoside **1**, obtained adopting the synthetic strategy depicted in scheme 3.

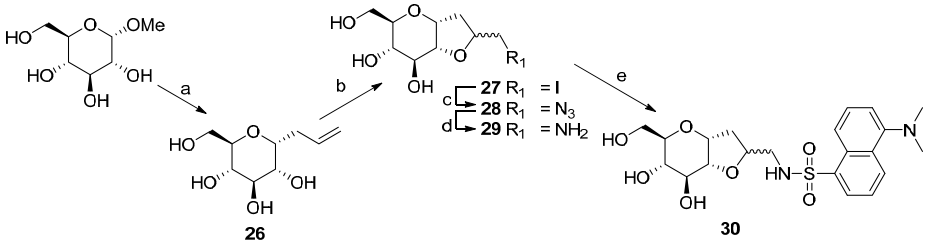


Scheme 5 - Synthesis of compound 5. Reagents and conditions: a) tert-butylcarbamate, Et_3SiH, TFA, CH_3CN, r.t., 3h; b) TFA 50% in CH_2Cl_2, r.t., 2 h; c) dansyl chloride, Et_3N, MeOH, r.t., 3 h; d) MeONa, MeOH, r.t., 2 h.

Compound **5** is a kind of dimer, in which two *C*-ethanamine glucopyranosyl residues are attached to a single dansyl unit. Since we were focused on the large-scale set up of compound **1** preparation, we performed the synthesis in gram scale. Our most plausible hypothesis for the origin of this compound is the creation, on a scaled-up reaction, of a glucosyl dimer during the carbamate formation on the peracetylated *C*-glucosylethanaldehyde, since this reaction is a kind of reductive amination, and it is reported that the formation of di- and tri- alkylated species during such reactions often occurs.

Compounds **6** and **7** are derivatives of **1**, consisting of a *C*-glucosidic bicycle structure, with a terminal amino group which was used for the dansyl linkage. The preparation of these compound was performed within the development of *C*-glucoside **1** analogues with improved water solubility, since the latest compound shows a low solubility in aqueous media (0,5 mg/mL, 1 mM) and aggregation phenomena over time at higher concentrations. The amphiphilicity of **1** is probably the cause of this character, thus, in our idea, a bicyclic scaffold could ensure a higher structural rigidity, which should prevent the formation of aggregates, allowing a better solubility in water. The exploited approach for the preparation of the bicyclic compounds has been previously developed by

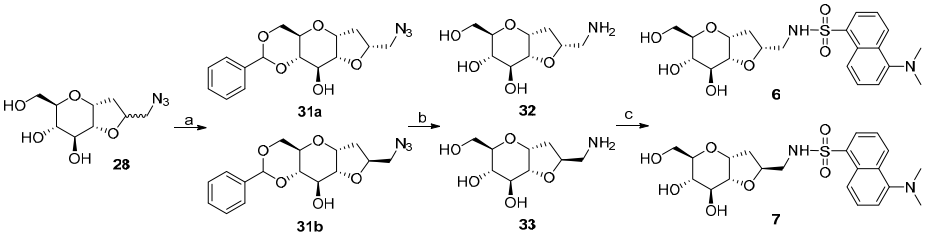
our group (see ref. of La Ferla, et al. [12a], Mari, et al. [12b]). A iodocyclization reaction between the hydroxyl group of *C*-2 position and the allylic olefin of α-*C*-allyl-glucopyranoside, mediated by iodine-donating species, such as I_2 or *N*-iodo-succinimide (NIS), occurs. A mixture of two diastereoisomer is generated, depending on the side of attack of the OH group on the iodonium structure formed. The resulted byciclic *C*-glucoside mixture bears an electrophilic site which can be exploited for a substitution with, for example, an azido group, easily convertible in a free amine. The two stereoisomers cannot be separated with simple silica gel chromatography; initially the diastereoisomeric mixture of the two iodine-bearing bicycles, described elsewere[13] was used for the synthesis, generating a mixture of the two dansyl-bearing bicycles, in a ratio of 70:30 (according to ^{1}H-NMR analysis) (Scheme 6).



Scheme 6 - Synthesis scheme of bicyclic dansyl-*C*-glucosides obtained as mixture of two diastereoisomers. Reagents and conditions: a) BSTFA, CH_3CN, reflux, 3 h, then AllylTMS, TMSOTf, 0°C to r.t., 16 h[12a]; b) NIS, DMF, r.t., 2 h[12a]; c) NaN_3, DMF, 80°C, 24 h[12a]; d) $Pd(OH)_2/C$, H_2 atm., MeOH, r.t., 12 h[12a]; d) dansyl chloride, Et_3N, MeOH, r.t., 3 h[13].

Actually, we noticed a good water solubility for the mixture, higher than for compound **1**. Furthermore, preliminary biological tests revealed a considerable anti-inflammatory effect both *in vivo* and *in vitro*: the change to a bicyclic scaffold does not seem to affect the biological activity (unpublished data). Therefore, we decided to perform a

synthesis with the aim to obtain the two separate diastereoisomers, in order to understand if the biological activity depends on only one stereoisomers, or if the difference in stereochemistry is not crucial for the effect. Moreover, the diastereoisomeric mixture resolution allows to assign the correct stereochemistry of the new chiral center. The synthetic plan is depicted below (Scheme 7).

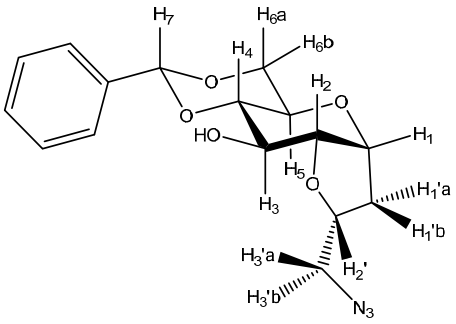


Scheme 7 - Synthesis of 6 and 7. Reagents and conditions: a) Benzaldehyde dimethylacetal, CSA, dry DMF, 70°C, 12 h, 78% (51% for 31a and 27% for 31b); b) $Pd(OH)_2/C$, H_2 atm., AcOH, r.t., 24 h, quant.; c) dansyl chloride, Et_3N, MeOH, r.t., 4 h (41% for 6 and 57% for 7).

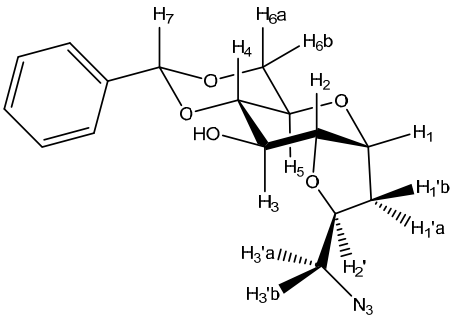
The separation of the two stereoisomers is performed protecting the hydroxyl in C4 and C6 positions of the bicyclic *C*-glycoside as benzylidene acetal. This allowed to an easily separation of the two stereoisomers with classic flash chromatography technique. The diastereoisomers were then use separately for the synthesis, performing the same transformations: a reduction of the azide into amino group with contemporary cleavage of the acetal protecting group and the attachment of the dansyl moiety. Compound **6** was obtained in higher amount compared to **7**, due to the diastereoisomeric excess of the iodocyclization reaction. To assign the correct stereochemistry of the new chiral center, NOESY analysis were performed on the azido intermediates **31a** and **31b** (Figure 40). In NOESY spectrum of **31b** an evident crosspeak between H2'and H3 is present, indicating their spatial proximity, which is possible for the diastereisomer with an (S)

configuration at C2'. This crosspeak is absent in the spectrum of the other diastereoisomer, as expected for the C2' R configuration, corresponding to compound **31a.**

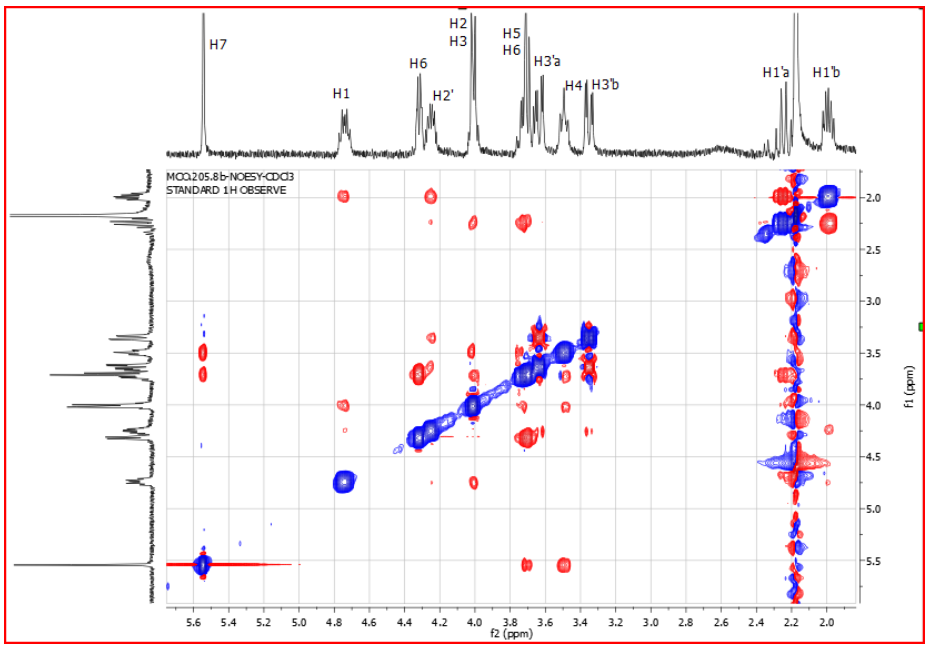
Compound 31a (R)



Compound 31b (S)



31a



31b

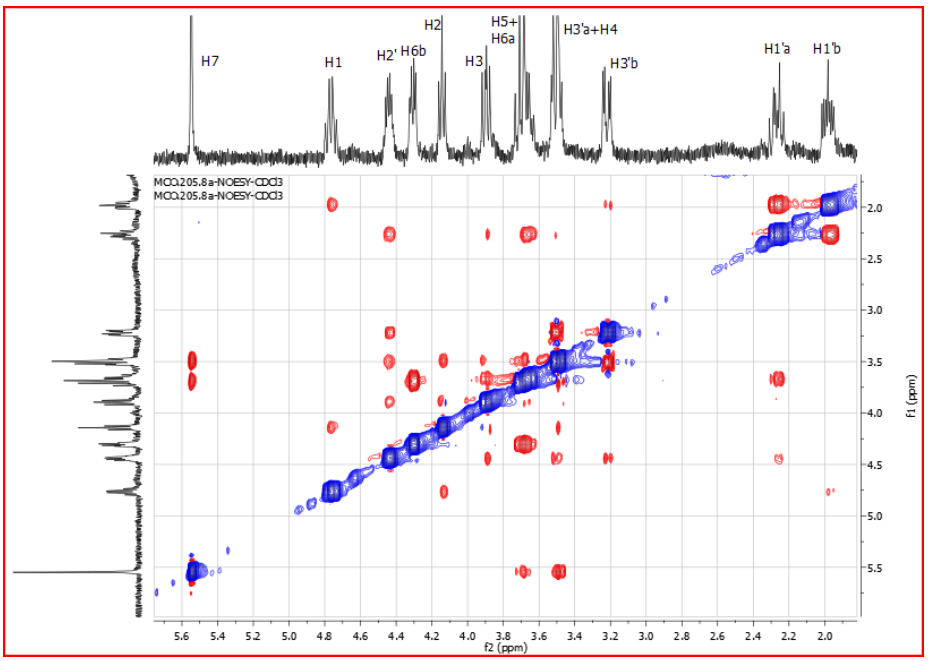


Figure 40 - NOESY experimnents on compounds 31a and 31b. In NOESY spectrum of 31b, a evident crosspeak between H2'and H3 is present, indicating their spatial proximity, which is possible for this diastereisomer and not for 31a.

Diastereoisomers differ from physical-chemical features: in our case the definition is perfectly in agreement with what observed. The water solubility is completely different between the two compound (**31a** is completely water-soluble, **31b** is completely water-unsoluble at concentration ranges of 0,5 – 5 mg/mL: the diastereoisomeric mixture previously prepared, containing the two molecules in 7:3 ratio, is soluble in water between these concentration ranges, since the 70% is constituted by **31a**). Surprisingly, we observed also a great difference in fluorescence of the two compound bearing the dansyl unit (dansyl chloride is used commonly as fluorescent derivatization reagent) (Figure 41).

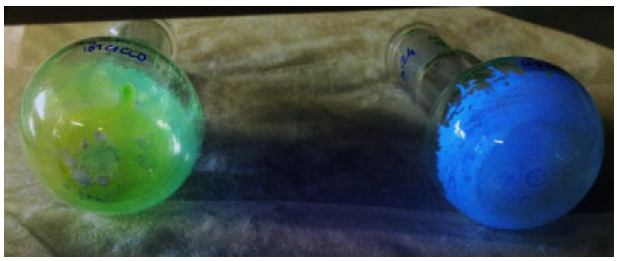


Figure 41 – Fluorescence emission (under a UVA lamp with a excitation wavelength of 365 nm) of compounds 31a (on the left) and 31b (on the right).

Compounds **1**-**6**, **30** (as mixture) and compounds **34**-**37** (Figure 42), which were previously synthesized and for which a biological assay was already performed[3], were tested for their protective activity on an *in vitro* model of doxorubicin-induced cytotoxicity, using SGLT1-overexpressing Caco-2 cell line (Figure 43). No test were conducted with compound **7** due to its insolubility in water.

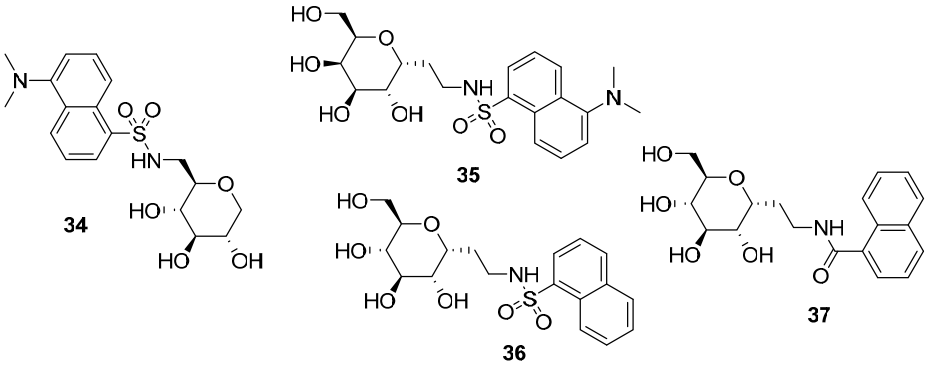


Figure 42 - Structure of compounds 34,35,36.

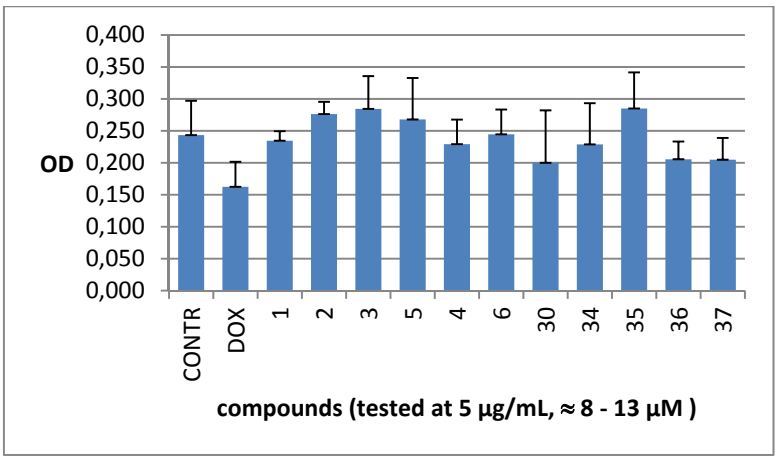


Figure 43 - Preliminary viability experiments with synthesized compounds. Neutral red assay for protective effect on Doxorubicin-induced cytotoxicity on Caco-2 cells overexpressing SGLT1.

Preliminary results seem to indicate that the presence of the aromatic moiety (dansyl of dimethylaminocinnamic residue) is important for the biological activity of the tested compounds, and less influence is associated to the sugar portion, since for almost all compounds a cytoprotection is observed at 5 µg/mL. However, with these preliminary results, the extent of the standard deviations are such as to not allow a reliable interpretation of the data. This experiment has thus to be performed again in order to obtain significant data. Meanwhile, these preliminary results suggested us to extend the members of the library including few compounds bearing the dansyl moiety and lacking the sugar entity. We decided to prepared other molecules bearing a dansyl residue attached to a completely different structure, in order to delucidate the importance of the different parts of the molecule involved in the biological effect and with the aim to obtain other information about the structure-activity relashionship. We planned the synthesis of the new compounds taking into account the probable water

solubility of the products. In fact, the synthesized molecules containing a dansyl attached on the amino functionality of glucosamine, on a hexaethylene glycol chain and on a glycine (Figure 44). The synthesis of **42** and **43** have involved a simple reaction between D-glucosamine and L-glycine, commercially available, and dansyl chloride, wherease the functionalization with a hexaethyleneglycol chain required the monomesylation of one of the two hydroxyl group, a following conversion to azide and its reduction into an amino functionality. These compounds, all water soluble, and the already synthesized ones will be tested for their cytoprotective activity.

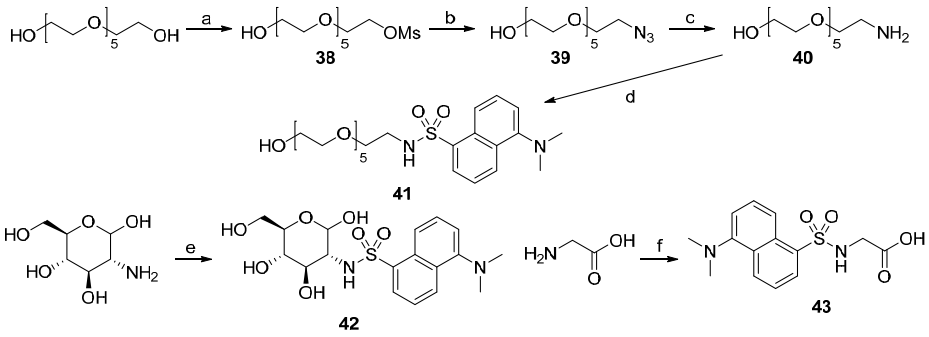


Figure 44 – Structure and synthesis of new prepared dansyl-derivative. Reagents and condition: a) Ag_2O, MsCl, CH_2Cl_2, r.t., 12 h, 47%; b) NaN_3, DMF, r.t., 36 h, 70%; c) H_2 atm., $Pd(OH)_2/C$, r.t., 15 h, quant.; d) dansyl chloride, Et_3N, MeOH, r.t., 3 h, 60%; e) dansyl chloride, Et_3N, MeOH, r.t., 12 h, 68%; f) dansyl chloride, NaOH 1M, THF, r.t., 3 h[14].

Conclusions

In this work we have described the design and synthesis of putative SGLT1 ligands as potential cytoprotective agents and molecular/chemical tools for studying the SGLT1-mediated cytoprotection process. The mechanism at the base of this phenomenon is still unclear, in particular how *C*-glucoside **1**, interacting with SGLT1, acts as anti-inflammatory/protective agent. Organic synthesis of its analogues can help this work, by preparing a library of molecules to use in structure-activity relationship studies. Testing the biological effect of analogues of compound **1,** it will be possible to build a putative pharmacophore map, thus identifying which groups are crucial for the binding to SGLT1, hence for the biological effect. The knowledge of these aspects represents a fundamental point; a rationalization of how the protective compounds bind to SGLT1 will be useful for the development of new and more active molecules.

Experimental Section

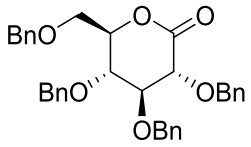
Cell cultures and treatments

Cell lines Caco-2/bbe and Caco-2/bbe permanently transfected with SGLT-1 (kind gift of Prof. Mark Donowitz, MD, Johns Hopkins University School of Medicine, Baltimore) were cultured in DMEM high-glucose medium (4,5g/L) (Euroclone, Pero, Italy) supplemented with 10% FBS (Euroclone), 1% glutamine (Euroclone), 15 nM sterile HEPES solution (Euroclone). For Caco-2/bbe/SGLT1 cells, 250 µg/ml G418 gentamicin bisulfate salt solution (Sigma-Aldrich) was added as antibiotic agent; for normal Caco-2/bbe penicillin/streptamicin solution (Euroclone) was added.

For viability assays 1 x 10^5 Caco-2 cells were plated on 96 flat-bottom well plates and treated after 24 hours with doxorubicin (100 µM) with the compounds at concentrations of 5 µg/mL (8 – 13 µM). Cell viability was evaluated after 48 h with Neutral Red assay kit (Sigma Aldrich), according to the datasheet.

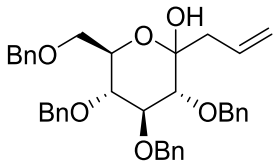
Synthesis of compounds

General remarks. All commercial chemicals were purchased from Sigma-Aldrich. All chemicals were used without further purification. All required anhydrous solvents were dried with molecular sieves for at least 24h prior to use. Thin layer chromatography (TLC) was performed on silica gel 60 F_{254} plates (Merck) with detection under UV light when possible, or by charring with a solution of $(NH_4)_6Mo_7O_{24}$ (21g), $Ce(SO_4)_2$ (1g), concentrated H_2SO_4 (31 mL) in water (500 mL) or with an ethanol solution of ninhydrin or with Dragendorff' spray reagent. Flash-column chromatography was performed on silica gel 230–400 mesh (Merck). ^1H and ^{13}C NMR spectra were recorded at 25°C, unless otherwise stated, with a Varian Mercury 400-MHz instrument. Chemical shift assignments, reported in parts per million, were referenced to the corresponding solvent peaks. Mass spectra were recorded on a QTRAP system with ESI source.

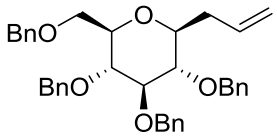


2,3,4,6-tetra-*O*-benzyl-D-gluconolactone **8**

The synthesis of **8** is described in Brenna, et al. [8].

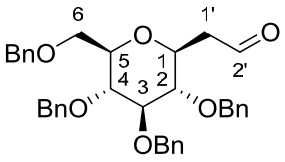


3-hydroxy-3-(2,3,4,6-tetra-*O*-benzyl-D-glucopyranosyl)-propene **9**



3-(2,3,4,6-tetra-*O*-benzyl-D-glucopyranosyl)-propene **10**

The synthesis of **9** and **10** are described in McGarvey, et al. [9].



2-(2,3,4,6-tetra-*O*-benzyl-β-D-glucopyranosyl)-ethanal **11**

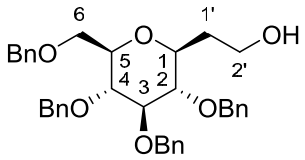
In a solution of **10** (1,42 g, 2,51 mmol) in CH_2Cl_2 (50 mL, 0,05 M) at -78°C, O3 was bubbled until a pale blue colour appears, indicating the end of the reaction. The reaction was followed by TLC (PE/AcOEt 8:2). Then the excess of O_3 was removed by purging the reaction with a stream of Argon at -78°C and then Ph_3P (6,15 mmol, 2,45 eq) were added. The

reaction was stirred for 24 h at r.t. The product was purified directly from the crude reaction by flash chromatography (PE/AcOEt 9:1 – PE/AcOEt 8,5:1,5) to provide aldehyde **11** (1,032 g, 1,82 mmol, 73%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 9.72 (s, 1H, C*H*O), 7.40 – 7.03 (m, 20H, C*H* Ar), 4.98 – 4.86 (m, 3H, CH_2Ph), 4.82 (d, *J* = 10.7 Hz, 1H, CH_2Ph), 4.63 – 4.55 (m, 3H, CH_2Ph), 4.51 (d, *J* = 12.2 Hz, 1H, CH_2Ph), 3.83 (td, *J* = 9.1, 4.5 Hz, 1H, H1), 3.77 – 3.63 (m, 4H, H2, H4, H6a, H6b), 3.52 – 3.44 (m, 1H, H5), 3.34 (t, *J* = 9.1 Hz, 1H, H3), 2.77 – 2.67 (m, 1H, H1'a), 2.57 (ddd, *J* = 16.2, 7.9, 2.4 Hz, 1H, H1'b).

^{13}C NMR (101 MHz, $CDCl_3$) δ 200.44 (C2' (*C*HO)), 138.47, 138.06, 137.69 (Cq Ar x 4), 129.79, 128.73, 128.68, 128.64, 128.58, 128.53, 128.34, 128.19, 128.05, 127.98, 127.92, 127.86, 127.83, 127.80 (C Ar x 20), 87.22, 81.25, 79.25, 78.39 (C1, C2, C3, C4, C5), 75.73, 75.21 (*C*H_2Ph x 4), 74.58 (C5), 68.75 (C6), 46.20 (C1').

$C_{36}H_{38}O_6$; calcd. mass 566,69; ESI-MS: m/z 567,61 $[M+H]^+$.



2-(2,3,4,6-tetra-*O*-benzyl-β-D-glucopyranosyl)-ethanol **12**

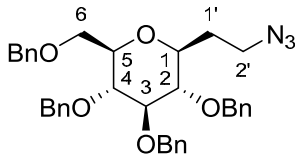
Compound **11** (736 mg, 1,3 mmol) was dissolved in a mixture of CH_2Cl_2/MeOH 1:1 (6 mL, 0,2 M) at r.t. $NaBH_4$ (5,19 mmol, 4 eq) was added and the reaction was stirred vigorously at r.t. The reaction was followed by TLC (PE/AcOEt 7:3). At the end of the reaction, the solvent was evaporated; the residue was suspended in a volume of a saturated solution of sodium carbonate and stirred vigorously for 30 min. The aqueous phase was extracted with AcOEt (3 x). The organic phases were combined, dried with sodium sulphate and filtrated. The product was

purified from the residue by flash chromatography (PE/AcOEt 7:3), to provide compound **11** (646 mg, 1,136 mmol, 90%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.38 – 7.23 (m, 18H, C*H* Ar), 7.23 – 7.13 (m, 2H, C*H* Ar), 4.93 – 4.87 (m, 3H, CH_2Ph), 4.83 (d, J = 10.8 Hz, 1H, CH_2Ph), 4.64 (d, J = 10.9 Hz, 1H, CH_2Ph), 4.59 – 4.49 (m, 3H, CH_2Ph), 3.79 (t, J = 5.4 Hz, 2H, H2'a,b), 3.73 – 3.64 (m, 2H, H6a, H3), 3.63 – 3.54 (m, 2H, H6b, H2), 3.54 – 3.44 (m, 2H, H1, H5), 3.35 (t, J = 9.2 Hz, 1H, H4), 2.12 – 2.00 (m, 1H, H1'a), 1.84 – 1.67 (m, 1H, H1'b).

^{13}C NMR (101 MHz, $CDCl_3$) δ 138.55, 138.02, 137.97, 137.95 (Cq Ar), 128.63, 128.61, 128.58, 128.56, 128.17, 128.11, 128.08, 127.99, 127.96, 127.86, 127.83 (C Ar x 20), 87.11 (C3), 81.85 (C2), 80.00 (C1), 78.72 (C4), 78.57 (C5), 75.75, 75.47, 75.21, 73.60 (*C*H_2Ph), 69.17 (C6), 61.68 (C2'), 33.80 (C1').

$C_{36}H_{40}O_6$; calcd. mass 568,61; ESI-MS: m/z 569,64 $[M+H]^+$.



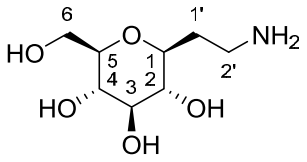
2-(2,3,4,6-tetra-*O*-benzyl-β-D-glucopyranosyl)-1-azidoethane **13**

Compound **12** (631 mg, 1,1 mmol) was dissolved in THF dry (5,5 mL, 0,2 M) and Ph_3P (3,3 mmol, 3 eq) was added. The solution was cooled to 0°C and DIAD (3,3 mmol, 3 eq) was added dropwise. After 10 min, 3,5 mmol (3.2 eq) of phosphoryldiphenil azide $(PhO)_2P(O)N_3$ are added at 0°C and the reaction was stirred for 5 h at r.t. The reaction was followed by TLC (PE/AcOEt 8:2) and the product was purified from the crude by flash chromatography (eluent PE/AcOEt 9:1) to provide compound **13** (621 mg, 1,03 mmol, 93%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.38 – 7.23 (m, 18H C*H* Ar), 7.20 – 7.13 (m, 2H C*H* Ar), 4.98 – 4.87 (m, 3H, CH_2Ph), 4.84 (d, *J* = 10.8 Hz, 1H, CH_2Ph), 4.69 – 4.61 (m, 2H, CH_2Ph), 4.58 (d, *J* = 10.8 Hz, 1H, CH_2Ph), 4.54 (d, *J* = 12.2 Hz, 1H, CH_2Ph), 3.76 – 3.64 (m, 4H, H2, H3, H4, H6a), 3.45 – 3.37 (m, 3H, H1,H5, H6b), 3.35 (dd, *J* = 9.2, 2.3 Hz, 1H, H2'a), 3.33 – 3.25 (m, 1H, H2'b), 2.17 – 2.04 (m, 1H, H1'a), 1.75 – 1.63 (m, 1H, H1'b).

^{13}C NMR (101 MHz, $CDCl_3$) δ 138.63, 138.23, 138.19, 138.00 (Cq Ar), 128.66, 128.61, 128.56, 128.52, 128.22, 128.09, 128.03, 127.95, 127.90, 127.87, 127.82 (C Ar), 87.36 (C3), 81.94 (C2), 78.93 (C1), 78.48 (C4), 76.39 (C5), 75.74, 75.39, 75.13, 73.62 (*C*H_2Ph), 68.95 (C6), 48.04 (C2'), 29.85 (C1').

$C_{36}H_{39}N_3O_5$; calcd. mass 593,72; ESI-MS: m/z 594,68 $[M+H]^+$.



2-(β-D-glucopyranosyl)-ethanamine **14**

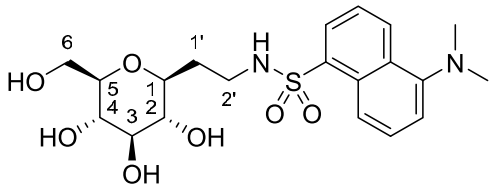
Azide **13** (610 mg, 1,026 mmol) was dissolved in a mixture of AcOEt/MeOH and the solution was degased under vacuum for 10 min. A catalytic amount $Pd(OH)_2$/C and 0,1 eq of AcOH were added and the reaction was stirred under H_2 atmosphere for 48 h. The reaction was followed by TLC (PE/AcOEt 8:2 and CH_2Cl_2/MeOH/NH_3 5:5:1) until the disappearance of the starting compound. The reaction was filtered through a celite pad and concentered. Compound **14** was obtained quantitatively.

^{1}H NMR (400 MHz, D_2O) δ 3.89 – 3.75 (m, 1H, H6a), 3.63 (dt, *J* = 12.1, 4.5 Hz, 1H, H5, H6b), 3.34 (ddt, *J* = 18.4, 12.7, 8.3 Hz, 4H, H1, H2, H4), 3.25 –

3.05 (m, 3H, H2'a,b, H3), 2.25 – 2.07 (m, 1H, H1'a), 1.84 – 1.69 (m, 1H, H1'b).

^{13}C NMR (101 MHz, D_2O) δ 79.27 (C5), 77.47 (C1), 76.94 (C3), 72.88 (C2), 69.49 (C4), 60.62 (C6), 45.20 (C2'), 36.95 (C1').

$C_8H_{17}NO_5$; calcd. mass 207,13; ESI-MS: m/z 208,20 $[M+H]^+$.



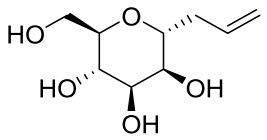
5-(N,N-Dimethylamino)-N-[2-(β-d-glucopyranosyl)ethyl]-1-naphthalensulfonamide **2**

To a solution of compound **14** (164 mg, 0,792 mmol) in a mixture of H_2O/THF 1:1 (10 mL, 0,08 M), 222 µL of Et_3N (1,583 mmol, 2 eq) are added at r.t.. After 10 min, 320 mg (1,187 mmol, 1,5 eq) of dansyl chloride are added and the reaction was stirred at r.t. for 3 h. The reaction was followed by TLC (CH_2Cl_2/MeOH/NH_3 5:5:1 and AcOEt/MeOH 9:1) and the product was purified by flash chromatography (eluent AcOEt/MeOH 9,5:0,5). 123 mg (0,28 mmol, 35%) of compound **2** are obtained.

^{1}H NMR (400 MHz, CD_3OD) δ 8.56 (d, *J* = 8.5 Hz, 1H, C*H* Ar), 8.36 (d, *J* = 8.7 Hz, 1H, C*H* Ar), 8.20 (d, *J* = 7.3 Hz, 1H, C*H* Ar), 7.64 – 7.51 (m, 2H, C*H* Ar), 7.27 (d, *J* = 7.5 Hz, 1H, C*H* Ar), 3.62 (dd, *J* = 11.8, 2.2 Hz, 1H, H6a), 3.50 (dd, *J* = 11.8, 5.5 Hz, 1H, H6b), 3.21 – 3.09 (m, 2H, H4, H3), 3.09 – 2.96 (m, 3H, H1, H2'a,b), 2.95 – 2.79 (m, 8H, (CH_3)$_2$N-, H5, H2), 2.04 – 1.91 (m, 1H, H1'a), 1.50 – 1.37 (m, 1H, H1'b).

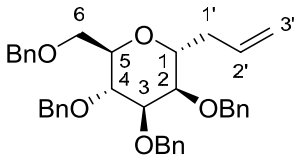
^{13}C NMR (101 MHz, CD_3OD) δ 153.20 (Cq Ar), 136.78 (Cq Ar), 131.25 (Cq Ar), 131.13 (C Ar), 130.94 (Cq Ar), 130.34 (C Ar), 129.08 (C Ar), 124.30 (C

Ar), 120.66 (C Ar), 116.45 (C Ar), 81.18 (H4), 79.64 (H3), 78.08 (C1), 75.26 (C2), 71.62 (C5), 62.78 (C6), 45.83 ((*C*H_3)$_2$N- x2), 40.58 (C2'), 32.90 (C1').
$C_{10}H_{28}N_2O_7S$; calcd. mass 440,51; ESI-MS: m/z 441,6 $[M+H]^+$.



C-allyl-α-D-mannopyranoside **15**

To a suspension of *O*-Me-α-D-mannopyranoside (2,024 g, 10,42 mmol) in CH_3CN dry (5 mL) were added 8,3 mL (31,27 mmol, 3 eq) of BSTFA; the suspension was heated to reflux and stirred for 3 h or otherwise when the sugar is completely dissolved. The reaction was then cooled to 0°C and AllylTMS and TMSOTf were added dropwise. The reaction was stirred at r.t. for 16 h and followed by TLC (AcOEt/MeOH 8:2). At the end of the reaction, 20 mL of water were added to the solution at 0°C; the reacxtion was neutralized with the addition of basic resin that was filtered off. The solution is concentrated and then washed with AcOEt. The aqueous phase containing the compound is evaporated and the residue (≈ 4 g) is directly used for the following step.
All spectral data was consistent with that reported in the literature[15].



3-(2,3,4,6-tetra-*O*-benzyl-D-mannopyranosyl)-propene **15**

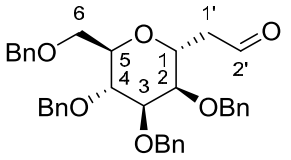
Crude compound **15** (≈ 4 g) was dissolved in dry DMF under argon. 2,7 g (68 mmol, 6,5 eq) of NaH (60% dispersion in mineral oil) are added at r.t.. After 10 min, benzyl bromide (7,4 mL, 63 mmol, 6 eq) is added

dropwise to the reaction that was stirred at r.t. for 5 h and then heated for 80°C for 24 h. The reaction is followed by TLC (PE/AcOEt 9:1) until the disappearance of starting material. The reaction was quenched with the addition of MeOH. The solution is then diluted with AcOEt and washed with H_2O (3x). The organic phase is then dried with sodium sulphate and evaporated. The crude is purified by flash chromatography (eluent PE/AcOEt 10:1), to provide compound **16** (4,37 g, 7,74 mmol, 75% over two steps).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.43 – 7.15 (m, 20H, C*H* Ar), 5.77 (dt, *J* = 16.4, 6.9 Hz, 1H, H8), 5.08 – 4.99 (m, 2H, H9a,b), 4.72 (d, *J* = 11.3 Hz, 1H, CH_2Ph), 4.64 – 4.50 (m, 7H, CH_2Ph), 4.06 (dd, *J* = 12.2, 6.3 Hz, 1H, H1), 3.90 – 3.82 (m, 2H, H4, H6a), 3.82 – 3.76 (m, 2H, H3, H5), 3.72 (dd, *J* = 10.2, 3.4 Hz, 1H, H6b), 3.64 (dd, *J* = 4.5, 3.1 Hz, 1H, H2), 2.43 – 2.28 (m, 2H, H1'a,b).

^{13}C NMR (101 MHz, $CDCl_3$) δ 138.57, 138.40, 138.39, 138.30 (Cq Ar), 134.45 (C8), 128.50, 128.44, 128.42, 128.14, 128.11, 128.01, 127.85, 127.82, 127.75, 127.59 (C Ar x 20), 117.32 (C9), 77.01 (C3), 75.27 (C2), 75.02 (C4), 73.96 (*C*H_2Ph), 73.84 (C5), 73.42 (*C*H_2Ph), 72.46 (C1), 72.19 (*C*H_2Ph), 71.65 (*C*H_2Ph), 69.29 (C6), 34.80 (C1').

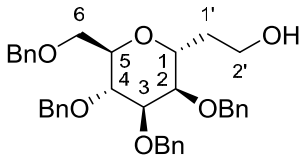
$C_{37}H_{40}O_5$; calcd. mass 564,72; ESI-MS: m/z 565,7 $[M+H]^+$.



2-(2,3,4,6-tetra-*O*-benzyl-α-D-mannopyranosyl)-ethanal **16**

A solution of compound **16** (2,27 g, 4 mmol) in dry CH_2Cl_2 (30 mL, ≈ 0,1 M) is cooled to -78°C and ozone was bubbled through the solution for approximately 3 h, until the solution turned blue. The reaction is then

deozonized bubbling argon through the solution and Ph_3P (3,3 g, 12,5 mmol, 3 eq) was added at -78°C. The solution is slowly warmed to r.t. and stirred overnight. The reaction was followed by TLC (PE/AcOEt 8:2). The crude product was directly used for the next reaction without further purifications. A ^{1}H-NMR performed on the crude confirmed the complete transformation of the allyl group into aldehyde function.



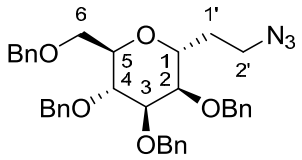
2-(2,3,4,6-tetra-*O*-benzyl-α-D-mannopyranosyl)-ethanol **18**

Crude aldehyde **17** was dissolved in a mixture of CH_2Cl_2/MeOH 1:1 (80 mL, ≈ 0,05 M). $NaBH_4$ (605 mg, 16 mmol, 4 eq related to compound **16**) was added at r.t. and the reaction is stirred at r.t. for 2 h. The product was extracted as for compound 11 and purified by flash chromatography (PE/AcOEt 6:4). 1363 mg (2,4 mmol, 60% yield over two steps) of alcohol **18** were obtained.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.39 – 7.23 (m, 18H, C*H* Ar), 7.23 – 7.16 (m, 2H, C*H* Ar), 4.63 – 4.45 (m, 8H, CH_2Ph), 4.21 – 4.13 (m, 1H, H1), 4.01 – 3.94 (m, 1H, H5), 3.84 – 3.74 (m, 4H, H3, H6a, H2'a,b), 3.71 (t, *J* = 5.7 Hz, 1H, H4), 3.66 – 3.55 (m, 2H, H2, H6b), 2.71 (s, 1H, O*H*), 1.93 – 1.81 (m, 1H, H1'a), 1.80 – 1.71 (m, 1H, H1'b).

^{13}C NMR (101 MHz, $CDCl_3$) δ 138.13, 138.11, 137.69, 137.66 (Cq Ar), 129.84, 128.56, 128.53, 128.52, 128.47, 128.11, 128.00, 127.92, 127.91, 127.88, 127.77 (C Ar), 76.29 (C2), 76.00 (C3), 74.91 (C4), 73.63 (C5), 73.40 (*C*H_2Ph), 72.49 (*C*H_2Ph), 72.25 (C1), 71.74 (*C*H_2Ph), 68.68 (C6), 61.56 (C2'), 32.25 (C1').

$C_{36}H_{40}O_6$; calcd. mass 568,71; ESI-MS: m/z 569,6 $[M+H]^+$, 591,5 $[M+Na]^+$.



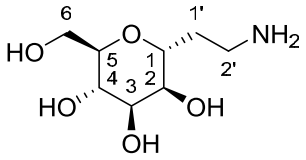
2-(2,3,4,6-tetra-*O*-benzyl-α-D-mannopyranosyl)-1-azidoethane **19**

Compound **18** (653 mg, 1,148 mmol) was converted into azido derivative **19** following the same procedure for compound **13**. The crude product was purified by flash chromatography (eluent PE/AcOEt 9:1) to afford 549 mg (0,925 mmol, 80%) of azido derivative **19**.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.39 – 7.19 (m, 20H C*H* Ar), 4.63 (d, *J* = 11.7 Hz, 1H, C*H* Ar), 4.60 – 4.50 (m, 7H, C*H* Ar), 4.09 – 4.02 (m, 1H, H1), 3.91 (dd, *J* = 10.6, 4.8 Hz, 1H, H5), 3.86 – 3.77 (m, 3H, H3, H6a, H4), 3.73 (dd, *J* = 10.1, 4.6 Hz, 1H, H6b), 3.58 (dd, *J* = 6.3, 2.8 Hz, 1H, H2), 3.49 – 3.32 (m, 2H, H2'a,b), 1.93 – 1.73 (m, 2H, H1'a,b).

^{13}C NMR (101 MHz, $CDCl_3$) δ 138.35, 138.14, 138.07, 138.02 (Cq Ar), 128.52, 128.50, 128.45, 128.12, 128.03, 128.00, 127.88, 127.81, 127.68 (C Ar), 75.95 (C2), 75.55 (C3), 74.54 (C4), 74.05 (C5), 73.40 (*C*H_2Ph), 73.27 (*C*H_2Ph), 72.36 (*C*H_2Ph), 71.52 (*C*H_2Ph), 68.89 (C1), 68.76 (C6), 48.06 (C2'), 29.93 (C1').

$C_{36}H_{39}N_3O_5$; calcd. mass 593,72; ESI-MS: m/z 594,6 $[M+H]^+$.



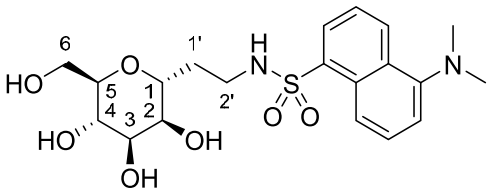
2-(α-D-mannopyranosyl)-ethanamine **20**

Compound **19** (493 mg, 0,83 mmol) was dissolved in a mixture of AcOEt/MeOH and the solution was degased under vacuum for 10 min. A catalytic amount $Pd(OH)_2$/C and 0,1 eq of AcOH were added and the reaction was stirred under H_2 atmosphere for 72 h. The reaction was

followed by TLC (PE/AcOEt 9:1 and AcOEt/MeOH/H_2O/AcOH 5:5:1:1) until the disappearance of the starting compound. The reaction was filtered through a celite pad and concentered. 110 mg (0,53 mmol, 64%) of compound **20** were obtained.

^{1}H NMR (400 MHz, D_2O) δ 4.04 – 3.91 (m, 1H), 3.89 – 3.81 (m, 1H), 3.81 – 3.74 (m, 2H), 3.74 – 3.67 (m, 1H), 3.60 (t, *J* = 8.8 Hz, 1H), 3.51 (t, *J* = 6.7 Hz, 1H), 3.17 – 3.01 (m, 2H), 2.16 (dt, *J* = 20.1, 11.3 Hz, 1H), 1.96 – 1.83 (m, 1H).

$C_8H_{17}NO_5$; calcd. mass 207,23; ESI-MS: m/z 208,2 $[M+H]^+$.

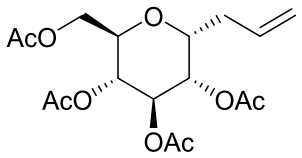


5-(N,N-Dimethylamino)-N-[2-(α-D-mannopyranosyl)ethyl]-1-naphthalensulfonamide **3**

To a solution of compound **20** (86 mg, 0,415 mmol) in MeOH (4 mL, 0,1 M), 116 µL of Et_3N (0,83 mmol, 2 eq) are added at r.t.. After 10 min, 168 mg (0,625 mmol, 1,5 eq) of dansyl chloride are added and the reaction was stirred at r.t. for 3 h. The reaction was followed by TLC (AcOEt/MeOH/H_2O/AcOH 5:5:1:1 and AcOEt/MeOH 9:1) and the product was purified by flash chromatography (eluent AcOEt/MeOH 8,5:1,5). 41 mg (0,1 mmol, 25%) of compound **3** are obtained.

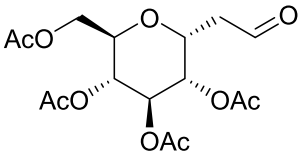
^{1}H NMR (400 MHz, CD_3OD) δ 8.56 (d, *J* = 8.5 Hz, 1H, C*H* Ar), 8.34 (d, *J* = 8.7 Hz, 1H, C*H* Ar), 8.20 (d, *J* = 7.2 Hz, 1H, C*H* Ar), 7.63 – 7.54 (m, 2H, C*H* Ar), 7.27 (d, *J* = 7.5 Hz, 1H, C*H* Ar), 3.86 – 3.79 (m, 1H, H1), 3.71 – 3.64 (m, 2H, H6a,b), 3.58 – 3.45 (m, 3H, H2, H3, H4), 3.35 (m, 1H, H5), 2.95 (t, *J* = 6.9 Hz, 2H, H2'a,b), 2.88 (s, 6H, $(CH_3)_2$N-), 1.86 – 1.73 (m, 1H, H1'a), 1.62 – 1.48 (m, 1H, H1'b).

^{13}C NMR (101 MHz, D_2O) δ 150.57 (Cq Ar), 133.01 (Cq Ar), 130.21 (C Ar), 130.07 (C Ar), 128.87 (Cq Ar), 128.75 (Cq Ar), 128.64 (C Ar), 123.95 (C Ar), 119.28 (C Ar), 116.11 (C Ar), 78.57 (C5), 77.03 (C1), 75.35 (C2), 72.86 (C4), 68.96 (C3), 59.92 (C6), 44.86 ((*C*H_3)$_2$N-), 38.14 (C2'), 30.07 (C1').
$C_{20}H_{28}N_2O_7S$; calcd. mass 440,51; ESI-MS: m/z 441,6 $[M+H]^+$.



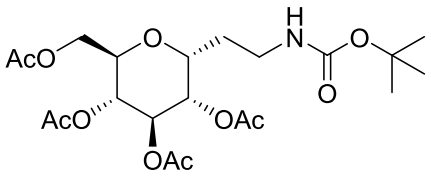
3-(2,3,4,6-tetra-*O*-acetyl-α-D-glucopyranosyl)-propene **21**

The synthesis and characterization for this compound are described in McGarvey, et al. [9]. All NMR data are in agreement with those published.



2-(2,3,4,6-tetra-*O*-acetyl-α-D-glucopyranosyl)-ethanal **22**

The synthesis and characterization for this compound are described in Stepanek, et al. [11] and Abdel-Rahman, et al. [16]. All NMR results are consistent with published data.



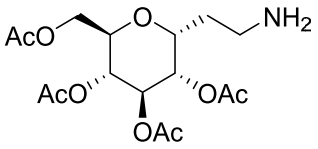
2-(2,3,4,6-tetra-*O*-acetyl-α-D-glucopyranosyl)-tert-butoxycarbonylaminoethane **23**

Compound **22** (800 mg, 2,14 mmol) was dissolved in CH_3CN dry (10 mL) under argon atmosphere. Tert-butyl carbamate (6,42 mmol, 3 eq.), triethylsilane (6,42 mmol, 3 eq.) and TFA (6,21 mmol, 2,9 eq) were successively added to the solution, and the reaction, followed by TLC (PE/AcOEt 7:3), was stirred at r.t.. After 3 h, Et_3N was added to neutralize the reaction and the solvent evaporated. The residue was purified by FC (eluent PE/AcOEt 7:3), affording compound 23 (1,52 mmol, 71% yield).

^1H NMR (400 MHz, $CDCl_3$) δ = 5.51 (m, 1H), 5.18 (dd, *J*=18.9, 9.9, 1H), 5.00 – 4.92 (m, 1H), 4.92 – 4.84 (m, 1H), 4.27 – 4.12 (m, 2H), 4.01 (dt, *J*=12.7, 6.6, 2H), 3.24 – 3.00 (m, 1H), 1.99 (m, 12H), 1.65 (s, 1H), 1.35 (s, 9H), 1.16 (t, *J*=7.0, 2H).

^{13}C NMR (101 MHz, $CDCl_3$) δ = 170.79 (s), 170.10 (s), 169.66 (s), 155.99 (s), 70.97 (s), 70.20 (s), 70.10 (s), 69.79 (s), 69.40 (s), 68.64 (s), 62.21 (s), 62.03 (s), 37.28 (s), 29.91 – 29.69 (m), 28.89 (s), 28.54 (s), 28.48 (s), 28.41 (s), 25.98 (s), 20.80 (s), 14.33 (s).

$C_{21}H_{33}NO_{11}$; calcd. mass 475,49; ESI-MS: m/z 476,5 $[M+H]^+$, 498,5 $[M+Na]^+$.

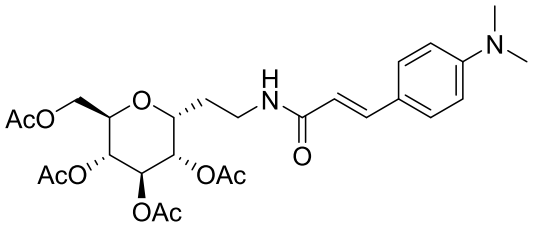


2-(2,3,4,6-tetra-*O*-acetyl-α-D-glucopyranosyl)-ethaneamine **24**

To a solution of **23** (310 mg, 0,65 mmol) in CH_2Cl_2 (2,5 mL), 2,5 mL of trifluoroacetic acid were added. The reaction was followed by TLC (AcOEt/MeOH/H_2O/AcOH 7:3:1:1) and after 2 h the solvent was evaporated. Compound **24** (quant. yield) was obtained directly without further purifications.

^{1}H NMR (400 MHz, $CDCl_3$) δ = 5.16 (t, *J*=8.9, 1H), 4.92 (dd, *J*=9.2, 5.7, 1H), 4.82 (dd, *J*=15.9, 7.1, 1H), 4.27 – 4.02 (m, 3H), 4.01 – 3.89 (m, 2H), 3.89 – 3.68 (m, 4H), 2.82 – 2.62 (m, 2H), 1.91 (dd, *J*=14.7, 6.4, 12H), 1.54 (dd, *J*=7.3, 4.2, 1H).

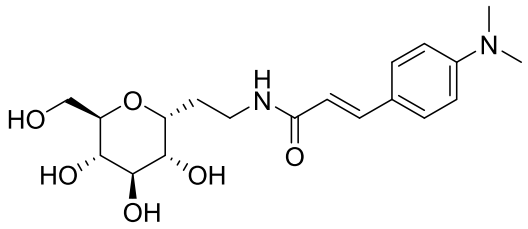
$C_{16}H_{25}NO_9$; calcd. mass 375,37; ESI-MS: m/z 376,5 $[M+H]^+$.



(E)-3-(4-(dimethylamino)phenyl)-N-(2,3,4,6-tetra-*O*-acetyl-α-D-glucopyranosyl)-acrylamide **25**

255 mg of compound **24** (0,68 mmol) were dissolved in 3 mL of DMF dry under argon atmosphere. Then N,N-dimethylaminocinnamic acid, HOBt, DIPEA were added; the reaction is cooled to 0°C and DIC was added. The reaction was stirred at r.t. and followed by TLC (PE/AcOEt 5:5). When TLC indicated the disappereance of the starting material, the crude was purified by TLC (PE/AcOEt 3:7), affording compound 25 (175 mg, 0,32 mmol, 47%).

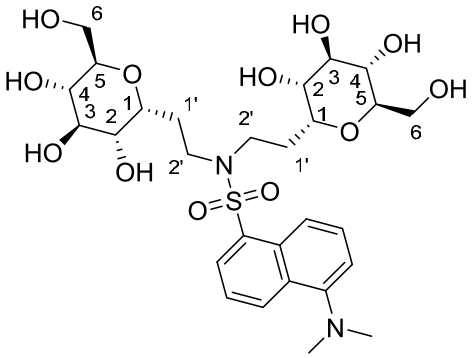
^{1}H NMR (400 MHz, $CDCl_3$) δ = 7.54 (d, *J*=15.5, 1H), 7.39 (t, *J*=11.3, 2H), 6.65 (d, *J*=8.7, 2H), 6.17 (d, *J*=15.5, 1H), 5.26 (t, *J*=8.4, 1H), 5.04 (dd, *J*=8.7, 5.5, 1H), 4.95 (t, *J*=8.3, 1H), 4.35 (dd, *J*=12.2, 5.8, 1H), 4.24 (s, 1H), 4.10 (dd, *J*=8.9, 5.8, 1H), 3.97 (s, 1H), 3.82 (d, *J*=6.6, 1H), 3.57 (dd, *J*=13.3, 6.7, 1H), 3.33 (dd, *J*=13.7, 6.8, 1H), 2.99 (s, *J*=6.5, 6H), 2.09 (s, 2H), 2.04 (s, 2H), 2.03 (s, 3H), 2.02 (s, 5H), 1.78 (d, *J*=14.9, 2H).
$C_{27}H_{36}N_2O_{10}$; calcd. mass 548,59; ESI-MS: m/z 549,6 $[M+H]^+$.



(E)-3-(4-(dimethylamino)phenyl)-N-(α-D-glucopyranosyl)-acrylamide **4**

To a solution of **25** (55 mg, 0,1 mmol) in MeOH (1 mL), 0,035 mL of a 1,45 M solution of MeONa (0,05 mmol, 0,5 eq) were added, and the reaction stirred at r.t.. until TLC (CH_2Cl_2/MeOH 8:2) shows the consumption of the starting material. Then, the reaction was neutralized with the addition of acdici resin (IRA-120 H+) which was removed; the resulting solution was concentrated and the residue purified by FC (eluent CH_2Cl_2/MeOH 8:2), obtaining product 4 (9 mg, 0,024 mmol, 24 % yield).

^{1}H NMR (400 MHz, d_6-DMSO) δ = 7.35 (d, *J*=8.7, 2H), 7.28 (d, *J*=15.6, 1H), 6.69 (d, *J*=8.8, 2H), 6.35 (d, *J*=15.7, 1H), 3.83 (d, *J*=7.6, 1H), 3.65 (d, *J*=11.5, 1H), 3.56 (s, 1H), 3.47 – 3.36 (m, 2H), 3.26 (ddd, *J*=20.2, 15.4, 7.4, 4H), 3.13 (s, 1H), 3.01 (t, *J*=8.9, 1H), 1.75 (d, *J*=7.2, 2H).
$C_{19}H_{28}N_2O_6$; calcd. mass 380,44; ESI-MS: m/z 381,3 $[M+H]^+$.



5-(N,N-Dimethylamino)-N,N-di-[2-(α-D-glucopyranosyl)ethyl]-1-naphthalensulfonamide **5**

Compound **5** was isolated as a by-product during the synthesis depicted in Scheme 5 (see Results and Discussion); 2 g of crude amine **24** (containing **24 bis** as side product) were dissolved in dry CH_2Cl_2 (15 mL). An excess of Et_3N and dansyl chloride were then added and the reaction stirred at r.t.. The reaction was followed by TLC (PE/AcOEt 6:4 and AcOEt/MeOH/AcOH 6:3:1). After 2 h, the reaction was concentrated and the residue subjected to FC (PE/AcOEt 6:4). 1,07 g (1,1 mmol) of the peracetylated dimer and 600 mg (0,99 mmol) of the monomer derivative (compound **1** in peracetylated form) were obtained, as confirmed by ESI-MS analysis (m/z 967,3 $[M+H]^+$ for the dimer and 609,2 $[M+H]^+$ for the monomer). The dimer was directly deprotected with MeONa (6 mL of 1 M solution in MeOH) in MeOH and CH_2Cl_2 (10 + 5 mL). After 1 h the pH was adjusted to neutrality adding IRA-120 H^+ resin. After its removal by filtration, the filtrate was concentrated, to obtain **5** with quantitative yield.

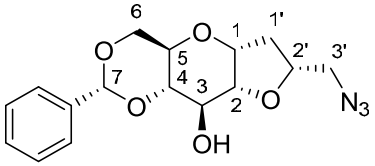
^{1}H NMR (400 MHz, CD_3OD) δ 8.56 (d, *J* = 8.4 Hz, 1H, C*H* Ar), 8.29 (d, *J* = 8.7 Hz, 1H, C*H* Ar), 8.21 (d, *J* = 7.3 Hz, 1H, C*H* Ar), 7.63 – 7.53 (m, 2H, C*H* Ar), 7.25 (d, *J* = 7.5 Hz, 1H, C*H* Ar), 3.94 – 3.81 (m, 2H, H1 (x2)), 3.75 (d, *J*

= 11.7 Hz, 2H, H6a (x2)), 3.65 – 3.47 (m, 6H, H6b (x2), H2 (x2), H2'a (x2)), 3.41 (t, *J* = 9.0 Hz, 2H, H3 (x2)), 3.37 – 3.27 (m, 4H, H5 (x2), H2'b (x2)), 3.20 (t, *J* = 9.1 Hz, 2H, H4 (x2)), 2.87 (s, 6H, (C*H*$_3$)$_2$N-), 2.03 – 1.75 (m, 4H, H1'a,b (x2)).

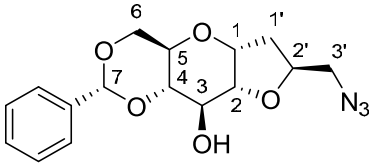
^{13}C NMR (101 MHz, CD_3OD) δ 153.15 (Cq Ar), 136.13 (Cq Ar), 131.52 (C Ar), 131.32 (Cq Ar), 131.25 (Cq Ar), 130.91 (C Ar), 129.18 (C Ar), 124.38 (C Ar), 120.64 (C Ar), 116.45 (C Ar), 75.16 (C3), 74.88 (C1), 74.53 (C5), 72.62 C2), 72.12 (C4), 62.99 (C6), 45.82 ((*C*H$_3$)$_2$N-), 45.71 (C2'), 24.92 (C1').

$C_{28}H_{42}N_2O_{12}S$; calcd. mass 630,71; ESI-MS: m/z 631,3 $[M+H]^+$ 653,3 $[M+Na]^+$.

The synthesis of compound **26**-**30** was already described in Rumio, et al. [13] and La Ferla, et al. [12a].



(2R,4aR,5aR,7R,8aR,9R,9aS)-7-(azidomethyl)-2-phenyloctahydrofuro[2',3':5,6]pyrano[3,2-d][1,3]dioxin-9-ol **31a**



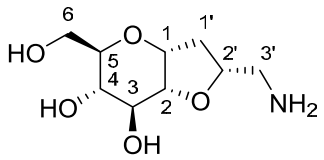
(2R,4aR,5aR,7S,8aR,9R,9aS)-7-(azidomethyl)-2-phenyloctahydrofuro[2',3':5,6]pyrano[3,2-d][1,3]dioxin-9-ol **31 b**

Mixture **28** (1115 mg, 4,54 mmol) was dissolved in DMF dry (5 mL, ≈ 1 M). Benzaldehyde dimethylacetal (5,45 mmol, 1,2 eq) and camphorsulphonic acid (2,27 mmol, 0,5 eq) are added at r.t.. The

reaction was stirred at 70°C for 12 h, and followed by TLC (PE/AcOEt 2:8). Et_3N (5 mL) was added to neutralize the acidity of the reaction and the solvent was evaporated. The crude was purified by flash chromatography (eluent PE/AcOEt 6:4 to PE/AcOEt 1:1), obtaining compound **31a** (777 mg, 2,33 mmol, 51% yield) and compound **31b** (405 mg, 1,21 mmol, 27% yield) with an overall yield of 78%.

31 a) ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.49 (dd, *J* = 6.5, 3.0 Hz, 2H, C*H* Ar), 7.42 – 7.32 (m, 3H, C*H* Ar), 5.52 (s, 1H, H7), 4.73 (dd, *J* = 15.9, 8.9 Hz, 1H, H1), 4.39 (dq, *J* = 8.3, 4.1 Hz, 1H, H2'), 4.29 (dd, *J* = 9.6, 4.1 Hz, 1H, H6a), 4.11 (t, *J* = 7.1 Hz, 1H, H2), 3.84 (dd, *J* = 9.6, 7.3 Hz, 1H, H3), 3.68 (t, *J* = 9.8 Hz, 1H, H6a), 3.65 – 3.58 (m, 1H, H5), 3.52 – 3.39 (m, 2H, H4, H3'a), 3.18 (dd, *J* = 12.9, 4.2 Hz, 1H, H3'b), 2.81 (s, 1H, OH), 2.22 (dt, *J* = 13.2, 9.1 Hz, 1H, H1'a), 1.94 (ddd, *J* = 13.3, 9.0, 4.5 Hz, 1H, H1'b).
^{13}C NMR (101 MHz, $CDCl_3$) δ 137.08 (*C*q Ar), 129.39 (*C*H Ar), 128.45 (*C*H Ar), 126.35 (*C*H Ar), 101.89 (*C*HPh), 80.16 (*C*H), 79.97 (*C*H), 75.46 (*C*H), 75.30 (*C*H), 71.02 (*C*H), 69.09 (CH_2), 63.83 (*C*H), 54.93 (CH_2), 29.46 (CH_2).
$C_{16}H_{19}N_3O_5$; calcd. mass 333,34; ESI-MS: m/z 334,4 $[M+H]^+$.

31 b) ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.50 (dd, *J* = 6.6, 2.9 Hz, 2H, C*H* Ar), 7.42 – 7.33 (m, 3H, C*H* Ar), 5.51 (s, 1H, H7), 4.66 (dt, *J* = 11.1, 7.0 Hz, 1H, H1), 4.32 – 4.22 (m, 1H, H6), 4.16 (td, *J* = 9.9, 4.1 Hz, 1H, H2'), 4.00 – 3.87 (m, 2H, H(2), H3), 3.71 – 3.63 (m, 2H, H5, H6), 3.55 (dd, *J* = 13.0, 3.8 Hz, 1H, H3'), 3.44 (dd, *J* = 11.6, 6.6 Hz, 1H, H4), 3.28 (dd, *J* = 13.0, 4.2 Hz, 1H, H3'), 2.14 (dd, *J* = 22.3, 11.1 Hz, 1H, H1'a), 1.98 – 1.86 (m, 1H, H1'b).
^{13}C NMR (101 MHz, $CDCl_3$) δ 137.08 (*C*qAr), 129.25 (*C*HAr), 128.33 (*C*HAr), 126.31 (*C*HAr), 101.77 (*C*HPh), 79.84 (*C*H), 79.11 (*C*H), 75.94 (*C*H), 75.87 (*C*H), 73.36 (*C*H), 68.98 (CH_2), 64.00 (*C*H), 54.14 (CH_2), 29.71 (CH_2).
$C_{16}H_{19}N_3O_5$; calcd. mass 333,34; ESI-MS: m/z 334,4 $[M+H]^+$.



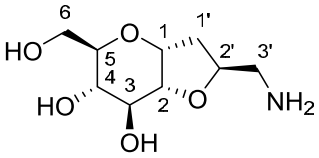
(2R,3aR,5R,6S,7S,7aR)-2-(aminomethyl)-5-(hydroxymethyl)hexahydro-2H-furo[3,2-b]pyran-6,7-diol **32**

To a solution of compound **31b** (94 mg, 0,28 mmol) in a mixture of MeOH/AcOEt (5 + 3 mL) were added few drops of glacial acetic acid and a catalytic amount of $Pd(OH)_2/C$. The reaction was stirred at r.t. under Hydrogen atmosphere until TLC (PE/AcOEt 1:1 and AcOEt/MeOH/H_2O/AcOH 6:4:1:1) indicated the absence of starting material. The reaction is then filtered through a pad of celite and the solution containing the product was concentrated. ^{1}H-NMR confirmed the purity of the deprotected bicyclic sugar that was directly used for the next reaction without further purification (60 mg, 0,27 mmol, quant. yield).

^{1}H NMR (400 MHz, CD_3OD) δ 4.61 (s, 1H, H1), 4.38 (d, *J* = 8.7 Hz, 1H, H2'), 3.98 – 3.90 (m, 1H, H2), 3.79 (dd, *J* = 12.0, 6.8 Hz, 1H, H6a), 3.74 – 3.63 (m, 2H, H3, H4), 3.60 – 3.52 (m, 1H, H5), 3.43 – 3.33 (m, 1H, H6b), 3.13 (dd, *J* = 13.0, 3.0 Hz, 1H, H3'a), 2.92 (dd, *J* = 13.1, 8.8 Hz, 1H, H3'b), 2.19 (ddd, *J* = 13.5, 6.3, 2.6 Hz, 1H, H1'a), 1.86 – 1.75 (m, 1H, H1'b).

^{13}C NMR (101 MHz, D_2O) δ 85.34 (*C*H), 80.54 (*C*H), 76.40 (*C*H), 76.09 (*C*H), 75.35 (*C*H), 70.23 (*C*H), 62.68 (CH_2), 45.17 (CH_2), 36.37 (CH_2).

$C_9H_{17}NO_5$; calcd. mass 219,24; ESI-MS: m/z 220,2 $[M+H]^+$.



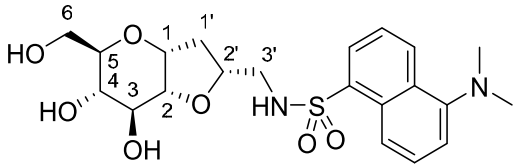
(2S,3aR,5R,6S,7S,7aR)-2-(aminomethyl)-5-(hydroxymethyl)hexahydro-2H-furo[3,2-b]pyran-6,7-diol **33**

The preparation of compound **33** was performed exactly as described for compound **32**. The catalytic hydrogenolysis (41 mg, 0,123 mmol) afforded 27 mg (0,123 mmol, quant. yield) of amine **33** that was used directly for the next reaction.

^{1}H NMR (400 MHz, CD_3OD) δ 4.46 (s, 1H, H1), 4.30 (dd, *J* = 8.7, 4.6 Hz, 1H, H2'), 4.01 (dd, *J* = 12.1, 8.3 Hz, 1H, H6a), 3.89 – 3.82 (m, 1H, H4), 3.75 (dt, *J* = 8.1, 3.6 Hz, 2H, H2, H5), 3.62 (dd, *J* = 12.1, 3.2 Hz, 1H, H6b), 3.54 (t, *J* = 4.9 Hz, 1H, H3), 3.13 (dd, *J* = 13.2, 3.6 Hz, 1H, H3'a), 3.05 (dd, *J* = 13.2, 5.0 Hz, 1H, H3'b), 2.51 – 2.38 (m, 1H, H1'a), 1.96 (s, 1H, H1'b).

^{13}C NMR (101 MHz, D_2O) δ 85.24 (*C*H), 80.52 (*C*H), 77.03 (*C*H), 75.40 (*C*H), 74.99 (*C*H), 70.02 (*C*H), 62.32 (*C*H_2), 45.98 (*C*H_2), 36.56 (*C*H_2).

$C_9H_{17}NO_5$; calcd. mass 219,24; ESI-MS: m/z 220,2 $[M+H]^+$.



N-(((2R,3aR,5R,6S,7S,7aR)-6,7-dihydroxy-5-(hydroxymethyl)hexahydro-2H-furo[3,2-b]pyran-2-yl)methyl)-5-(dimethylamino)naphthalene-1-sulfonamide **6**

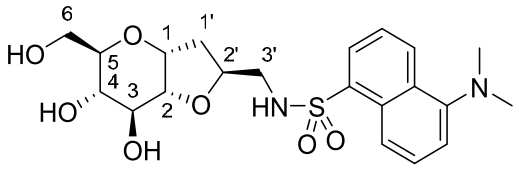
Biclyclic amine **32** (60 mg, 0,27 mmol) was dissolved in MeOH (5 mL); Et_3N (0,57 mmol, 2,1 eq) and dansyl chloride (0,41 mmol, 1,5 eq) were added and the reaction was stirred at r.t. until TLC (AcOEt/MeOH 9:1 and AcOEt/MeOH/H_2O/AcOH 6:4:1:1) indicated the end of the reaction. The

crude reaction was subjected to FC (eluent AcOEt/MeOH 9,5:0,5) which afforded 51 mg (0,11 mmol, 41% yield) of compound **6**.

^{1}H NMR (400 MHz, CD_3OD) δ 8.56 (d, *J* = 8.5 Hz, 1H, C*H* Ar), 8.36 (d, *J* = 8.7 Hz, 1H, C*H* Ar), 8.19 (dd, *J* = 7.3, 1.1 Hz, 1H, C*H* Ar), 7.64 – 7.53 (m, 2H, C*H* Ar), 7.28 (d, *J* = 7.6 Hz, 1H, C*H* Ar), 4.40 (t, *J* = 7.6 Hz, 1H, H1), 4.11 (s, 1H, H2'), 3.75 – 3.60 (m, 3H, H6a, H2, H6b), 3.50 (dd, *J* = 9.5, 6.5 Hz, 1H, H3), 3.46 (dd, *J* = 10.7, 4.0 Hz, 1H, H5), 3.28 – 3.20 (m, 1H, H4), 3.00 (dd, *J* = 13.7, 4.6 Hz, 1H, H3'a), 2.93 (dd, *J* = 13.7, 5.3 Hz, 1H, H3'b), 2.88 (s, 6H, -N(C*H*$_3$)$_2$), 1.91 (ddd, *J* = 13.2, 6.5, 3.6 Hz, 1H, H1'a), 1.77 – 1.63 (m, 1H, H1'b).

^{13}C NMR (101 MHz, CD_3OD) δ 153.15 (Cq Ar), 137.14 (Cq Ar), 131.16 (Cq Ar), 131.12 (*C* Ar), 130.90 (Cq Ar), 130.01 (C Ar), 129.10 (C Ar), 124.30 (C Ar), 120.54 (C Ar), 116.41 (C Ar), 84.35 (C2'), 79.21 (C1), 77.47 (C2), 74.94 (C4), 74.31 (C3), 69.52 (C5), 62.29 (C6), 47.46 (C3'), 45.80 ((*C*H$_3$)$_2$N-), 35.01 (C1').

$C_{21}H_{28}N_2O_7S$; calcd. mass 452,52; ESI-MS: m/z 453,4 $[M+H]^+$.



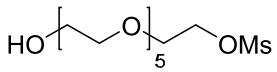
N-(((2S,3aR,5R,6S,7S,7aR)-6,7-dihydroxy-5-(hydroxymethyl)hexahydro-2H-furo[3,2-b]pyran-2-yl)methyl)-5-(dimethylamino)-naphthalene-1-sulfonamide **7**

Compound **33** was converted into dansyl-sulphonamide derivative **7** with the same procedure described for the diastereoisomer **32**. Compound **33** (27 mg, 0,123 mmol) were converted into compound **7** (32 mg, 0,07 mmol) with a yield of 57%.

^{1}H NMR (400 MHz, CD_3OD) δ 8.56 (d, *J* = 8.5 Hz, 1H, C*H* Ar), 8.36 (d, *J* = 8.7 Hz, 1H, C*H* Ar), 8.20 (dd, *J* = 7.3, 1.1 Hz, 1H, C*H* Ar), 7.59 (td, *J* = 8.8, 7.6 Hz, 2H, C*H* Ar), 7.28 (d, *J* = 7.2 Hz, 1H, C*H* Ar), 4.44 (dd, *J* = 12.0, 5.3 Hz, 1H, H1), 3.94 – 3.83 (m, 1H, H2'), 3.72 (dd, *J* = 12.0, 6.4 Hz, 1H, H6a), 3.64 (dd, *J* = 12.4, 3.2 Hz, 1H, H6b), 3.62 – 3.53 (m, 2H, H2, H3), 3.48 – 3.41 (m, 1H, H5), 3.36 – 3.27 (m, 1H, H4), 3.02 (dd, *J* = 5.4, 2.4 Hz, 2H, H3'a, b), 2.88 (s, 6H, -N(CH_3)$_2$), 2.00 (dt, *J* = 13.7, 7.0 Hz, 1H, H1'a), 1.71 (ddd, *J* = 13.1, 7.5, 5.3 Hz, 1H, H1'b).

^{13}C NMR (101 MHz, CD_3OD) δ 153.14 (Cq Ar), 136.95 (Cq Ar), 131.16 (Cq Ar), 131.14 (C Ar), 130.87 (Cq Ar), 130.06 (C Ar), 129.20 (C Ar), 124.28 (C Ar), 120.49 (C Ar), 116.43 (C Ar), 83.95 (C2'), 78.88 (C1), 78.23 (C2), 75.48 (C4), 74.92 (C3), 69.49 (C5), 62.30 (C6), 47.99 (C3'), 45.79 ((*C*H_3)$_2$N-), 34.88 (C1').

$C_{21}H_{28}N_2O_7S$; calcd. mass 452,52; ESI-MS: m/z 453,4 $[M+H]^+$.



Hexaethyleneglycol monomethylsulphonate **38**

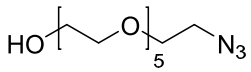
500 mg (1,77 mmol) of commercially available hexaethyleneglycol and 616 mg of Ag_2O (2,7 mmol, 1,5 eq) were dissolved in 18 mL (0,1 M) of CH_2Cl_2 under argon atmosphere and MsCl (2,12 mmol, 1,2 eq) was added to this suspension. The reaction was stirred at r.t. and followed by TLC (AcOEt/MeOH 9:1). After 12 h, the suspension was filtered through a celite pad to remove silver oxide and washed with CH2Cl2. The solution was concentrated and chromatographed on silica gel (eluent CH_2Cl_2/MeOH 9,5:0,5). 300 mg of monomesyl hexaethyleneglycol (0,83 mmol) were obtained with a yield of 47%.

^{1}H NMR (400 MHz, $CDCl_3$) δ 4.42 – 4.34 (m, 2H, CH_2OMs), 3.78 – 3.73 (m, 2H, CH_2OH), 3.73 – 3.69 (m, 2H, C$H_2$$CH_2$OMs), 3.69 – 3.61 (m, 16H, -

OC*H*$_2$C*H*$_2$O-), 3.61 – 3.55 (m, 2H, C*H*$_2$CH$_2$OH), 3.08 (s, 3H, C*H*$_3$SO$_2$), 2.77 (s, 1H, OH).

^{13}C NMR (101 MHz, $CDCl_3$) δ 72.70, 70.68, 70.65, 70.64, 70.62, 70.59, 70.32, 69.50 (CH_2 PEG x 10), 69.09 (*C*H$_2$OMs), 61.77 (*C*H$_2$OH), 37.82 (*C*H$_3$SO$_2$).

$C_{13}H_{28}O_9S$; calcd. mass 360,42; ESI-MS: m/z 361,2 $[M+H]^+$; 383,2 $[M+Na]^+$; 399,1 $[M+K]^+$



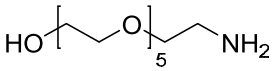
17-azido-3,6,9,12,15-pentaoxaheptadecan-1-ol **39**

To a solution of compound **38** (283 mg, 0,786 mmol) in DMF, NaN_3 (1,97 mmol, 2,5 eq) was added and the reaction was stirred at r.t. for 36 h. When TLC (AcOEt/MeOH 9,5:0,5) indicated the absence of starting material and the presence of a new spot with a Rf ≈ 0,4, the reaction was filtered through a celite pad to remove the excess of sodium azide and washed with CH_2Cl_2. The solution was then concentrated and the product purified by FC (AcOEt/MeOH 9,5:0,5) affording 165 mg (0,53 mmol, 70% yield) of compound **39**.

^{1}H NMR (400 MHz, $CDCl_3$) δ 3.75 – 3.69 (m, 2H), 3.69 – 3.62 (m, 18H), 3.62 – 3.56 (m, 2H), 3.38 (t, *J* = 5.0 Hz, 2H, C*H*$_2$N$_3$), 2.90 (s, 1H, OH).

^{13}C NMR (101 MHz, CD_3OD) δ 73.66, 71.62, 71.58, 71.54, 71.50, 71.38, 71.15 (*C*H$_2$ PEG x 10), 62.21 (*C*H$_2$OH), 51.76 (*C*H$_2$N$_3$).

$C_{12}H_{25}N_3O_9$; calcd. mass 307,35; ESI-MS: m/z 308,3 $[M+H]^+$; 330,3 $[M+Na]^+$; 346,3 $[M+K]^+$



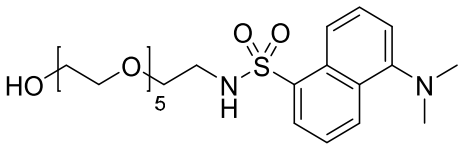
17-amino-3,6,9,12,15-pentaoxaheptadecan-1-ol **40**

Compound **39** (151 mg, 0,49 mmol) was dissolved in MeOH (5 mL, 0,1 M) and the solution was degased with a stream of argon for 10 min. A catalytic amount of $Pd(OH)_2/C$ was added and the suspension was vigorously stirred under a hydrogen atmosphere. The hydrogenation was followed by TLC (CH_2Cl_2/MeOH 9:1 and CH_2Cl_2/MeOH 7:3 + 1% NH_3 (aq)). After 15 h, the mixture was filtered on celite and the solvent was evaporated. The residue, containing the product **40** in a pure form, was used directly for the next reaction without further purification (quant. yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 3.75 – 3.69 (m, 2H, C*H*$_2$OH), 3.69 – 3.57 (m, 20H, C*H*$_2$ PEG), 3.32 (s, 3H, O*H*, N*H*$_2$), 2.98 – 2.90 (m, 2H, C*H*$_2$NH$_2$).

^{13}C NMR (101 MHz, CD_3OD) δ 73.67 (*C*H$_2$CH$_2$NH$_2$), 71.59, 71.54, 71.50, 71.39, 71.38, 71.33, 71.22, 71.05, 70.53 (*C*H$_2$ PEG x 9), 62.07 (*C*H$_2$OH), 41.37 (*C*H$_2$NH$_2$).

$C_{12}H_{27}NO_6$; calcd. mass 281,35; ESI-MS: m/z 282,3 $[M+H]^+$.



5-(dimethylamino)-N-(17-hydroxy-3,6,9,12,15-pentaoxaheptadecyl)naphthalene-1-sulfonamide **41**

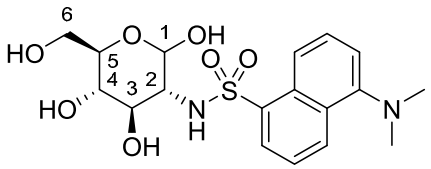
To a solution of compound **40** (100 mg, 0,356 mmol) in MeOH (4 mL, ≈ 0,1 M), Et_3N (0,534 mmol, 1,5 eq) and dansyl chloride (0,427 mmol, 1,3 eq) were added. The reaction, followed by TLC (CH_2Cl_2/MeOH 9:1 and CH_2Cl_2/MeOH 7:3 + 1% NH_3 (aq)) was stirred at r.t.. After 3 h, the solvent was removed in vacuo and the residue chromatographed (eluent

CH_2Cl_2/MeOH 9,5:0,5) affording 110 mg of compound **41** (0,214 mmol, 60%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 8.51 (d, *J* = 8.5 Hz, 1H, C*H* Ar), 8.32 (d, *J* = 8.6 Hz, 1H, C*H* Ar), 8.23 (dd, *J* = 7.3, 1.1 Hz, 1H, C*H* Ar), 7.63 – 7.44 (m, 2H, C*H* Ar), 7.17 (d, *J* = 7.2 Hz, 1H, C*H* Ar), 5.90 (s, 1H, N*H*SO_2), 3.75 – 3.69 (m, 2H, C*H*$_2$OH), 3.64 (d, J = 4.8 Hz, 10H, C*H*$_2$ PEG), 3.62 – 3.56 (m, 4H, C*H*$_2$ PEG), 3.52 – 3.44 (m, 2H, C*H*$_2$ PEG), 3.41 – 3.33 (m, 4H, C*H*$_2$ PEG), 3.11 (dd, J = 10.0, 5.2 Hz, 2H, C*H*$_2$ PEG), 2.88 (s, 6H, -N(C*H*$_3$)$_2$).

^{13}C NMR (101 MHz, D_2O) δ 150.94 (Cq Ar), 134.97 (Cq Ar), 129.67 (C Ar), 129.02 (Cq Ar), 128.95 (Cq Ar), 128.71 (C Ar), 128.28 (C Ar), 123.51 (C Ar), 118.99 (C Ar), 115.31 (C Ar), 71.62 (*C*H_2 PEG), 69.54 (*C*H_2 PEG), 69.45 (*C*H_2 PEG), 69.41 (*C*H_2 PEG), 69.37 (*C*H_2 PEG), 69.27 (*C*H_2 PEG), 69.04 (*C*H_2 PEG), 68.60 (*C*H_2 PEG), 60.23 (*C*H_2NHSO_2), 44.70 (-N(*C*H_3)$_2$).

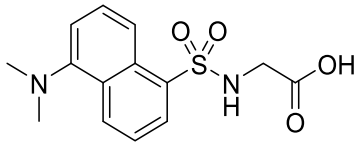
$C_{24}H_{38}N_2O_8S$; calcd. mass 514,63; ESI-MS: m/z 515,7 $[M+H]^+$.



5-(dimethylamino)-N-(2-deoxyglucosyl)naphthalene-1-sulfonamide **42**

To a solution of commercially available D-glucosamine hydrochloride (125 mg, 0,58 mmol) in MeOH (2,3 mL, 0,25 M) were added Et_3N (1,74 mmol, 3 eq) and dansyl chloride (0,7 mmol, 1.2 eq). The reaction was stirred at r.t. and followed by TLC (AcOEt/MeOH 9:1 and AcOEt/MeOH/H_2O/AcOH 6:4:1:1). After 12 h, TLC indicated the disappearance of starting compound, and the solvent was evaporated. The residue was purified by FC (AcOEt/MeOH 9:1) which afford compound **42** (163 mg, 0,4 mmol, 68% yield).

^{1}H NMR (400 MHz, CD_3OD) δ 8.54 (d, *J* = 8.5 Hz, 1H, C*H* Ar), 8.41 (d, *J* = 8.7 Hz, 1H, C*H* Ar), 8.32 (dd, *J* = 7.3, 1.0 Hz, 1H, C*H* Ar), 7.63 – 7.51 (m, 2H, C*H* Ar), 7.26 (d, *J* = 7.5 Hz, 1H, C*H* Ar), 4.59 (d, *J* = 3.4 Hz, 1H, H1), 3.72 – 3.64 (m, 2H, H4, H6a), 3.64 – 3.56 (m, 2H, H5, H6b), 3.23 (t, *J* = 9.3 Hz, 1H, H3), 3.10 (dd, *J* = 10.3, 3.4 Hz, 1H, H2), 2.87 (s, 6H, -N(CH_3)$_2$).
^{13}C NMR (101 MHz, CD_3OD) δ 153.04 (Cq Ar), 138.23 (Cq Ar), 131.17 (Cq Ar), 131.05 (C Ar), 130.96 (Cq Ar), 129.86 (C Ar), 129.01 (C Ar), 124.28 (C Ar), 120.93 (C Ar), 116.38 (C Ar), 92.90 (C1), 72.70, 72.48, 72.18 (C3 C4, C5), 62.57 (C6), 59.84 (C2), 45.80 (-N(*C*H_3)$_2$).
$C_{18}H_{24}N_2O_7S$; calcd. mass 412,46; ESI-MS: m/z 413,5 $[M+H]^+$.



((5-(dimethylamino)naphthalen-1-yl)sulfonyl)glycine **43**
The synthesis of compound 43 is already described in Machida, et al. [14].
^{1}H NMR (400 MHz, d_6-DMSO) δ 8.44 (d, *J* = 8.5 Hz, 1H, C*H* Ar), 8.28 (d, *J* = 8.6 Hz, 1H, C*H* Ar), 8.10 (d, *J* = 7.2 Hz, 1H, C*H* Ar), 7.63 – 7.54 (m, 2H, C*H* Ar), 7.24 (d, *J* = 7.5 Hz, 1H, C*H* Ar), 3.59 (s, 2H, -NHCH_2COOH), 2.82 (s, 6H, -N(CH_3)$_2$).
^{13}C NMR (101 MHz, d_6-DMSO) δ 170.45 (*C*OOH), 151.30 (Cq Ar), 136.39 (Cq Ar), 129.37 (C Ar), 129.10 (Cq Ar), 129.07 (Cq Ar), 127.87 (C Ar), 127.79 (C Ar), 123.54 (C Ar), 119.40 (C Ar), 115.11 (C Ar), 45.12 (-N(*C*H_3)$_2$), 43.85 (-NH*C*H_2COOH).

References

[1] S. K. Banerjee, K. R. McGaffin, N. M. Pastor-Soler, F. Ahmad, *Cardiovascular Research* **2009**, *84*, 111-118.

[2] V. Gorboulev, A. Schuermann, V. Vallon, H. Kipp, A. Jaschke, D. Klessen, A. Friedrich, S. Scherneck, T. Rieg, R. Cunard, M. Veyhl-Wichmann, A. Srinivasan, D. Balen, D. Breljak, R. Rexhepaj, H. E. Parker, F. M. Gribble, F. Reimann, F. Lang, S. Wiese, I. Sabolic, M. Sendtner, H. Koepsell, *Diabetes* **2012**, *61*, 187-196.

[3] B. La Ferla, V. Spinosa, G. D'Orazio, M. Palazzo, A. Balsari, A. A. Foppoli, C. Rumio, F. Nicotra, *ChemMedChem* **2010**, *5*, 1677-1680.

[4] M. Palazzo, S. Gariboldi, L. Zanobbio, S. Selleri, G. F. Dusio, V. Mauro, A. Rossini, A. Balsari, C. Rumio, *Journal of Immunology* **2008**, *181*, 3126-3136.

[5] L. C. H. Yu, J. R. Turner, A. G. Buret, *Experimental Cell Research* **2006**, *312*, 3276-3286.

[6] L. C. H. Yu, A. N. Flynn, J. R. Turner, A. G. Buret, *Faseb Journal* **2005**, *19*, 1822-1835.

[7] E. M. Wright, D. D. F. Loo, B. A. Hirayama, *Physiological Reviews* **2011**, *91*, 733-794.

[8] E. Brenna, C. Fuganti, P. Grasselli, S. Serra, S. Zambotti, *Chemistry-a European Journal* **2002**, *8*, 1872-1878.

[9] G. J. McGarvey, C. A. LeClair, B. A. Schmidtmann, *Organic Letters* **2008**, *10*, 4727-4730.

[10] D. Dubé, A. A. Scholte, *Tetrahedron Letters* **1999**, *40*, 2295-2298.

[11] P. Stepanek, O. Vich, L. Kniezo, H. Dvorakova, P. Vojtisek, *Tetrahedron-Asymmetry* **2004**, *15*, 1033-1041.

[12] aB. La Ferla, F. Cardona, I. Perdiga, F. Nicotra, *Synlett* **2005**, 2641-2642; bS. Mari, F. J. Canada, J. Jimenez-Barbero, A. Bernardi, G. Marcou, I. Motto, I. Velter, F. Nicotra, B. La Ferla, *European Journal of Organic Chemistry* **2006**, 2925-2933.

[13] C. Rumio, M. Palazzo, A. Balsari, F. Nicotra, B. La Ferla, **2008**.

[14] S. Machida, K. Usuba, M. A. Blaskovich, A. Yano, K. Harada, S. M. Sebti, N. Kato, J. Ohkanda, *Chemistry – A European Journal* **2008**, *14*, 1392-1401.

[15] K. C. Nicolaou, C. K. Hwang, M. E. Duggan, *Journal of the American Chemical Society* **1989**, *111*, 6682-6690.

[16] A. A. H. Abdel-Rahman, E. S. H. El Ashry, R. R. Schmidt, *Carbohydr. Res.* **1999**, *315*, 106-116.

Chapter 3. Synthesis of a labeled SGLT1 ligand for *in vitro* and *in vivo* trafficking studies

Abstract

Recently, a novel immunological function was established for the sodium-glucose co-transporter 1 (SGLT1), a protein involved in sugars absorption in small intestine. High-glucose dosage and pharmacological concentrations of C-glucoside **1** showed a protective role of both *in vivo* and *in vivo* systems treated with several inflammatory insults; experimental evidences suggest the involvement of SGLT1 in this process, being this protein "over-activated" by high glucose and/or *C*-glucoside **1**; the mechanism of action at the basis of the protection is still unclear, and in order to increase our knowledge about the phenomenon we have developed a synthesis for the preparation of a radiolabelled-tritiated analogue of *C*-glucoside **1**, as radiotracer for cellular – tissue localization.

Introduction

SGLT1 (sodium glucose co-transporter 1) represents the principal transport protein for glucose and galactose absorption in mammalian intestine[1]. It acts translocating two sodium ions and one molecule of glucose or galactose across the apical membrane of intestinal epithelial cells, which constitute the brush border membrane of small intestine. Beyond its crucial physiological function, recently an immunological role for this transporter has been established[2]. Several works and experimental evidences suggest its involvement and activation in glucose-mediated protection of *in vivo* and *in vitro* models of inflammation states, caused by several insults, like Lipopolysaccharides (LPS)[2], acetaminophen, D-glucosamine and alpha-amanitin[3]. High glucose concentrations (5 g/L *in vitro* and 2,5 g/kg *in vivo*) completely

block the inflammatory response, and this fact seems to be correlated and related to a hyperactivation of SGLT1. The obvious drawback of this treatment, that is, its impact in metabolism, has prompted towards the development of glycomimetics able to act as glucose but at pharmacological concentration. Among a library of *C*-glycosides, we found that compound **1** showed the best protective and anti-inflammatory activity both in vivo and in vitro at dramatic lower concentrations (5 μg/L *in vitro* and 25 μg/Kg *in vivo*), compared to D-glucose[4]. *C*-glycosides are glycomimetics which do not undergo to metabolic changes, thus avoiding the side effects associated with a high glucose concentration therapy.

In these years our work has been focused on the study of the protective role of compound **1** against several inflammatory states, also in vivo, as Chron disease, ulcerative colitis and mucositis, trying to understand and clarify the mechanism of action and the molecular features at the basis of its function. In particular, our attention has been focused on the role of SGLT1, since we have hypothesized a completely new biological role for this protein. To date, we do not possess any information about putative binding site and interaction data about SGLT1 and the bioactive *C*-glycoside.

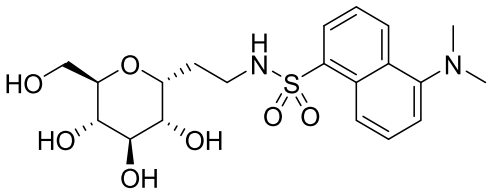


Figure 45 - Structure of *C*-glycoside 1.

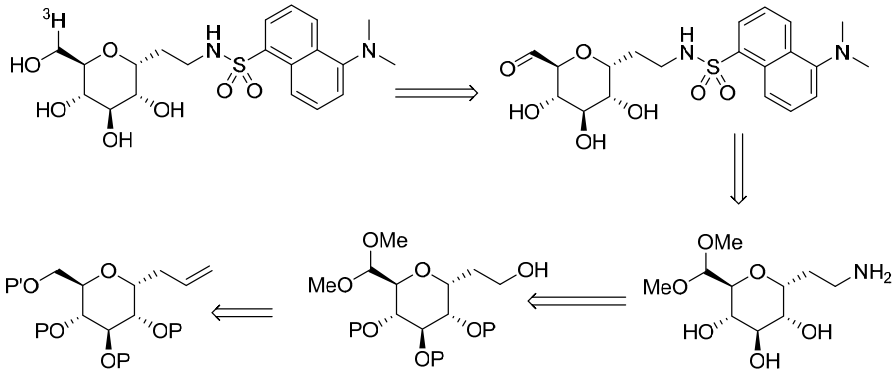
In our opinion, one of the most important aspect is to understand the nature of the interaction between SGLT1, the molecular target, and *C*-glycoside **1**. In order to do this, we decided to synthesize a radiolabeled

C-glycoside **1** derivative, to exploit in radioactivity experiments. The use of radiolabeled compounds is a reliable approach for trafficking studies, widely exploited in medicinal and biological chemistry[5] [6]. This allows to understand the biological destiny of the molecule, in particular if it is transported inside the cell and thus absorbed.

Results and discussion

C-glycoside **1** (Figure 45) is a glucose analogue bearing a dansyl residue connected to a *C*-ethyleneamine spacer fixed at the anomeric position. We developed different synthetic approaches for this compound, both with and without protecting groups. In order to prepare a radiolabeled derivative, we decide to modify the molecule introducing a tritium ($^3[H_2]$) atom in position *C*6 of the sugar region, since it is certainly the most accessible and workable point of the molecule. The synthetic procedure for this purpose was designed in order to introduce the labeling at the end of the synthesis so to limit the manipulation of the radiolabelled entity. The aim was to generate a precursor bearing an aldehyde at C6, that could be exploited for the introduction of the tritium through a reduction with tritiated form of sodium borohydride ($NaB^3[H_2]_4$), a cheap and commercially available radiolabeled reagent. The synthetic scheme and reaction conditions were set up using the corresponding deuterated reagents. For this synthesis a protecting group strategy (Scheme 9) was exploited. As shown in the retrosynthetic scheme (Scheme 8), the desired deprotected aldehyde precursore could be obtained from the corresponding acetal, bearing a *C*-ethyleneamine, which can be chemoselectively derivatized with the dansyl moiety. The latest intermediate could be easily obtained from the corresponding protected

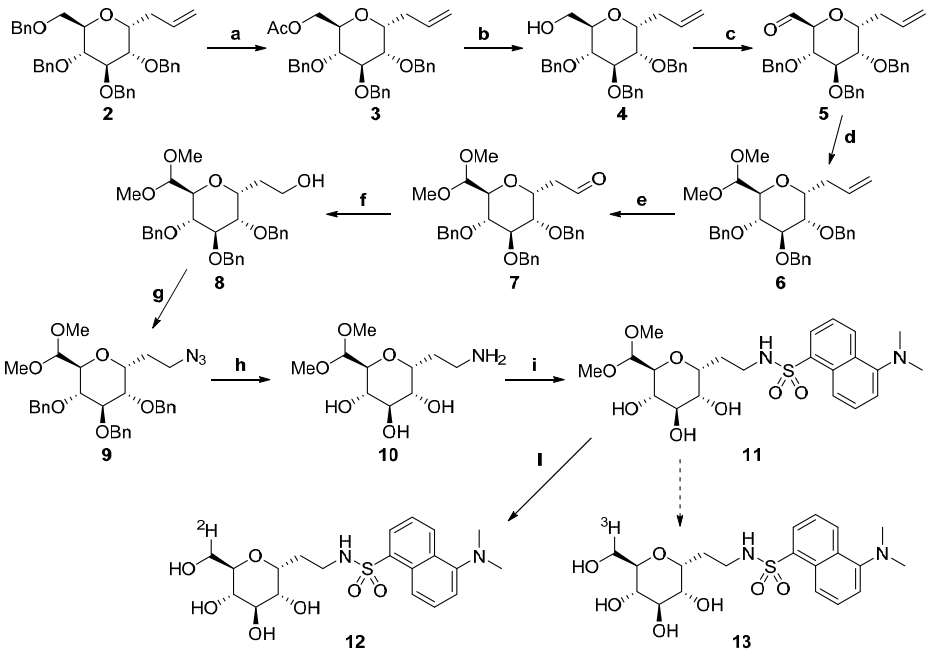
alcohol. Finally the C6 acetal could be synthesized from the orthogonally protected *C*-allyl glucose.



Scheme 8 – Retrosynthetic strategy.

Starting from 2,3,4,6-tetra-*O*-benzyl-α-*C*-allyl-D-glucopyranoside **2**[7] an acetolysis reaction was performed, affording the 6-*O*-acetyl derivative **3**. The following Zemplèn deacetylation[8] allowed to obtain the deprotected primary alcohol **4**, which was rapidly oxidized to the aldehyde **5** with Dess-Martin periodinane. The product was not purified and used directly for the next reaction, where it was converted into the dimethylacetal derivative **6**, according to a procedure of Deleuze, et al. [9]. Once obtained the protected aldehyde, the synthetic attention was moved to the conversion of the allylic group, first oxidizing it into the aldehydic derivative **7**, then reducing it to alcohol **8**. This group was successively converted into the azido derivative **9** by a Mitzunobu reaction; a catalytic hydrogenation/hydrogenolysis on this compound allowed the removal of the benzyl ethers and the azide reduction, affording deprotected amine **10**. This was reacted with dansyl chloride to generate the dansyl derivative **11**, whose dimethylacetal function was subjected to acidic cleavage, to restore the aldehyde group, directly

reacted with sodium borodeuteride to give the $^{2}[H]_2$-derivative **12** of *C*-glycoside **1**. This passage was carried out in order to set up the chemical condition for the Tritium atom insertion. Consequently, intermediate **11** can be used at the same manner to generate, with $NaBT_4$, the tritiated analogue **13**.



Scheme 9 - Reagents and conditions: a) Ac_2O/TFA 4:1, 0°C, 1,5 h, 88%; b) MeONa, MeOH, r.t., 1 h, quant.; c) Dess-Martin periodinane, CH_2Cl_2, r.t., 1,5 h; d) CSA, MeOH, 50°C, 2h, 62% (two steps); e) OsO_4, $NaIO_4$, H_2O/THF/Acetone, r.t., 24 h; f) $NaBH_4$, CH_2Cl_2/EtOH, r.t., 2,5 h, 65% (two steps); g) Ph_3P, DIAD, $(PhO)_2PON_3$, THF dry, 0°C to r.t., 2 h, 60%; h) H2 atm., $Pd(OH)_2$/C, MeOH/AcOEt, r.t., 48 h; i) dansylCl, Et_3N, H_2O/THF, r.t., 2 h, 22% (two steps); l) HCl, H_2O, 2h then $NaBD_4$.

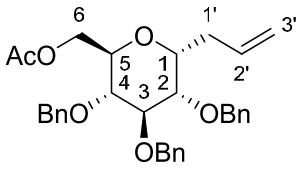
Conclusions

The synthesis of the deuterated derivative of compound **1** was successfully carried out. In order to generate the corresponding radiolabelled compound, for *in vivo* and *in vitro* trafficking studies, the last reaction will be performed employing the corresponding tritiated reagent. These studies will allow to understand if the protective *C*-glucoside, which seems to be a SGLT1 agonist, simply binds to the transporter or if it is also translocated inside cells. This represents a fundamental information to add to our still incomplete knowledge map about the novel established biological role of SGLT1.

Experimental Section

General remarks.

All commercial chemicals were purchased from Sigma-Aldrich. All chemicals were used without further purification. All required anhydrous solvents were dried with molecular sieves for at least 24h prior to use. Thin layer chromatography (TLC) was performed on silica gel 60 F_{254} plates (Merck) with detection under UV light when possible, or by charring with a solution of $(NH_4)_6Mo_7O_{24}$ (21g), $Ce(SO_4)_2$ (1g), concentrated H_2SO_4 (31 mL) in water (500 mL) or with an ethanol solution of ninhydrin or with Dragendorff' spray reagent. Flash-column chromatography was performed on silica gel 230–400 mesh (Merck). ^{1}H and ^{13}C NMR spectra were recorded at 25°C, unless otherwise stated, with a Varian Mercury 400-MHz instrument. Chemical shift assignments, reported in parts per million, were referenced to the corresponding solvent peaks. Mass spectra were recorded on a QTRAP system with ESI source.



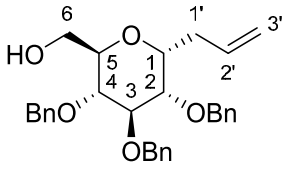
2,3,4-tri-*O*-benzyl-6-*O*-acetyl-α-*C*-allyl-glucopyranoside **3**

A mixture of Ac_2O/TFA 4:1 (70 mL), prepared at 0°C under argon atmosphere was added via a double tip needle to a round bottom flask containing 2 g (3,55 mmol) of 2,3,4,6-tetra-*O*-benzyl-α-D-glucopyranose. The solution was stirred vigorously at 0°C and the reaction followed by TLC (PE/AcOEt 8,5:1,5). After 1,5 h, no more starting compound is present and the solution is poured into ice-water and stirred for 10 min. The aqueous solution was extracted with AcOEt (3x) and the organic phase was then washed once with a sodium hydrogen carbonate saturated solution and twice with distilled water. After anhydrification, filtration and concentration of the remaining organic layer, the crude was purified by FC (PE/AcOEt 9:1), affording compound **3** (1,61 g, 3,12 mmol, 88% yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.43 – 7.24 (m, 15H, C*H* Ar), 5.79 (ddt, *J* = 17.2, 10.2, 6.9 Hz, 1H, H2'), 5.18 – 5.05 (m, 2H, H3'a,b), 4.98 (d, *J* = 10.8 Hz, 1H, CH_2Ph), 4.89 (d, *J* = 10.7 Hz, 1H, CH_2Ph), 4.83 (d, *J* = 10.8 Hz, 1H, CH_2Ph), 4.72 (d, *J* = 11.6 Hz, 1H, CH_2Ph), 4.65 (d, *J* = 11.6 Hz, 1H, CH_2Ph), 4.57 (d, *J* = 10.8 Hz, 1H, CH_2Ph), 4.24 (d, *J* = 3.5 Hz, 2H, H6a,b), 4.11 (dt, *J* = 9.3, 5.9 Hz, 1H, H1, H1), 3.85 (t, *J* = 9.0 Hz, 1H, H3), 3.76 (dd, *J* = 9.4, 5.8 Hz, 1H, H2), 3.70 (dt, *J* = 9.8, 3.5 Hz, 1H, H5), 3.48 (dd, *J* = 9.8, 8.7 Hz, 1H, H4), 2.50 (t, *J* = 8.2 Hz, 2H, H1'a,b), 2.04 (s, 3H, CH_3CO).

^{13}C NMR (101 MHz, $CDCl_3$) δ 170.92 (CH_3*C*O), 138.59 (Cq Ar), 138.15 (Cq Ar), 137.83 (Cq Ar), 134.40 (C2'), 128.80, 128.65, 128.59, 128.58, 128.49, 128.29, 128.11, 128.05, 128.00, 127.95, 127.85 (C Ar x 15), 117.24 (C3'),

82.37 (C3), 80.04 (C2), 77.89 (C4), 75.65, 75.21 (*C*H_2Ph), 73.68 (C1), 73.23 (*C*H_2Ph), 69.68 (C5), 63.62 (C6), 29.89 (C1'), 21.00 (*C*H_3CO).
$C_{32}H_{36}O_6$; calcd. mass: 516,63; MS-ESI: m/z 517,64 [M+H^+].



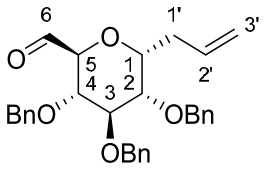
2,3,4-tri-*O*-benzyl-α-*C*-allyl-glucopyranoside **4**

Compound **3** (1335 mg, 2,58 mmol) was subjected to deacetylation using the standard procedure described in literature. Briefly, the acetyl ester derivative was dissolved in 10 mL of a CH_2Cl_2/MeOH mixture and 2,6 mL of a sodium methoxide solution (1 M) were added. The reaction was stirred at r.t. and followed by TLC (PE/AcOEt 8:2). After 1 h the solution was neutralized with the addition of IRA-120H^+ resin, which was then filtered and the organic solution was concentrated, to afford alcohol **4** in a pure form (quant. yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ = 7.38 – 7.27 (m, 15H, C*H* Ar), 5.77 (td, *J*=17.0, 6.8, 1H, H2'), 5.10 (dd, *J*=13.7, 7.6, 2H, H3'a,b), 4.94 (d, *J*=10.9, 1H, CH_2Ph), 4.85 (dd, *J*=19.3, 10.9, 2H, CH_2Ph), 4.71 (d, *J*=11.6, 1H, CH_2Ph), 4.63 (d, *J*=11.7, 2H, CH_2Ph), 4.10 – 4.01 (m, 1H, H1), 3.86 – 3.68 (m, 3H, H2, H3, H6a), 3.68 – 3.59 (m, 1H, H5), 3.58 – 3.46 (m, 2H, H4, H6b), 2.52 – 2.44 (m, 2H, H1'a,b).

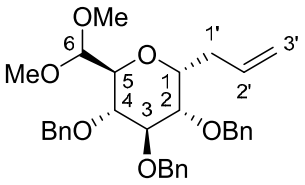
13C NMR (101 MHz, $CDCl_3$) δ = 138.72, 138.21, 138.11 (Cq Ar), 134.58 (C2'), 128.62, 128.57, 128.52, 128.20, 128.03, 127.99, 127.96, 127.91, 127.76 (C Ar), 117.34 (C3'), 82.31 (*C*H sugar), 80.19 (*C*H sugar), 78.13 (*C*H sugar), 75.56 (*C*H_2Ph), 75.26 (*C*H_2Ph), 73.71 (*C*H sugar), 73.27 (*C*H_2Ph), 71.65, 62.36 (C6), 30.07 (C1').

$C_{30}H_{34}O_5$; calcd. mass: 474,6; MS-ESI:m/z 475,5 [M+H^+], 497,6 [M+Na]$^+$.



2,3,4-tri-*O*-benzyl-6-oxo-α-*C*-allyl-glucopyranoside **5**

Alcohol **4** (830 mg, 1,75 mmol) was dissolved in CH_2Cl_2 (15 mL) and 2,6 mmol (1,5 eq) of Dess-Martin periodinane were added. The reaction was stirred at r.t. and followed by TLC (PE/AcOEt 8:2). After 1,5 h, TLC indicated the absence of the starting material and the formation of a new spot with a higher Rf, just above those of the starting alcohol. The reaction is then quenched by the addition of 15 mL of satd. solution of $NaHCO_3$ and 15 mL of 10% aqueous sodium thiosulphate, and the reaction was vigorously stirred for 10 min. The mixture was partitioned between aqueous and organic phases, and the first was extracted (3x) with CH_2Cl_2. The combined organic phases were dried over sodium sulphate and concentrated. The crude residue was analyzed by ^{1}H-NMR (δ 9.73 ppm, $CDCl_3$) to check the presence of the aldehydic signal and immediately used for the next reaction.



2,3,4-tri-*O*-benzyl-6,6-dimethoxy-α-*C*-allyl-glucopyranoside **6**

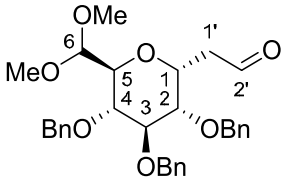
The crude aldehyde **5** (912 mg) was dissolved in dry MeOH (20 mL) and 405 mg of camphor sulphonic acid (1,75 mmol, 1 eq relative to alcohol **4**) were added. The reaction was heated and stirred at 50°C until TLC (PE/AcOEt 8:2) showed no more starting aldehyde. After 2 h, the reaction was cooled to r.t. and 10 mL of $NaHCO_3$ satd. solution and 20

mL of water were added. The solution was then extracted with AcOEt (3x). The organic layer was separated, dried and the product purified by FC (PE/AcOEt 9:1 to 8,5:1,5). 564 mg of dimethylacetal derivative **6** (1,09 mmol, 62% yield over two steps) were obtained.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.40 – 7.22 (m, 15H, C*H* Ar), 5.90 – 5.77 (m, 1H, H8), 5.20 – 5.05 (m, 2H, H9a,b), 4.88 (d, *J* = 11.1 Hz, 1H, CH_2Ph), 4.82 (d, *J* = 11.0 Hz, 1H, CH_2Ph), 4.77 (d, *J* = 11.1 Hz, 1H, CH_2Ph), 4.68 (d, *J* = 11.6 Hz, 1H, CH_2Ph), 4.64 (d, *J* = 11.2 Hz, 1H, CH_2Ph), 4.60 (d, *J* = 11.7 Hz, 1H, CH_2Ph), 4.53 (d, *J* = 2.3 Hz, 1H, H6), 4.14 (dt, *J* = 10.1, 4.9 Hz, 1H, H1), 3.81 (t, *J* = 8.4 Hz, 1H, H3), 3.75 – 3.66 (m, 2H, H2, H5), 3.62 (t, *J* = 8.5 Hz, 1H, H4), 3.40 (s, *J* = 10.9 Hz, 3H, CH_3O-), 3.38 (s, 3H, CH_3O-), 2.60 – 2.40 (m, 2H, H7a,b).

^{13}C NMR (101 MHz, $CDCl_3$) δ 138.73 (Cq Ar), 138.35 (Cq Ar), 138.28 (Cq Ar), 134.87 (C8), 128.58, 128.52, 128.51, 128.23, 128.00, 127.95, 127.90, 127.74 (C Ar x 15), 117.14 (C9), 102.28 (C6), 81.39 (C3), 79.29 (C2), 77.97 (C4), 75.20, 74.79 (*C*H_2Ph), 73.58 (C1), 73.08 (*C*H_2Ph), 71.85 (C5), 55.32 (*C*H_3O-), 55.05 (*C*H_3O-), 30.51 (C7).

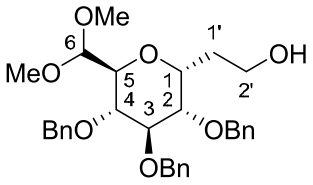
$C_{32}H_{38}O_6$; calcd. mass: 518,65; MS-ESI: m/z 519,70 [M+H^+].



2-(2,3,4-tri-*O*-benzyl-6,6-dimethoxy-α-glucopyranosyl)-ethanal **7**

Compound **6** (560 mg, 1,08 mmol) were dissolved in a H_2O/THF/Acetone solution (4,5:4,5:3 mL). $NaIO_4$ (5,4 mmol, 5 eq) were added and the suspension was stirred at r.t. for 30 min. Then 0,054 eq of OsO_4 (solution in tBuOH) were dropped in the reaction that was vigorously stirred for at r.t., following the formation of the product by TLC (PE/AcOEt 6:4). After

24 h, the reaction is concentrated and the aqueous residue was extracted with AcOEt (3x) and the organic phase back-extracted twice with water and brine. The organic solution was dried over Na_2SO_4 and concentrated *in vacuo*, affording crude aldehyde **7** (601 mg) which was used directly for the next reaction without purification. ^{1}H-NMR of a sample of the crude revealed the absence of the allylic bond and the formation of the aldehyde group (δ 9,71 ppm).



2-(2,3,4-tri-*O*-benzyl-6,6-dimethoxy-α-glucopyranosyl)-ethanol **8**

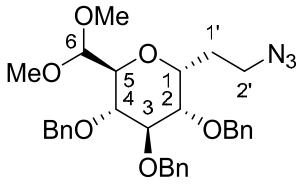
To a solution of crude compound **7** in CH_2Cl_2/EtOH (10 + 5 mL), 4,32 mmol (4 eq, relative to compound **6**) of $NaBH_4$ were added. The reaction was stirred at r.t. and followed by TLC (PE/AcOEt 7:3). After 2,5 h, the solvent was evaporated, the residue resuspended in a solution of saturated Na_2CO_3 and stirred for 20 min. The product was extracted from the aqueous phase with AcOEt (3x), the organic phase was dried, concentrated and purified by FC (eluent PE/AcOEt 5:5). 365 mg (0,698 mmol) of compound **8** were obtained (65% yield over two steps).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.41 – 7.24 (m, 15H, C*H* Ar), 4.83 – 4.75 (m, 2H, CH_2Ph), 4.72 – 4.66 (m, 2H, CH_2Ph), 4.66 – 4.59 (m, 2H, CH_2Ph, C6), 4.59 – 4.52 (m, 1H, CH_2Ph), 4.23 – 4.12 (m, 1H, H1), 3.90 – 3.72 (m, 4H, H3, H4, H5, H8a), 3.63 – 3.53 (m, 2H, H2, H8b), 3.42 (s, 3H, CH_3O-), 3.38 (s, 3H, CH_3O-), 2.16 – 2.03 (m, 1H, H7a), 1.79 – 1.65 (m, 1H, H7b).

^{13}C NMR (101 MHz, $CDCl_3$) δ 138.41, 138.31, 138.08 (Cq Ar), 128.56, 128.55, 128.54, 128.18, 128.12, 128.02, 127.99, 127.95, 127.92, 127.89, 127.85 (C Ar), 102.14 (C6), 79.97 (C3), 78.42 (C2), 76.80 (C4), 74.71

(*C*H_2Ph), 74.19 (*C*H_2Ph), 73.48 (C1), 73.16 (*C*H_2Ph), 72.25 (C5), 61.48 (C8), 56.67 (*C*H_3O-), 54.54 (*C*H_3O-), 28.72 (C7).

$C_{31}H_{38}O_7$; calcd. mass: 522,64; MS-ESI: m/z 523,66 [M+H^+].



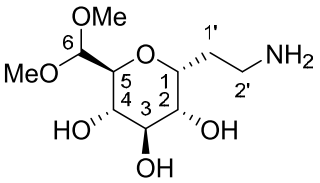
2-(2'-azidoethyl)-2,3,4-tri-*O*-benzyl-6,6-deoxy-dimethoxy-α-D-glucopyranoside **9**

Ph_3P (1,957 mmol, 3 eq) was added to a solution of alcohol **8** (341 mg, 0,652 mmol) in dry THF (3,23 mL). The solution was cooled to 0°C and DIAD (1,957 mmol, 3 eq) was added dropwise. After the formation of a white precipitate, $(PhO)_2PON_3$ (2,088 mmol, 3,2 eq) was added and the reaction was stirred at r.t., following the disappearance of the starting material by TLC (PE/AcOEt 7:3). After 2 h, the solvent was evaporated and the crude was loaded on silica gel, performing a FC (eluent PE/AcOEt 9:1 to 6:4) which afforded the desired product **9** (209 mg, 0,38 mmol) with 60% yield.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.37 – 7.24 (m, 15H, C*H* Ar), 4.82 – 4.75 (m, 2H, CH_2Ph), 4.74 – 4.61 (m, 3H, CH_2Ph), 4.60 – 4.52 (m, 2H, CH_2Ph, C6), 4.13 – 4.04 (m, 1H, H1), 3.75 (t, *J* = 7.7 Hz, 1H, H3), 3.70 – 3.56 (m, 3H, H2, H4, H5), 3.50 – 3.30 (m, 8H, CH_3O- x 6, H8a,b), 2.09 – 1.94 (m, 1H, H7a), 1.91 – 1.79 (m, 1H, H7b).

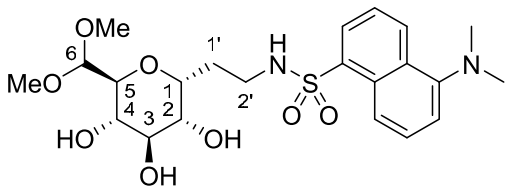
^{13}C NMR (101 MHz, $CDCl_3$) δ 138.49, 138.33, 138.07 (Cq Ar), 128.59, 128.57, 128.56, 128.15, 128.03, 127.94, 127.91, 127.84 (C Ar), 102.11 (C6), 80.27 (C3), 78.43 (C2), 77.01 (C4), 74.86 (*C*H_2Ph), 74.40 (*C*H_2Ph), 73.12 (*C*H_2Ph), 72.58 (C1), 70.57 (C5), 55.74 (*C*H_3O-), 54.97 (*C*H_3O-), 48.06 (C8), 29.84 (C7).

$C_{31}H_{37}N_3O_6$; calcd. mass: 547,65; MS-ESI: m/z 548,66 [M+H$^+$].



α-*C*-(1'-ethylamino)-6,6-dimethoxy-D-glucopyranoside **10**

Compound **9** (209 mg, 0,38 mmol) was dissolved in a mixture of MeOH/AcOEt (5 + 5 mL) and the solution was degased under vacuum. A catalytic amount of $Pd(OH)_2/C$ was added and the reaction was stirred vigorously at r.t. under hydrogen atmosphere. After 48 h, the catalyst was removed by filtration through a pad of celite and the filtrate was concentrated. ^{1}H-NMR analysis was performed on the crude to confirm the complete debenzylation of the molecule. The crude amine was used directly for the next reaction.



α-*C*-[(ethylen-2'-(*N,N*-Dimethylamino)-*N*-naphthalensulfonamidyl]-6,6-dimethoxy-D-glucopyranoside **11**

Crude amine **10** (69 mg) was dissolved in a mixture of water/THF. Et_3N (0,76 mmol, 2 eq related to compound **9**) and 0,57 mmol (1,5 eq) of dansyl chloride were added to the reaction that was stirred at r.t. and followed by TLC (CH_2Cl_2/MeOH/NH_3 (aq) 5:5:1 and AcOEt/MeOH 9:1). After 2 h, the solvent was removed and the product purified by FC (eluent AcOEt/MeOH 9,5:0,5) obtaining 39,7 mg (0,082 mmol, 22% yield over two steps) of a mixture of compound 11 and the hydrated form of

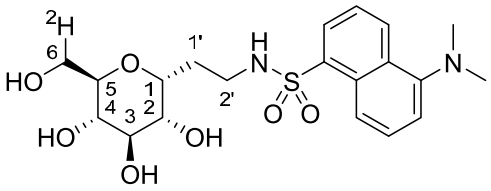
the aldehyde, in a ratio, as determined by ^{1}H-NMR, of ≈ 70:30 (dimethylacetal:hydrated aldehyde).

1H NMR (400 MHz, CD_3OD) δ 8.56 (t, J = 7.6 Hz, 1H), 8.34 (d, J = 8.7 Hz, 1H), 8.20 (dd, J = 7.3, 1.0 Hz, 1H), 7.63 – 7.51 (m, 2H), 7.25 (d, J = 7.5 Hz, 1H), 4.64 (d, J = 3.7 Hz, 1H), 3.82 – 3.73 (m, 1H), 3.55 – 3.35 (m, 10H), 2.94 (t, J = 6.9 Hz, 2H), 2.87 (s, 6H), 1.79 – 1.60 (m, 2H).

Signals of the aldehyde in hydrated form: δ 8.15 (dd, J = 7.3, 0.9 Hz, 1H, CH Ar), 4.62 (d, J = 3.7 Hz, 1H, H6), 3.91 – 3.83 (m, 1H, H6), 3.15 – 3.05 (m, 2H, H2'a,b), 1.98 – 1.81 (m, 2H, H1'a,b).

^{13}C NMR (101 MHz, CD_3OD, mixture of signals of 11 and its hydrated aldehyde as contaminant) δ 153.17, 153.11, 145.10, 145.09, 136.76, 135.75, 135.58 (Cq Ar), 131.41, 131.17 (C Ar), 130.91 (Cq Ar), 130.74, 130.33, 129.12, 129.06, 124.31, 120.86, 120.43, 116.41 (C Ar), 104.69, 104.16 (C6), 74.61, 74.46, 74.44, 74.16, 73.93, 72.24, 72.16, 71.72, 71.46 (*C*H sugar), 56.40, 56.32, 55.65, 55.32 (*C*H_3OCH-), 45.79 ((*C*H_3)$_2$N-), 41.59 (C2'), 26.98, 25.51 (C1').

$C_{22}H_{32}N_2O_8S$; calcd. mass: 484,56; ESI-MS: m/z 485,3 $[M+H]^+$, 507,4 $[M+Na]^+$, 479,3 $[M+H]^+$ (aldehyde in hydrated form).



[$^{2}H_2$]-α-*C*-[(ethylen-2'-(*N,N*-Dimethylamino)-*N*-naphthalensulfonamidyl]-D-glucopyranoside **12**

Compound **11** (162) (1 mg) was dissolved in 0,5 mL of distilled water; 0,05 mL of HCl 2 M were added to the solution, stirred at r.t. for 24 h. The reaction was followed by TLC (CH_2Cl_2/MeOH), which indicated the formation of a new product with a minor Rf respect to starting

compound. ESI-MS spectrum of the crude revealed a m/z signal of 439,3 $[M+H]^+$, confirming the conversion of the dimethylacetal into aldehyde group. The pH of the solution was then adjusted to 6 adding drops of NaOH 4 M. $NaBD_4$ (0,009 mmol, 4 eq) was added to the solution, stirred at r.t. for 1 h. TLC (CH_2Cl_2/MeOH 8:2) indicated the complete transformation of the aldehydic intermediate into a more polar compound, with a Rf identical to that of the H*6*-analogue of the compound (compound **1**). A Mass Spectrometry analysis of the solution revealed the presence of a molecule with a m/z of 442,4 (M + 1).
$C_{20}H_{27}DN_2O_7S$; calcd. mass: 441,52 ESI-MS: m/z 442,4 $[M+H]^+$, 464,4 $[M+Na]^+$.

References

[1] E. M. Wright, B. A. Hirayama, D. F. Loo, *Journal of Internal Medicine* **2007**, *261*, 32-43.

[2] M. Palazzo, S. Gariboldi, L. Zanobbio, S. Selleri, G. F. Dusio, V. Mauro, A. Rossini, A. Balsari, C. Rumio, *Journal of Immunology* **2008**, *181*, 3126-3136.

[3] L. Zanobbio, M. Palazzo, S. Gariboldi, G. F. Dusio, D. Cardani, V. Mauro, F. Marcucci, A. Balsari, C. Rumio, *Am. J. Pathol.* **2009**, *175*, 1066-1076.

[4] B. La Ferla, V. Spinosa, G. D'Orazio, M. Palazzo, A. Balsari, A. A. Foppoli, C. Rumio, F. Nicotra, *ChemMedChem* **2010**, *5*, 1677-1680.

[5] C. Zona, B. La Ferla, *Journal of Labelled Compounds & Radiopharmaceuticals* **2011**, *54*, 629-632.

[6] Y. Kuang, N. Salem, D. J. Corn, B. Erokwu, H. Tian, F. Wang, Z. Lee, *Molecular Pharmaceutics* **2010**, *7*, 2077-2092.

[7] G. J. McGarvey, C. A. LeClair, B. A. Schmidtmann, *Organic Letters* **2008**, *10*, 4727-4730.

[8] Z. Wang, in *Comprehensive Organic Name Reactions and Reagents*, John Wiley & Sons, Inc., **2010**.

[9] A. Deleuze, M. Sollogoub, Y. Bleriot, J. Marrot, P. Sinay, *European Journal of Organic Chemistry* **2003**, 2678-2683.

Chapter 4. Generation of gold nanoparticles decorated with synthetic ligands of co-transporters SGLT-1 and B^0AT1 for the investigation of the multivalent-synergistic effect.

Abstract

D-Glucose, *C*-glucoside **1** and L-glutamine have been derivatized with suitable thiol ending spacers-linkers and have been used for the surface functionalization of gold nanoparticles (AuNPs). In particular *C*-glucoside **1**, D-glucose and L-glutamine decorated nanoparticles have been generated as molecular tools to investigate the existence of a possible multivalency effects towards the co-transporters, respectively the sodium glucose co-transporter 1 (SGLT1) and sodium glutamine co-transporter (B^0AT1). Moreover, gold nanoparticles decorated with a combination of the three ligands have been generated as tool study potential synergistic effects. In vitro biological evaluations are ongoing on Caco 2 cell line expressing both SGLT1 and B^0AT1 transporters.

Introduction

Recent papers outline the protective effect of sodium-glucose cotransporter (SGLT1) engagement with high oral doses of D-glucose and non metabolizable 3-*O*-methyl-D-glucopyranose (3OMG) on damages induced by Toll-Like Receptors (TLRs) ligands in intestinal epithelial cells, in a murine model of septic shock and in LPS-induced liver injury, as well as liver injury and death induced by an overdose of acetaminophen[1]. This was found being due to glucose-induced down-regulation of systemic production of inflammatory cytokines and enhanced production of anti-inflammatory cytokines. However, the high amounts of D-glucose and 3OMG necessary to induce anti-inflammatory effects represent a serious limitation for a hypothetical therapeutic application. In a previous work we identified a glucose derivative, *C*-glucoside **1** (Figure 46), exerting anti-inflammatory activity at a molar concentration

five orders of magnitude lower than glucose[2]. On the other hand, glutamine is the most abundant aminoacid in the total body free aminoacid pool and there has recently been much interest in its role in nutrition. Glutamine is the major energy source for intestinal epithelium and it is involved in maintenance of intestinal epithelial homeostasis. The protective anti-inflammatory effects of glutamine are also well established[3].

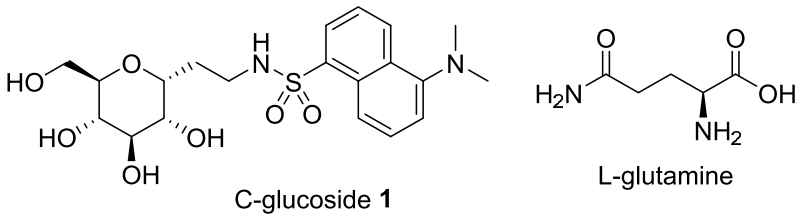


Figure 46 - Structure of the two ligands objected of this work, *C*-glucoside 1 and L-glutamine.

Glucose and glutamine are major nutrients absorbed by the intestinal epithelium via SGLT1 and sodium-glutamine cotransporter 1 (B^0AT1) respectively, both expressed on enterocytes. One of the most important events that compromise the intestinal epithelium biological functions (like nutrients absorptions and barrier effect against pathogens invasion) is the lack of the correct permeability caused by alterations of the junctional systems of epithelial cells in response to different damages. The fundamental aspect that links SGLT1 activation and L-glutamine uptake is the ability to modulate tight junctions activating intracellular signalling pathways, involved in junctional open-close mechanisms and remodelling[4]. Thus it could be of great relevance the simultaneous activation of both receptors, which could lead to a synergistic protective effect.

Multivalent interactions are considerably stronger than the individual bonding of a corresponding number of monovalent ligands to a

multivalent receptor and are often used in biological systems[5], particularly for cellular recognition processes and in signal transduction. Receptor dimerization/clustering is a well-known and ubiquitous phenomenon. Several reports have suggested that functional SGLT1 is an oligomeric protein, resulting by homodimerization of two identical subunits[6]. Evidence also suggest that the B^0AT1 transporter is mainly assembled as a cluster[7], and, moreover, that the B^0AT1 transporter localized in the small intestine displays a low affinity constant for the best aminoacid substrate[7]. These findings suggest that the study of multivalent effects of the single receptor ligand *C*-glycoside **1** or glutamine could lead to a significant increase in the observed activity.

With these preliminary data in hand, and with the above mentioned rational, our research topic consists in the study and evaluation of possible multivalent and synergistic effects operated by glucose-like ligands (*C*-glucoside **1** and other derivatives) in combination with glutamine and/or glutamine-like compounds in the preservation and/or recovery of the intestinal epithelium damaged by inflammation (induced by LPS and/or its chemotherapeutics agents). To reach our goal we selected gold nanoparticles as multivalent ligand presenting scaffold, for straightforward decoration with molecules bearing a suitable thiol functionalized spacer-linker. Moreover, the possibility to simultaniously introduce different thiol ligands, makes them suitable for the generation of molecular tools for the study of potential synergistic effects, exploiting the already prepared ligands without additional orthogonal functional group derivatization.

Results and Discussion

In the first phase of the work, our intent was to prepare a set of different surface-functionalized gold nanoparticles in order to study both the multivalent and the synergistic effect of **1** and L-Gln. We designed and prepared a set of thiol-functionalized derivatives of both the ligands (Figure 47) in order to have the suitable compounds for the preparation of the corresponding gold nanoparticles (AuNP) (Figure 48)[8]. We designed the synthesis of the derivatives taking into account the attachment of a suitable spacer / linker bearing a terminal thiol group. This linker ensures a correct spatial separation between the core of the gold nanoparticle and the ligand, which is necessary for the interaction of the bioactive molecules with the receptor of interest. The linker we chose is made of a triethyleneglycol (TEG) chain and a pentenyl (C5) spacer. Polyethylene glycol is bio-friendly polymer, it is extensively used in pharma industries, for its low toxicity, excellent solubility in aqueous media, and seems to be neither immunogenic nor allergenic[9] [10]. Moreover, PEG represents an inert polymer (is a polyether structure) and is not metabolized. PEGylation is an important and emerging aspect of the drug delivery field; the conjugation of PEG to drugs leads to an increase of pharmacokinetics and pharmacodynamics properties of the molecule to which it is conjugated. The hydrophobicity of the C5 chain is necessary to allow a good attachment of the ligand to the core of the Au-NP.

Concerning compound **1** derivative, we decided to connect the whole linker to the hydroxyl group on C-6 position of the pyranose ring, for two main reasons: firstly, this position bears a primary hydroxyl group, more easily functionalizable rather than the other secondary OH groups of the ring; secondly, several works suggest that the C6 OH group of D-glucose is not crucial for the binding to SGLT1 interaction and translocation site.

In our opinion, a presence of a linear chain should not affect the biological activity of the molecule. For L-glutamine ligands, we designed two different derivatives, bearing respectively the TEG-C5 chain on the nitrogen atom of the amide of lateral chain or connecting it directly to the C_α position of the aminoacids, with a new amide function. The choice to synthesized two L-Gln derivatives with the TEG chains on two different positions of the aminoacidic structure was made because we have no information about how L-Gln interacts with B^0AT1 (i.e. which groups are involved in the binding). The synthesis of both *C*-glucoside **1** and L-Gln derivative proceeded with a protecting group strategy, starting from commercially available *O*-Me-α-D-glucopyranoside and protected L-glutamine compounds. The presence of the ending thiol group was ensured and achieved by transforming a terminal olefinic appendage, on the C5 chain, into a thioester group. This was obtained exploiting the so called "Thiol-ene coupling" (TEC) approach[11]. Many synthetic procedures allow the conversion of a terminal alkene into a thioether linkage, including the reaction of the thiol-bearing compound and the alkenyl group in presence of a radical initiator, as AIBN, often heating the reaction to reflux, under argon atmosphere and with degased solvents. We decided to exploit the use of a radical photoinitiatior, called 2,2-dimethoxy-2-acetophenone (DPAP) which allows the thiol-ene reaction in a faster and more simple manner. It permits the use of not-degased solvents and water, the presence of air, and the proceeding the reaction at room temperature; only a short (generally 30 min – 1 hour) UV-light exposure at 365 nm (using a simple UV-light lamp), thus close to visible wavelengths, is required.

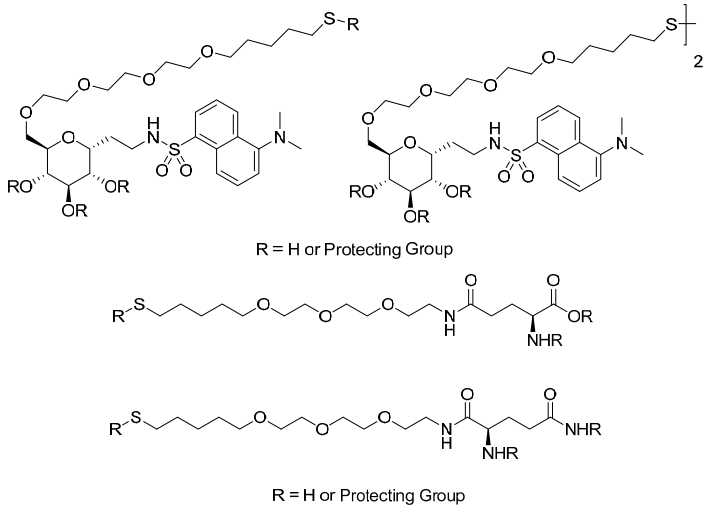


Figure 47 – Thiol-functionalized ligands for the preparation of the Au-NPs.

During the second phase, the objective consisted in the preparation, purification and structural characterization of gold nanoparticles (Figure 48) loaded with:

- Compound **1** derivative;
- L-glutamine derivatives;
- both compound **1** and L-glutamine derivatives.

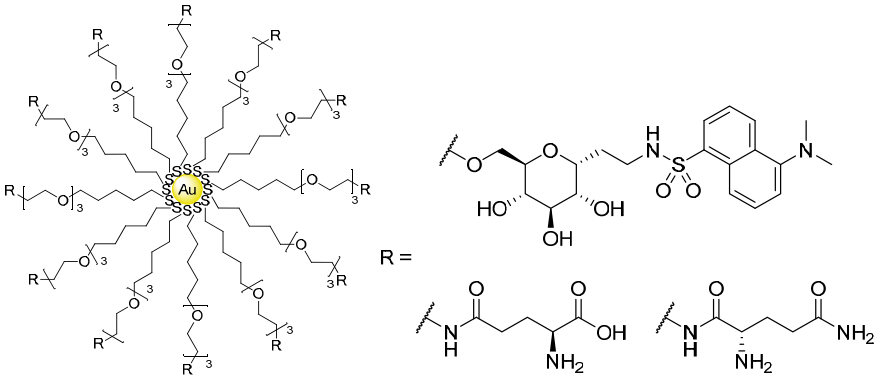
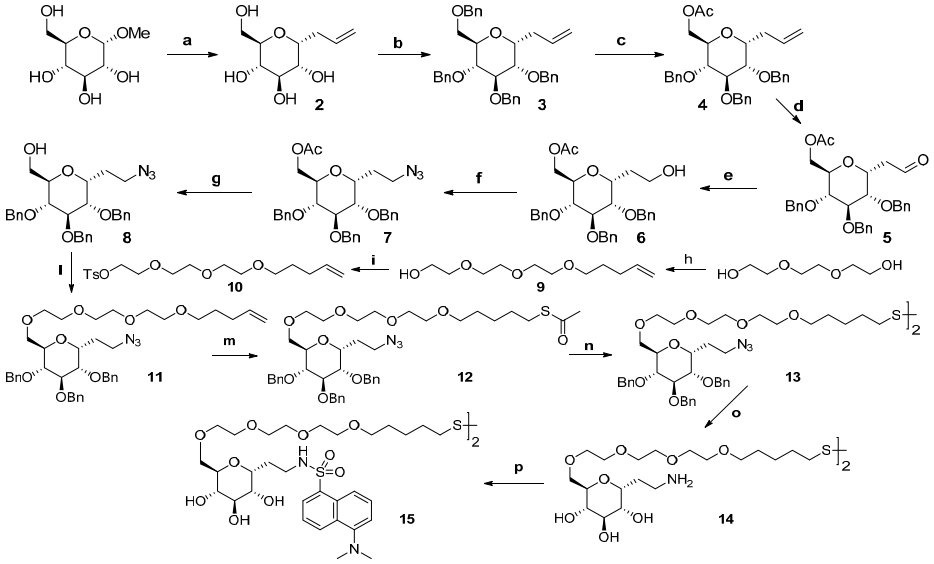


Figure 48 - General structure of the designed nanoparticles.

Preparation of ligands

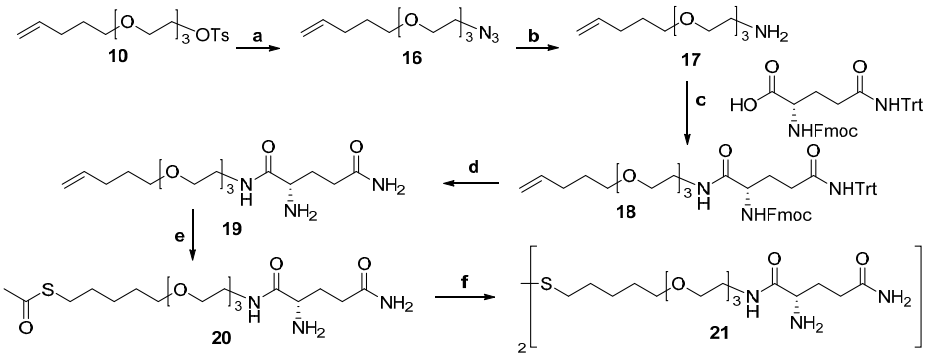
Concerning the synthesis of compound **1** derivative, the following synthetic steps have been carried out (Scheme 10). *O*-Me-α-D-glucopyranoside was subjected to a *C*-allylation according to the procedure described in Cardona and La Ferla [12] (compound **2**). Then the respective perbenzylated compound **3** was obtained (McGarvey, et al. [13]), and an acetolysis reaction in presence of trifluoroacetic acid and acetic anhydride allowed the removal of benzyl protecting group from the hydroxyl group in C-6 position and the contemporary transformation into an acetylester group (compound **4**). Then the allylic appendage was oxidized to aldehyde **5**, reduced to alcohol **6** and converted to azidoethyl derivative **7**. The acetyl group was successively removed from the OH in C6 position (**8**), which was alkylated with the electrophilic PEG-C5 chain bearing a methyl-p-toluen-sulphonate function as leaving group. The spacer was rapidly synthesized starting from TEG, then monoalkylated with a 4-pentenyl chain (**9**). The reaction with tosyl chloride afforded **10**. In our first planned synthesis, we subjected this compound to BCl_3-promoted benzyl deprotection, but we observed also the cleavage of the polyethyleneglycol chain. Hence, we decided first of all to subject the TEGylated sugar derivative to a radical-mediated reaction. Upon UV exposure at 365 nm of wavelength and adding a radical photoinitiator, as DPAP (2,2-dimethoxyphenyl-2-acetophenone), the allyl group was converted into the thioester over 1 h of reaction, to afford compound **12**. Successively, we tried both catalytic hyrogenolysis and Birch conditions for the azide reduction and benzyl cleavage, but, surprisingly, in both cases a complete removal of the thioester group, generating a pentyl ending chain, occurred. By the way, the desulfurization of thioester by SET (single electron transfer) processes is described and rationalized in literature[14]. Therefore, we changed again the strategy,

by deprotecting the thioester in basic conditions and promoting the disulphide bridge formation, in order to obtain a sort of "protected" thiol (compound **13**). Birch reaction on this compound led to the desired fully deprotected dimer **14**, of which the two amine functionalities were transformed into dansylsulphonamides, affording the final dimer **15**. This compound was used directly for the generation of the nanoparticles: it is possible to use both thiol-free and disulphide bearing molecules, since the addition of an excess of hydride reagents (often $NaBH_4$) leads to the disulphide bridge reduction.



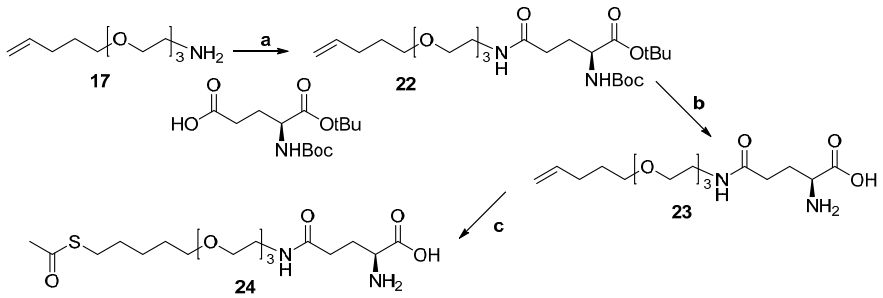
Scheme 10 - Reagents and conditions: a) i. BSTFA, CH_3CN, reflux, 3 h; ii. AllylTMS, TMSOTf, 0°C to r.t., 12 h[12]; b) BnBr, NaH, DMF dry, r.t., 3 h[13]; c) Ac_2O, TFA, 0°C, 1 h, 88%; d) i. O_3, CH_2Cl_2, -78°C, 1 h; ii. Ph_3P, CH_2Cl_2, -78°C to r.t. 24 h, 82%; e) $NaBH_4$, MeOH, 0°C, 2 h, 95%; f) DIAD, Ph_3P, $(PhO)_2PON_3$, THF dry, 0°C to r.t., 2 h, 90%; g) MeONa, MeOH dry, r.t., 1 h, 95%; h) 5-Bromo-1-pentene, NaOH, neat, 100°C, 14 h, 74%; i) TsCl, Et_3N, DIPEA, CH_2Cl_2, r.t., 3 h, 72%; l) NaH, DMF dry, r.t., 12 h, 83%; m) AcSH, DPAP, CH_2Cl_2, r.t., UV light (365 nm), 1 h, 85%; b) MeONa, MeOH, r.t., 16 h; c) Na, NH_3(l), THF dry, -78°C, 15 min, then NH_4Cl, r.t., 65%; d) dansyl chloride, Et_3N, MeOH, r.t., 5 h, 39%.

The synthesis of the L-glutamine derivatives proceeded with the following steps (Scheme 11 and Scheme 12). The tosyl group on the C5-TEG chain was converted into the azide function, which was reduced to amine (**17**). A coupling between this compound and Fmoc-L-Gln-Trt-OH afforded the derivative **18**, which was subjected to protecting group cleavage both in basic and acid conditions (compound **19**). The thiol-ene coupling with thioacetic acid gave thioester **20** rapidly cleaved in basic condition to afford dimer **21** (as resulted from NMR analysis), directly used for the nanoparticle preparation. The coupling of the amine **17** with Boc-L-Glu-O-tBu (Boc-L-glutamic acid 1-tert-butyl ester) gave intermediate **22**, which was deprotected by acidic conditions. The successive TEC with thiolacetic acid afforded compound **24**. In this case the thioester deprotection was not preformed since it requires not a simple methanolysis, but the use of other mild conditions, in order to avoid the Cα-racemization of the glutamine. This process did not occur for the first aminoacidic derivative (21) since after the coupling, the racemizable carbon atom of the glutamine is located on the α position to the new amide formed, and not to a carboxylic acid group.



Scheme 11 - Reagents and conditions: a) NaN_3, DMF dry, 80°C, 48 h, 78%; b) i. Ph_3P, THF, reflux, 3h; ii. H_2O, r.t., 12 h, 93%; c) EDC, HOBt, Et_3N, DMF dry, 0°C to r.t., 12 h,

79%; d) i. CH_2Cl_2, TFA, r.t., 2 h; ii. piperidine, DMF, r.t., 12 h, 96%; e) AcSH, DPAP, MeOH, r.t., UV light (365 nm), 1 h, 47%; f) MeONa, MeOH, r.t., 2 h.



Scheme 12 - Reagents and conditions: a) EDC, HOBt, Et_3N, DMF dry, 0°C to r.t., 12 h, 95%; b) CH_2Cl_2, TFA, r.t., 2 h, 30%; b) AcSH, DPAP, MeOH, r.t., UV light (365 nm), 1 h, 53%.

Preparation of the Au Nanoparticles

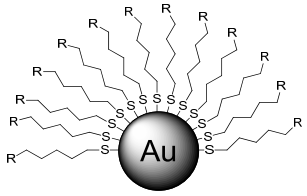
After the completion of the synthesis of the ligands of interest, the work was focused on the preparation of various types of gold nanoparticles, using the prepared ligands and a glucose derivative, *O*-(5'-thiopentyl)-β-D-glucopyranoside (Glc-C5-SH), synthesized according to Buskas, et al. [15] and Barrientos, et al. [16]

We used this glycoderivative to modulate and improve the water solubility of the nanoparticles, since compound **1** shows a low solubility in water (approximately 0,5 mg/mL). Furthermore, Glc-C5-SH should improve the multivalent effect of these derivatives, rather than interfere with their biological effects, since SGLT1-mediated anti-inflammatory effect is observed with huge amount of orally administered D-glucose.

The preparation of the gold nanoparticles (Au-NPs) involved a one-pot process, in which the thiol-ending ligands was dissolved in MeOH with a gold (III) compound, hydrogen tetrachloro aurate ($HAuCl_4$), to generate a kind of Au(I) polymer, on which thiol compounds are grafted. The successive addition of a reducing agent, as $NaBH_4$, provokes the passage

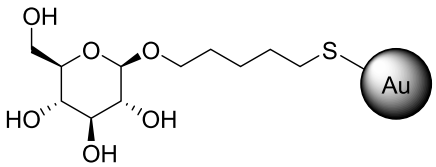
from Au(I) to Au^0, with the generation of the gold nanostructure, made of tens/hundreds of Au atoms and decorated, on the external side, with the ligands of interest[16] [8b]. Experimental details and procedures on gold nanoparticle preparation are described in the experimental section. The prepared Au-NPs are the follow:

General structure



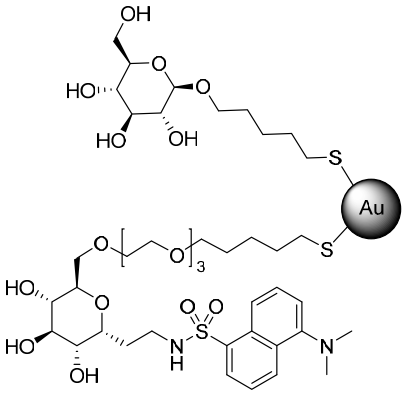
R = *C*-glycoside **1** or L-glutamine derivatives

AuNP1



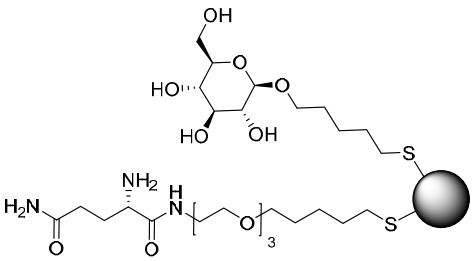
Au-NPs with Glc-C5-SH

AuNP2



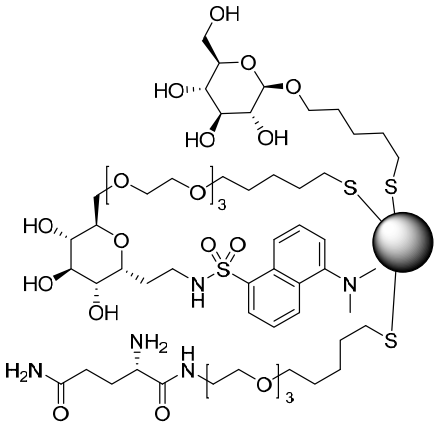
Au-NPs with Glc-C5-SH and 15

AuNP3



Au-NPs with Glc-C5-SH and **21**

AuNP4



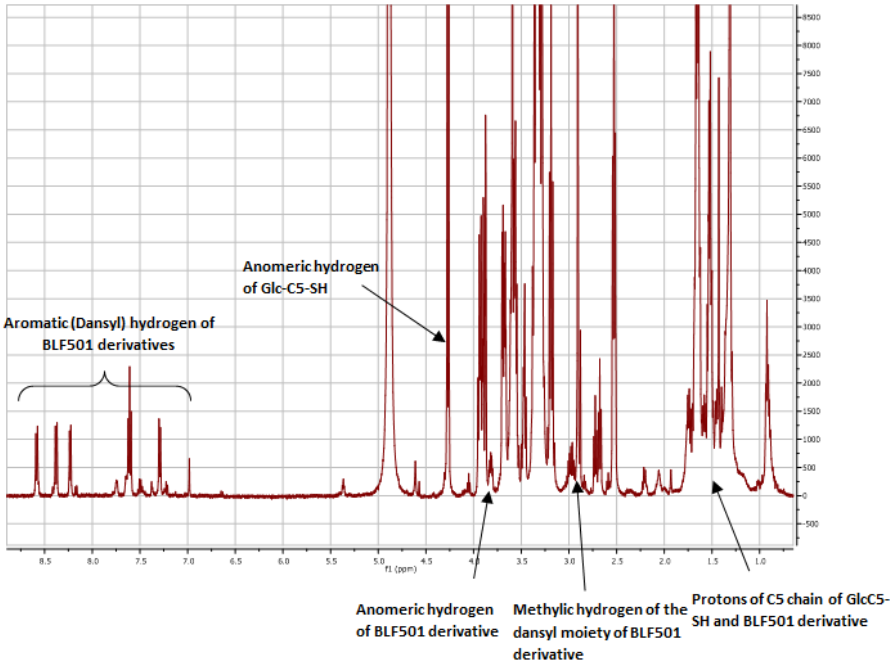
Au-NPs with Glc-C5-SH, **15** and **21**

The obtained nanoparticles, in a black-powder form, have been characterized according to the common experimental techniques, as ^{1}H NMR (with or without water signal suppression), TEM (Transmission Electron Microscopy) analysis, UV analysis. The NMR analysis is made dissolving the powder constituted by nanoparticles in deuterium oxide; the obtained spectrum shows broad typical signals of molecules grafted onto a solid structure/colloidal structure. If the ligands attached are one of different types, a preliminary ^{1}H-NMR analysis of the ligands-containing mixture has to be carried out, in order to check the real ligands ratio with which they are then loaded on the nanostructure. Comparing the nanoparticle spectrum with the spectrum of the ligand mixture before the nanoparticle formation it is possible to understand, in a qualitatively manner, the correct loading of the ligands the NP. However, the attachment of a particular ligand on a NP depends on several factors, including the chemical nature of the compound, its dimension, the spacer length, therefore the ligand loading must be always verified. For a quantitative data, a proton spectrum of the

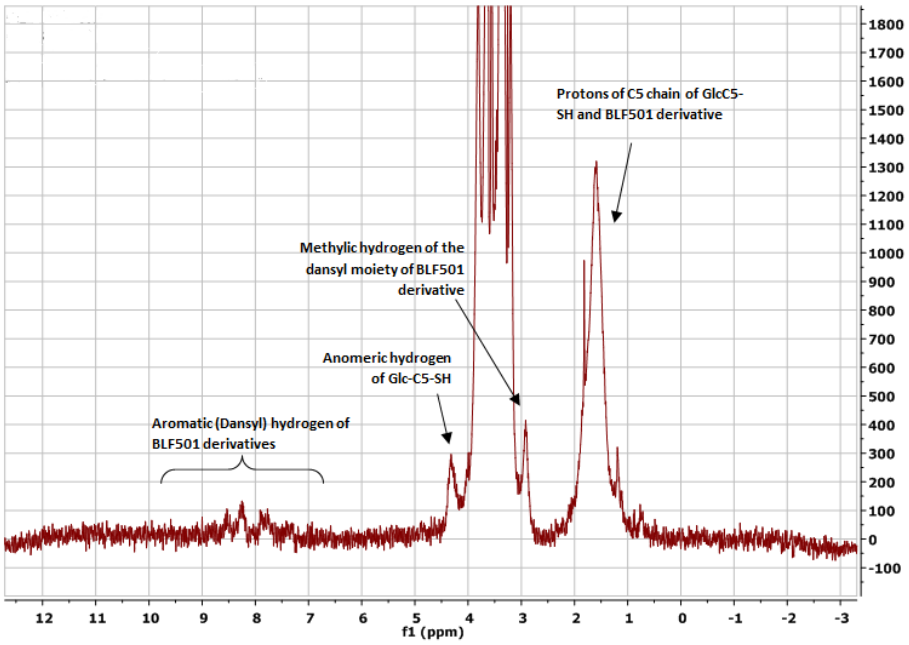
washing solution after nanoparticle formation (see procedure in experimental section) must be performed; comparing the resulting ratio of this mixture with the initial one is possible to determine in which ratio the ligands have been loaded.

Nanoparticle characterization

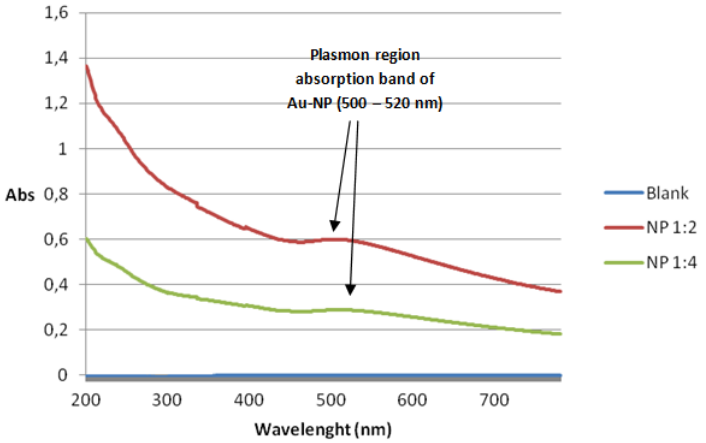
^{1}H-NMR of the solution of GlcC5-SH and *C*-glycoside **1** derivative (in CD_3OD) before the AuNP2 preparation. Some diagnostic signals are depicted. From the integration of the signal the ratio between the two ligands is 80:20 (Glc-C5-SH and *C*-glycoside **1** derivative).



^{1}H-NMR spectrum (with water suppression signal sequence) of the final AuNP2.

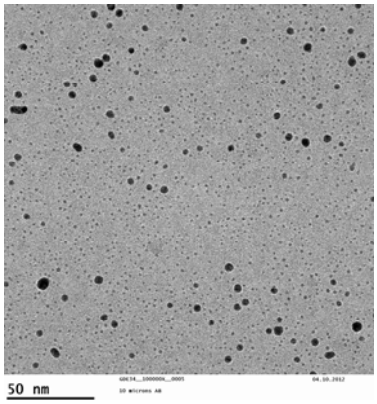


UV spectra of the AuNP2 dissolved in water; dilution 1:2 (1 mg in 2 mL of water, in red) and 1:4 (in green). Blue line is referred to water (blank).

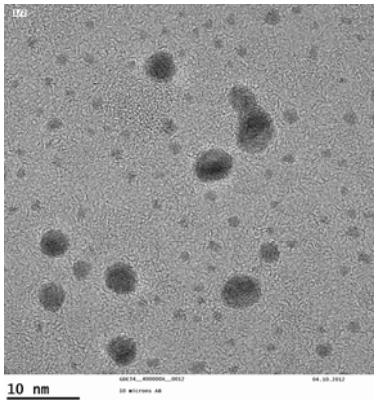


TEM analysis – transmission electron micrographs. On the left, micrograph of AuNP1. A double distribution of size (approximately 90% of ≈2 nm size and 10% of ≈5 nm size) can be seen. On the right, high magnification which shows the crystalline structure of the bigger gold nanoclusters. Below, the TEM analysis for AuNP 2,3,4 are presented.

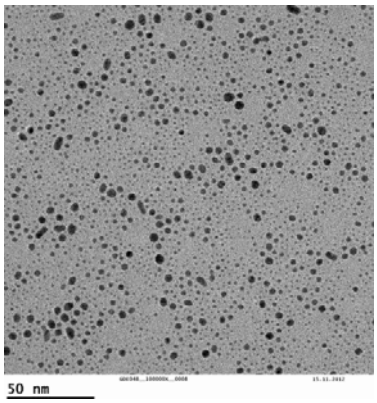
AuNP1



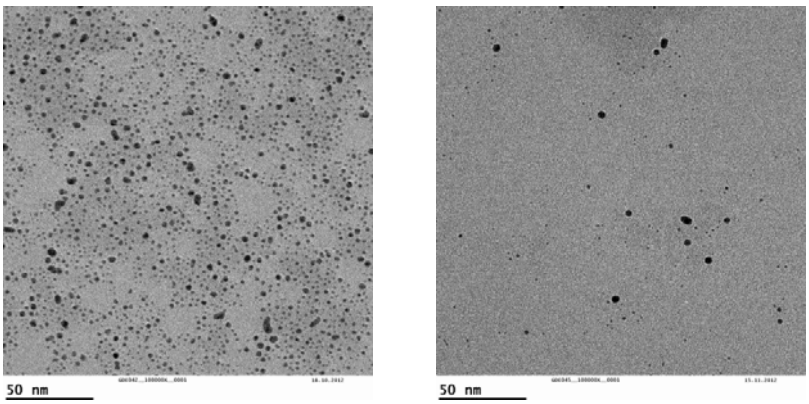
AuNP1



AuNP2



AuNP3 **AuNP4**



Further characterizations for the obtained nanoparticles will be performed.

Conclusions

In this work we have presented the preparation and synthesis of different ligands that are derivatives of *C*-glycoside **1** and L-glutamine, both protective agents against inflammatory insults *in vitro* and *in vivo*, as results from several literature data. These compounds were functionalized with a triethyleneglycol chain and a five-carbon spacer with a terminal thiol group in order to ensure the attachment on the gold nanoparticle, exploiting the affinity between the gold atom and thiol group, and to provide an adequate distance of these molecule from the core of the gold nanoparticle, allowing the binding of the ligands to their respective receptors. These ligands were used for the preparation of gold nanoparticle. A glucose derivative, with a C5 chain bearing a terminal thiol group is also used in the preparation of the nanoparticle to ensure the water solubility of the NPs. The Au-NPs will be used for biological studies in order to determine the multivalent/synergic activity

of the ligands attached to the nano-structure. The anti-inflammatory and protective role of the nanoparticles will be verified, in order to observe if a nano-structure with ligands attached in a multivalent fashion can increase the anti-inflammatory activity of the bioactive molecules. Furthermore, if a multivalent effect will be observed, decorated nanoparticles will be used as chemical tools in order to understand the nature of the multivalent phenomenon. On the basis of the obtained results, new nanoparticles could be designed and prepared, to enhance or modulate the anti-inflammatory activity. The nanoparticles could be tested against different kind of inflammation in which the protective role of SGLT1 and B^0AT1 is confirmed, as Chron' disease or IBDs (Intestinal Bowel Diseases).

Experimental Section

General remarks.

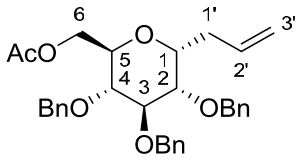
All commercial chemicals were purchased from Sigma-Aldrich and Alfa Aesar. All chemicals were used without further purification. Thin layer chromatography (TLC) was performed on silica gel 60 F_{254} plates (Merck) with detection under UV light when possible, or by charring with a solution of $(NH_4)_6Mo_7O_{24}$ (21g), $Ce(SO_4)_2$ (1g), concentrated H_2SO_4 (31 mL) in water (500 mL) or with an ethanol solution of ninhydrin. Flash-column chromatography was performed on silica gel 230–400 mesh (Merck) or on Isolera Four Flash Chromatography System (Biotage™). ^1H and ^{13}C NMR spectra were recorded at 25°C, unless otherwise stated, with a Varian Mercury 400-MHz instrument and on a Bruker DRX-300 spectrometer. Chemical shift assignments, reported in parts per million, were referenced to the corresponding solvent peaks. Mass spectra were recorded on a QTRAP system with ESI source. TEM examination was

carried out at 200 KeV with Philips CM200 microscope. UV spectra were obtained with a UV/vis Perkin-Elmer Lambda 12 spectrophotometer.

Synthesis of ligands

*Synthesis of C-glucoside derivative **15***

For compounds **2** and **3**, the synthesis and all characterization are already described in Cardona and La Ferla [12] and McGarvey, et al. [13].



2,3,4-tri-*O*-benzyl-6-*O*-acetyl-α-*C*-allyl-glucopyranoside **4**

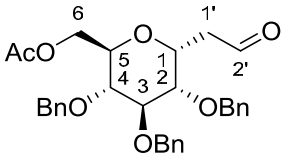
A mixture of Ac_2O/TFA 4:1 (70 mL), prepared at 0°C under argon atmosphere was added via a double tip needle to a round bottom flask containing 2 g (3,55 mmol) of 2,3,4,6-tetra-*O*-benzyl-α-D-glucopyranose. The solution was stirred vigorously at 0°C and the reaction followed by TLC (PE/AcOEt 8,5:1,5). After 1,5 h, no more starting compound is present and the solution is poured into ice-water and stirred for 10 min. The aqueous solution was extracted with AcOEt (3x) and the organic phase was then washed once with a sodium hydrogen carbonate saturated solution and twice with distilled water. After anhydrification, filtration and concentration of the remaining organic layer, the crude was purified by FC (PE/AcOEt 9:1), affording compound **4** (1,61 g, 3,12 mmol, 88% yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.43 – 7.24 (m, 15H, C*H* Ar), 5.79 (ddt, *J* = 17.2, 10.2, 6.9 Hz, 1H, H2'), 5.18 – 5.05 (m, 2H, H3'a,b), 4.98 (d, *J* = 10.8 Hz, 1H, CH_2Ph), 4.89 (d, *J* = 10.7 Hz, 1H, CH_2Ph), 4.83 (d, *J* = 10.8 Hz, 1H, CH_2Ph), 4.72 (d, *J* = 11.6 Hz, 1H, CH_2Ph), 4.65 (d, *J* = 11.6 Hz, 1H, CH_2Ph),

4.57 (d, *J* = 10.8 Hz, 1H, C*H*$_2$Ph), 4.24 (d, *J* = 3.5 Hz, 2H, H6a,b), 4.11 (dt, *J* = 9.3, 5.9 Hz, 1H, H1, H1), 3.85 (t, *J* = 9.0 Hz, 1H, H3), 3.76 (dd, *J* = 9.4, 5.8 Hz, 1H, H2), 3.70 (dt, *J* = 9.8, 3.5 Hz, 1H, H5), 3.48 (dd, *J* = 9.8, 8.7 Hz, 1H, H4), 2.50 (t, *J* = 8.2 Hz, 2H, H1'a,b), 2.04 (s, 3H, C*H*$_3$CO).

^{13}C NMR (101 MHz, $CDCl_3$) δ 170.92 (CH_3*C*O), 138.59 (Cq Ar), 138.15 (Cq Ar), 137.83 (Cq Ar), 134.40 (C2'), 128.80, 128.65, 128.59, 128.58, 128.49, 128.29, 128.11, 128.05, 128.00, 127.95, 127.85 (C Ar x 15), 117.24 (C3'), 82.37 (C3), 80.04 (C2), 77.89 (C4), 75.65, 75.21 (*C*H$_2$Ph), 73.68 (C1), 73.23 (*C*H$_2$Ph), 69.68 (C5), 63.62 (C6), 29.89 (C1'), 21.00 (*C*H$_3$CO).

$C_{32}H_{36}O_6$; calcd. mass: 516,63; MS-ESI: m/z 517,6 [M+H$^+$].



2-(2,3,4-tri-O-benzyl-6-acetyl-α-D-glucopyranosyl)-ethanal **5**

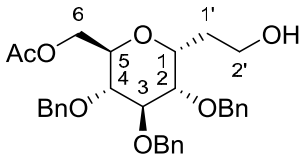
In a solution of **4** (3,53 g, 6,84 mmol) in CH_2Cl_2 (100 mL, 0,07 M) at -78°C, O_3 was bubbled until a pale blue colour appears, indicating the end of the reaction. The reaction was followed by TLC (PE/AcOEt 8:2). Then the excess of O3 was removed by purging the reaction with a stream of Argon at -78°C and then Ph_3P (17,1 mmol, 2,5 eq) were added. The reaction was stirred for 24 h at r.t. The product was purified directly from the crude reaction by flash chromatography (PE/AcOEt 8:2 – PE/AcOEt 7:3) to afford aldehyde **5** (2,92 g, 5,63 mmol, 82% yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 9.72 (s, 1H, H2' (C*H*O)), 7.43 – 7.24 (m, 15H, C*H* Ar), 4.93 (d, *J* = 10.9 Hz, 1H, C*H*$_2$Ph), 4.87 (d, *J* = 10.8 Hz, 1H, C*H*$_2$Ph), 4.83 (d, *J* = 10.9 Hz, 1H, C*H*$_2$Ph), 4.75 – 4.66 (m, 2H, C*H*$_2$Ph, H1), 4.63 (d, *J* = 11.5 Hz, 1H, C*H*$_2$Ph), 4.58 (d, *J* = 10.8 Hz, 1H, C*H*$_2$Ph), 4.30 – 4.18 (m, 2H, H6), 3.82 – 3.74 (m, 2H, H2, H3), 3.74 – 3.63 (m, 1H, H5), 3.53 – 3.43

(m, 1H, H4), 2.89 (dd, *J* = 15.8, 5.0 Hz, 1H, H1'a), 2.74 (ddd, *J* = 16.2, 8.6, 2.8 Hz, 1H, H1'b), 2.05 (s, 3H, C*H*$_3$CO).

^{13}C NMR (101 MHz, $CDCl_3$) δ 199.49 (C2' (*C*HO)), 170.83 (CH_3*C*O), 138.29, 137.67, 137.65 (Cq Ar), 128.66, 128.64, 128.61, 128.20, 128.18, 128.14, 127.99, 127.94 (C Ar), 81.90 (C2), 78.88 (C3), 77.30 (C4), 75.54 (*C*H$_2$Ph), 75.07 (*C*H$_2$Ph), 73.67 (*C*H$_2$Ph), 70.95 (C5), 69.58 (C1), 63.20 (C6), 41.26 (C1'), 20.95 (*C*H$_3$CO).

$C_{31}H_{34}O_7$; calcd. mass: 516,61; MS-ESI: m/z 519,6 [M+H$^+$].

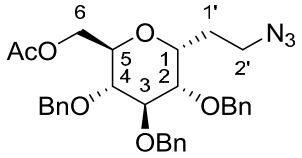


2-(2,3,4-tri-O-benzyl-6-acetyl-α-D-glucopyranosyl)-ethanol **6**

Compound **5** (2800 mg, 5,39 mmol) was dissolved in MeOH dry under argon. The solution was cooled to 0°C and $NaBH_4$ (2,164 mmol, 0,4 eq) was added in three portions. The reaction was stirred at 0°C and followed by TLC (PE/AcOEt 6:4). After 4 h, glacial acetic acid was added at 0°C until the pH is 3. Then the reaction is concentrated and the residue was redissolved in dichloromethane and washed with a HCl 1 N (1x) and brine (2x). The organic phase was dried, filtrated and evaporated. The crude was purified by flash chromatography (eluent PE/AcOEt 5,5:4,5 + 1% MeOH) to afford compound **6** (2672 mg, 5,14 mmol, 95% yield).

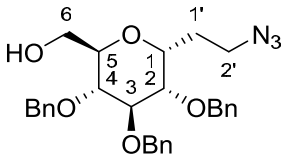
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.40 – 7.26 (m, 15H, m, C*H* Ar), 4.95 (d, *J* = 10.9 Hz, 1H, C*H*$_2$Ph), 4.88 (d, *J* = 10.9 Hz, 1H, C*H*$_2$Ph), 4.81 (d, *J* = 10.9 Hz, 1H, C*H*$_2$Ph), 4.73 (d, *J* = 11.6 Hz, 1H, C*H*$_2$Ph), 4.64 (d, *J* = 11.6 Hz, 1H, C*H*$_2$Ph), 4.59 (d, *J* = 10.9 Hz, 1H, C*H*$_2$Ph), 4.29 (dd, *J* = 11.7, 2.0 Hz, 1H, H6a), 4.26 – 4.14 (m, 2H, H1, H2), 3.87 – 3.75 (m, 4H, H2'a,b, H3, H5),

3.72 (dd, *J* = 9.2, 5.8 Hz, 1H, H6b), 3.43 (t, *J* = 9.0 Hz, 1H, H4), 2.44 (s, 1H, OH), 2.10 – 1.98 (m, 4H, C*H*$_3$CO, H1'a), 1.98 – 1.86 (m, 1H, H1'b).
^{13}C NMR (101 MHz, $CDCl_3$) δ 170.86 (CH_3*C*O), 138.41, 137.92, 137.74 (Cq Ar), 128.59, 128.57, 128.55, 128.13, 128.06, 128.03, 128.00, 127.98, 127.85 (C Ar), 81.87 (*C*H sugar), 79.46 (*C*H sugar), 77.82 (*C*H sugar), 75.49 (*C*H_2Ph), 75.00 (*C*H_2Ph), 73.37 (*C*H sugar), 73.33 (*C*H_2Ph), 70.30 (*C*H sugar), 63.77 (C6), 60.81 (C2'), 27.88 (C1'), 20.91 (*C*H_3CO).
$C_{31}H_{36}O_7$; calcd. mass: 520,62; MS-ESI: m/z 521,4 [M+H^+], 543,5 [M+Na^+].



2-(2,3,4-tri-O-benzyl-6-acetyl-α-D-glucopyranosyl)-1-azidoethane **7**
Alcohol **6** (2598 mg, 5 mmol) was dissolved in THF dry under argon. Ph_3P (15 mmol, 3 eq) was added and the solution was cooled to 0°C. DIAD and, after 10 min, $(PhO)_2PON_3$ were added dropwise to the reaction that was stirred at r.t.. The disappearance of the starting compound was followed by TLC (PE/AcOEt 7:3) and after 2 h the crude reaction was purified by flash chromatography (eluent PE/AcOEt 9:1). 2446 mg (4,48 mmol, 90% yield) of azide **7** were obtained.
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.38 – 7.26 (m, 15H, C*H* Ar), 4.94 (d, *J* = 10.9 Hz, 1H, C*H*$_2$Ph), 4.86 (d, *J* = 10.8 Hz, 1H, C*H*$_2$Ph), 4.80 (d, *J* = 10.8 Hz, 1H, C*H*$_2$Ph), 4.71 (d, *J* = 11.6 Hz, 1H, C*H*$_2$Ph), 4.62 (d, *J* = 11.6 Hz, 1H, C*H*$_2$Ph), 4.56 (d, *J* = 10.9 Hz, 1H, C*H*$_2$Ph), 4.23 (d, *J* = 3.4 Hz, 2H, H6a,b), 4.16 – 4.07 (m, 1H, H1), 3.82 – 3.70 (m, 2H, H2, H3), 3.68 – 3.60 (m, 1H, H5), 3.50 – 3.27 (m, 3H, H4, H2'a,b), 2.05 (s, 3H, C*H*$_3$CO), 1.96 (dd, *J* = 13.7, 6.5 Hz, 2H, H1'a,b).

^{13}C NMR (101 MHz, $CDCl_3$) δ 170.90 (CH_3CO), 138.45, 137.95, 137.76 (Cq Ar), 128.66, 128.61, 128.23, 128.13, 128.06, 127.92 (C Ar), 82.07 (C3), 79.48 (C2), 77.71 (C4), 75.60 (CH_2Ph), 75.17 (CH_2Ph), 73.38 (CH_2Ph), 71.34 (C1), 70.19 (C5), 63.64 (C6), 47.96 (C2'), 24.64 (C1'), 21.93 (CH_3CO). $C_{31}H_{35}N_3O_6$; calcd. mass: 545,64; MS-ESI: m/z 546,5 [$M+H^+$].



2-(2,3,4-tri-O-benzyl-α-D-glucopyranosyl)-1-azidoethane **8**

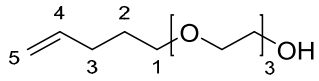
Pure azido derivative **7** (2420 mg, 4,44 mmol) was dissolved in MeOH (25 mL, ≈ 0,2 M) and ≈ 2 mL of a solution of 2,2 M sodium methoxide (2,2 M) in MeOH were added. The reaction was stirred at r.t. for 12 h. After the complete absence of starting compound checked by TLC (PE/AcOEt 7:3), an amount of acidic resin (IRA-120 H^+) was added the the reaction to neutralize the basicity of the solution. Then the resin was filtered off and the solution was concentrated, affording compound **8** (2116 mg, 4,21 mmol, 95% yield) which did not require further purification.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.41 – 7.27 (m, 15H, C*H* Ar), 4.94 (d, *J* = 10.9 Hz, 1H, CH_2Ph), 4.88 (d, *J* = 10.9 Hz, 1H, CH_2Ph), 4.82 (d, *J* = 10.9 Hz, 1H, CH_2Ph), 4.72 (d, *J* = 11.6 Hz, 1H, CH_2Ph), 4.67 – 4.60 (m, 2H, CH_2Ph), 4.11 (dd, *J* = 13.2, 7.2 Hz, 1H, H1), 3.83 – 3.74 (m, 2H, H6a, H3), 3.74 – 3.64 (m, 2H, H2, H6b), 3.56 – 3.45 (m, 2H, H4, H5), 3.45 – 3.37 (m, 1H, H2'a), 3.37 – 3.27 (m, 1H, H2'b), 1.98 (dd, *J* = 13.8, 7.0 Hz, 2H, H1'a,b), 1.87 (s, 1H, OH).

^{13}C NMR (101 MHz, $CDCl_3$) δ 138.60, 138.05, 138.02 (Cq Ar), 128.65, 128.57, 128.17, 128.09, 128.04, 128.01, 127.84 (C Ar), 82.11 (C3), 79.69

(C2), 77.85 (C4), 75.55 (*C*H_2Ph), 75.22 (*C*H_2Ph), 73.44 (*C*H_2Ph), 72.14 (C5), 71.38 (C1), 62.34 (C6), 48.04 (C2'), 24.71 (C1').

$C_{29}H_{33}N_3O_5$; calcd. mass: 503,60; MS-ESI: m/z 504,4 [M+H^+], 526,3 [M+H^+].



2-(2-(2-(pent-4-en-1-yloxy)ethoxy)ethoxy)ethan-1-ol **9**

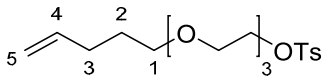
Compound **9** was prepared according to a procedure described in Svarovsky, et al. [17].

A solution of 1,006 mL of NaOH 50% (12,6 mmol) was added to 10 g of triethyleneglycol at 100°C. The reaction was stirred for 30 min at 100°C and 5-bromo-1-pentene (12,6 mmol, 0,19 eq) was added and the reaction was stirred at 100°C for 24 h. The reaction was followed by TLC (CH2Cl2/MeOH 9:1). After 24 h, the reaction was diluted with water and extracted six times with AcOEt. The organic phases were combined, dried over Na_2SO_4, evaporated and the crude was purified by FC (eluent CH_2Cl_2/MeOH 11:1). 2,039 g (9,34 mmol, 74% yield) of monopentenylated triethylenglycol were obtained.

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.75 (ddt, *J* = 16.9, 10.1, 6.6 Hz, 1H, H4), 4.96 (dd, *J* = 17.2, 1.6 Hz, 1H, H5a), 4.90 (d, *J* = 10.2 Hz, 1H, H5b), 4.08 (t, *J* = 5.1 Hz, 1H, OH), 3.66 (d, *J* = 3.7 Hz, 2H, C*H*$_2$ PEG), 3.63 – 3.57 (m, 6H, C*H*$_2$ PEG), 3.57 – 3.50 (m, 4H, C*H*$_2$ PEG), 3.42 (t, *J* = 6.7 Hz, 2H, H1a,b), 2.05 (q, *J* = 7.3 Hz, 2H, H3a,b), 1.68 – 1.57 (m, 2H, H2a,b).

^{13}C NMR (101 MHz, $CDCl_3$) δ 138.15 (C4), 114.84 (C5), 72.73, 70.80, 70.53, 70.44, 70.12, 69.94 (*C*H_2 PEG), 61.49 (C1), 30.15 (C3), 28.61 (C2).

$C_{11}H_{22}O_4$; calcd. mass: 218,29; MS-ESI: m/z 219,2 [M+H^+].



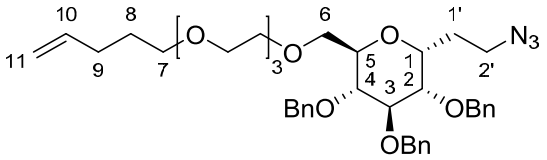
2-(2-(2-(pent-4-en-1-yloxy)ethoxy)ethoxy)ethyl-4-methylbenzenesulfonate **10**

To a solution of compound **9** (1,62 g, 7,421 mmol) and DMAP (0,148 mmol, 0,02 M) in CH_2Cl_2 (40 mL, ≈ 0,2 M), p-toluensulphonyl chloride (14,84 mmol, 2 eq) and Et_3N (18,55 mmol, 2,5 eq) were added. The reaction was stirred at r.t. and followed by TLC (CH_2Cl_2/MeOH 9,5:0,5). After 10 h, TLC indicated the disappearance of the starting material, and 40 mL of water were added. The organic phase was separated and the aqueous phase was washed three times with CH_2Cl_2. The combined organic phases were combined, dried and concentrated. The crude was purified by FC (CH_2Cl_2/MeOH 9,9:0,1), affording tosylate **10** (1,976 g, 5,3 mmol, 72 % yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.79 (d, *J* = 8.2 Hz, 2H, CH Ar), 7.33 (d, *J* = 8.0 Hz, 2H, CH Ar), 5.80 (ddt, *J* = 16.8, 10.1, 6.6 Hz, 1H, H4), 5.00 (dd, *J* = 17.1, 1.5 Hz, 1H, H5a), 4.94 (d, *J* = 9.4 Hz, 1H, H5b), 4.20 – 4.09 (m, 2H, CH_2CH_2OTs), 3.72 – 3.64 (m, 2H, CH_2CH_2OTs), 3.64 – 3.52 (m, 8H, -OCH_2CH_2O-), 3.45 (t, *J* = 6.7 Hz, 2H, H1a,b), 2.44 (s, 3H, CH_3PhSO_3), 2.10 (q, *J* = 7.2 Hz, 2H, H3a,b), 1.72 – 1.60 (m, 2H, H2a,b).

^{13}C NMR (101 MHz, $CDCl_3$) δ 144.90 (Cq Ar), 138.36 (C4), 133.04 (Cq Ar), 129.92, 128.08 (C Ar), 114.82 (C5), 70.86, 70.81, 70.74, 70.64, 70.17, 69.36, 68.77 (CH_2 PEG), 30.32 (C3), 28.84 (C2), 21.77 (CH_3PhSO_3).

$C_{18}H_{28}O_6S$; calcd. mass: 372,48; MS-ESI: m/z 373,4 [$M+H^+$].



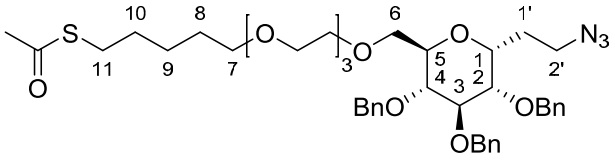
2-(2,3,4-tri-O-benzyl-6-(2-(2-(2-(pent-4-en-1-yloxy)ethoxy)ethoxy)ethyl)-α-D-glucopyranosyl)-1-azidoethane **11**

Compound **8** (1024 mg, 2,036 mmol) was dissolved in dry DMF (20 mL, 0,1 M) under argon at r.t.. 3,054 mmol (1,5 eq) of NaH (60% dispersion in mineral oil) were added and after 10 min a solution of tosylate **10** (3,9 mmol, 1,9 eq) in 5 mL of dry DMF was added dropwise. The reaction was stirred at r.t. under argon atmosphere and followed by TLC (PE/AcOEt 7:3). After 12 h, some drops of MeOH are added to the reaction to quench NaH and the solvent is evaporated. The residue was diluted with water and extracted three times with AcOEt. The organic phase was dried, filtered, evaporated and the crude purified by FC (eluent PE/AcOEt 7:3). 1,065 g (1,7 mmol, 83% yield) of compound **11** were obtained.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.35 – 7.22 (m, 15H, C*H* Ar), 5.79 (ddt, *J* = 16.9, 10.1, 6.6 Hz, 1H, H10), 4.99 (dd, *J* = 17.1, 1.6 Hz, 1H, H11a), 4.93 (d, *J* = 12.0 Hz, 1H, H11b), 4.90 (d, *J* = 11.1 Hz, 1H, CH_2Ph), 4.84 (d, *J* = 10.9 Hz, 1H, CH_2Ph), 4.79 (d, *J* = 10.9 Hz, 1H, CH_2Ph), 4.69 (d, *J* = 11.6 Hz, 1H, CH_2Ph), 4.63 (d, *J* = 11.0 Hz, 1H, CH_2Ph), 4.59 (d, *J* = 11.6 Hz, 1H, CH_2Ph), 4.15 – 4.05 (m, 1H, H1), 3.75 – 3.70 (m, 2H, C*H* sugar), 3.68 – 3.54 (m, 13H, C*H* sugar, CH_2 PEG), 3.54 – 3.47 (m, 3H, C*H* sugar), 3.42 (t, *J* = 6.7 Hz, 2H, H7a,b), 3.39 – 3.25 (m, 2H, H2'a,b), 2.08 (dd, *J* = 14.2, 7.3 Hz, 2H, H9a,b), 1.99 – 1.89 (m, 2H, H1'a,b), 1.70 – 1.60 (m, 2H, H8a,b).

^{13}C NMR (101 MHz, $CDCl_3$) δ 138.70 (Cq Ar), 138.39 (C10), 138.34 (Cq Ar), 138.09 (Cq Ar), 128.58, 128.55, 128.50, 128.04, 128.01, 127.98, 127.88, 127.74 (C Ar), 114.80 (C11), 82.31 (*C*H sugar), 79.63 (*C*H sugar), 77.98 (*C*H sugar), 75.59 (*C*H_2Ph), 75.16 (*C*H_2Ph), 73.37 (*C*H_2Ph), 71.67 (*C*H sugar),

71.43 (*C*H sugar), 71.10, 70.78, 70.73, 70.67, 70.59, 70.29, 70.15 (*C*H-*C*H_2 sugar, *C*H_2 PEG), 48.02 (C2'), 30.32 (C9), 28.84 (C8), 24.47 (C1').
$C_{40}H_{53}N_3O_8$; calcd. mass: 703,88; MS-ESI: m/z 704,7 [M+H^+].



2-(2,3,4-tri-O-benzyl-6-(1-thioacetylpentyl-(5-(2-(2-(2-hydroxyethoxy)ethoxy)ethyl)))-α-D-glucopyranosyl)-1-azidoethane **12**

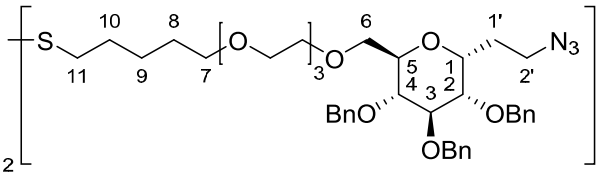
To a solution of compound **11** (384 mg, 0,545 mmol) in CH_2Cl_2 (5,4 mL, 0,1 M) contained in a sealed glass vial, thioacetic acid (0,656 mmol, 1,2 eq) and DPAP (0,055 mmol, 0,1 eq) were added. The reaction was stirred at r.t. under UV light explosion at 365 nm (using an UVA lamp of 4 W located 2 cm away from the glass vial). The reaction was followed by TLC (PE/AcOEt 7:3). After 2 h, the TLC indicated the absence of the starting material and the product was purified from the crude reaction by FC (eluent PE/AcOEt 7:3), affording thioester **12** (357 mg, 0,46 mmol, 85% yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.36 – 7.27 (m, 15H, C*H* Ar), 4.91 (d, *J* = 10.9 Hz, 1H, CH_2Ph), 4.85 (d, *J* = 10.9 Hz, 1H, CH_2Ph), 4.81 (d, *J* = 10.9 Hz, 1H, CH_2Ph), 4.70 (d, *J* = 11.6 Hz, 1H, CH_2Ph), 4.67 – 4.58 (m, 2H, CH_2Ph), 4.16 – 4.07 (m, 1H, H1), 3.75 – 3.71 (m, 2H, C*H* sugar), 3.69 – 3.47 (m, 16H, C*H* sugar, CH_2 PEG), 3.44 – 3.27 (m, 4H, H7a,b, H2'a,b), 2.85 (t, *J* = 7.3 Hz, 2H, H11a,b), 2.31 (s, 3H, CH_3COS), 2.00 – 1.90 (m, 2H, H1'a,b), 1.62 – 1.52 (m, 4H, H10a,b, H8a,b), 1.46 – 1.34 (m, 2H, H9a,b).

^{13}C NMR (101 MHz, $CDCl_3$) δ 196.10 (CH_3*C*OS), 138.73, 138.37, 138.12 (Cq Ar), 128.61, 128.58, 128.52, 128.07, 128.03, 128.01, 127.91, 127.76 (C Ar), 82.34 (*C*H sugar), 79.67 (*C*H sugar), 78.02 (*C*H sugar), 75.61

(C*H*$_2$Ph), 75.19 (C*H*$_2$Ph), 73.40 (C*H*$_2$Ph), 71.71 (*C*H sugar), 71.46 (*C*H sugar), 71.22, 71.12, 70.75, 70.70, 70.61, 70.34, 70.18 (*C*H$_2$ sugar, *C*H$_2$ PEG), 48.06 (C2'), 30.79 (*C*H$_3$COS), 29.49 (C8), 29.23 (C10), 29.16 (C11), 25.49 (C9), 24.51 (C1').

$C_{42}H_{57}N_3O_9S$; calcd. mass: 779,99; MS-ESI: m/z 780,6 $[M+H^+]$, 752,6 $[M-N_2+H]^+$, 802,6 $[M+Na]^+$, 818,6 $[M+K]^+$.

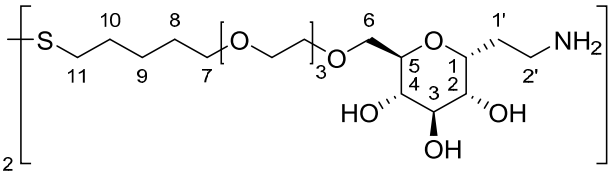


2-(2,3,4-tri-O-benzyl-6-(1-thio-(5-(2-(2-(2-hydroxyethoxy)ethoxy)ethyl)))-α-D-glucopyranosyl)-1-azidoethane (dimer) **13**

Pure thioester **12** (38 mg, 0,049 mmol) was dissolved in MeOH (1 mL, 0,05 M) and 1,3 mg (0,024 mmol, 0,5 eq) of solid sodium methylate were added. The reaction is stirred at r.t., free to air exposure to promote the oxidation of the thiol group into disulfide bridge. After 24 h, TLC (PE/AcOEt 6,5:3,5) indicated the end of the reaction, with the complete absence of starting compound. The reaction is concentrated to dryness and ^{1}H-NMR indicated the complete removal of acetyl group and the formation of a disulfide bond (triplet signal at 2.65 ppm). The residue was re-suspended in CH_2Cl_2 and the undissolved material was filtered off and discarded. The crude was used directly for the next reaction.

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.37 – 7.24 (m, 15H, CH Ar), 4.94 – 4.90 (m, 1H, C*H*$_2$Ph), 4.85 (d, *J* = 10.9 Hz, 1H, C*H*$_2$Ph), 4.81 (d, *J* = 10.9 Hz, 1H, C*H*$_2$Ph), 4.72 – 4.68 (m, 1H, C*H*$_2$Ph), 4.65 – 4.59 (m, 2H, C*H*$_2$Ph), 4.12 (dt, *J* = 9.8, 5.0 Hz, 1H, H1), 3.73 (dd, *J* = 8.5, 6.0 Hz, 2H, H2, H6a), 3.70 – 3.49 (m, 16H, CH2 PEG x 12, H3, H4, H5, H6b, H6b), 3.44 – 3.29 (m, 4H, H7a,b, H2'a,b), 2.71 – 2.62 (m, 2H, H11a,b), 2.01 – 1.89 (m, 2H, H1'a,b), 1.68 (dt,

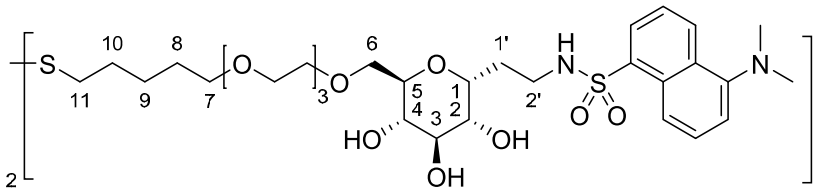
J = 15.0, 7.5 Hz, 2H, H10a,b), 1.62 – 1.52 (m, 2H, H8a,b), 1.47 – 1.37 (m, 2H, H9a,b).



2-(6-(1-thio-(5-(2-(2-(2-hydroxyethoxy)ethoxy)ethyl)))-α-D-glucopyranosyl)-1-azidoethane (dimer) **14**

Compound **14** was prepared by a Birch reduction of protected derivative **13**, according to a procedure described by Dettmann and Ziegler [18], slightly modified. Some drops of liquid ammonia were collected in a two-necked round bottomed flask cooled to – 78°C and a small piece of sodium (approx. 20 – 30 mg) was added. The solution turned blue immediately and after 10 minutes, a THF solution (1,4 mL) of crude compound **13** as monomer) was added. The reaction was stirred at -78°C for 15 min, then NH_4Cl (≈ 100 mg) was added and the reaction was warmed to r.t.. The solvent was evaporated and the crude was purified by Reverse-Phase C18 Cartridge, and 14 mg (0,032 mmol as monomer, 65% yield over two steps) of **14** were obtained.

^{1}H NMR (500 MHz, MeOD) δ 4.11 (ddd, *J* = 11.6, 5.1, 2.7 Hz, 1H, H1), 3.79 – 3.58 (m, 18H, C*H* sugar, CH_2 PEG), 3.51 (t, *J* = 6.6 Hz, 2H, H7a,b), 3.23 – 3.18 (m, 1H, H2'a), 3.17 – 3.10 (m, 1H, H2'b), 2.74 – 2.67 (m, 2H, H11a,b), 2.15 – 2.05 (m, 1H, H1'a), 1.99 – 1.90 (m, 1H, H1'b), 1.76 – 1.68 (m, 2H, H10a,b), 1.66 – 1.57 (m, 2H, H8a,b), 1.51 – 1.43 (m, 2H, H9a,b).

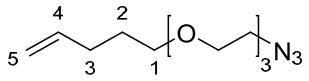


5-(N,N-dimethylamino)-N-(2-(6-(1-thio-(5-(2-(2-(2-hydroxyethoxy)ethoxy)ethyl)))-α-D-glucopyranosyl)-naphthalene-1-sulfonamide (dimer) **15**

Compound **14** (10 mg, 0,023 mmol as monomer) was dissolved in MeOH; Et_3N (10 µL, 3 eq) and dansyl chloride (16 mg, 2,6 eq) were added and the reaction was stirred at r.t. for 10 h. The reaction was followed by TLC (AcOEt/MeOH 9:1 and AcOEt/MeOH/H2O/AcOH 5:5:1:1). After 10 h, the crude was purified by FC (AcOEt/MeOH 8:2) and then with Reverse-Phase C18 Cartridge, affording 6 mg (0,009 mmol, 39% yield) of compound **15**.

^{1}H NMR (500 MHz, MeOD) δ 8.54 (t, J = 14.1 Hz, 1H, C*H* Ar), 8.36 (d, J = 8.7 Hz, 1H, C*H* Ar), 8.21 (dd, J = 7.3, 1.1 Hz, 1H, C*H* Ar), 7.59 (t, J = 7.7 Hz, 2H, C*H* Ar), 7.27 (d, J = 7.3 Hz, 1H, C*H* Ar), 3.86 – 3.77 (m, 1H, H1), 3.68 – 3.63 (m, 2H, H6a, CH_2 TEG x 1), 3.62 – 3.55 (m, 8H, CH_2 TEG), 3.55 – 3.51 (m, 2H, H6b, CH_2 TEG x 1), 3.49 – 3.42 (m, 3H, H2, H7a,b), 3.39 – 3.34 (m, 2H, H3, H5), 3.18 – 3.11 (m, 1H, H4), 3.00 – 2.91 (m, 2H, H2'a,b), 2.88 (s, 6H, (CH_3)$_2$N-), 2.65 (t, J = 7.2 Hz, 2H, H11a,b), 1.77 – 1.71 (m, 2H, H1'a,b), 1.69 – 1.62 (m, 2H, H10a,b), 1.60 – 1.51 (m, 2H, H8a,b), 1.47 – 1.39 (m, 4H, H9a,b).

*Synthesis of L-glutamine derivative **21** and **24***



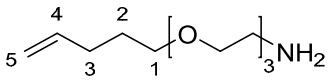
5-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)pent-1-ene **16**

To a solution of Tosylate **10** (1976 mg, 5,3 mmol) in dry DMF (35 mL), 15,9 mmol (3 eq) of NaN_3 were added. The reaction was stirred at r.t. for 2 h and then at 80°C for 10 h. The reaction was followed by TLC (PE/AcOEt 8:2). At the end of the reaction, the suspension was diluted with water and extracted with AcOEt (3x). The organic phases were combined, dried over sodium sulphate and evaporated. The residue was purified by FC (PE/AcOEt 8:2) affording azide **16** (1012 mg, 4,16 mmol, 78% yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.80 (ddt, *J* = 16.9, 10.2, 6.7 Hz, 1H, H4), 5.01 (dd, *J* = 17.1, 1.6 Hz, 1H, H5a), 4.95 (d, *J* = 10.2 Hz, 1H, H5b), 3.72 – 3.61 (m, 8H, -OC*H*$_2$C*H*$_2$-), 3.61 – 3.54 (m, 2H, -OC*H*$_2$C*H*$_2$-), 3.46 (t, *J* = 6.7 Hz, 2H, H1a,b), 3.38 (t, *J* = 5.0 Hz, 2H, -OCH$_2$C*H*$_2$N$_3$), 2.10 (q, *J* = 7.3 Hz, 2H, H3a,b), 1.73 – 1.61 (m, 2H, H2a,b).

^{13}C NMR (101 MHz, $CDCl_3$) δ 138.40 (C4), 114.80 (C5), 70.82, 70.81, 70.77, 70.21, 70.15 (*C*H$_2$ PEG), 50.78 (*C*H$_2$N$_3$), 30.34 (C3), 28.86 (C2).

$C_{11}H_{21}N_3O_3$; calcd. mass: 243,32; MS-ESI: m/z 244,2 [M+H$^+$].



2-(2-(2-(pent-4-en-1-yloxy)ethoxy)ethoxy)ethan-1-amine **17**

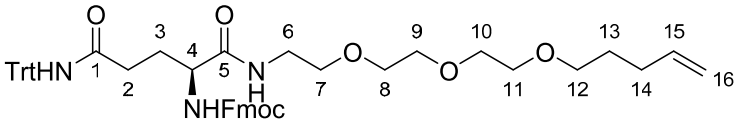
Compound **16** (1153 mg, 4,73 mmol) was dissolved in THF (30 mL) and Ph_3P (5,68 mmol, 1,2 eq) was added. The reaction was stirred at r.t. for 2 h, then 5,67 mmol (1,2 eq) of distilled water were added and the reaction was heated to 60°C and stirred for 30 min, following the

disappearance of the starting material by TLC (PE/AcOEt 8:2 and CH_2Cl_2/MeOH 9:1). The solvent was evaporated and the crude purified by FC (CH_2Cl_2/MeOH 9:1 to CH_2Cl_2/MeOH/NH_3 (aq) 7:3:1), affording amine **17** (0,951 g, 4,4 mmol, 93% yield).

^{1}H NMR (400 MHz, CD_3OD) δ 5.83 (ddt, *J* = 13.5, 10.1, 6.7 Hz, 1H, H4), 5.02 (dd, *J* = 17.1, 1.1 Hz, 1H, H5a), 4.95 (d, *J* = 10.1 Hz, 1H, H5b), 3.69 – 3.53 (m, 12H, C*H*$_2$ PEG), 3.52 – 3.45 (m, 2H, C*H*$_2$$NH_2$), 2.12 (dd, *J* = 14.4, 7.2 Hz, 2H, H3a,b), 1.72 – 1.58 (m, 2H, H2a,b).

^{13}C NMR (101 MHz, CD_3OD) δ 139.44 (C4), 115.23 (C5), 73.67, 71.56, 71.50, 71.26, 71.12, 70.39 (C1, *C*H_2 PEG), 41.42 (*C*$H_2$$NH_2$), 31.33 (C3), 30.03 (C2).

$C_{11}H_{23}NO_3$; calcd. mass: 217,31; MS-ESI: m/z 218,3 [M+H$^+$].



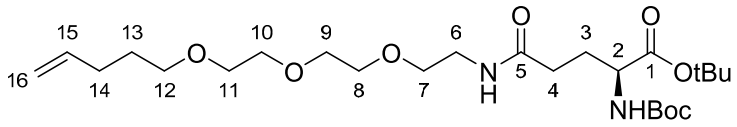
(9H-fluoren-9-yl)methyl-(S)-(3,7-dioxo-1,1,1-triphenyl-11,14,17-trioxa-2,8-diazadocos-21-en-6-yl) carbamate **18**

Compound **17** (133 mg, 0,612 mmol) was dissolved in dry CH_2Cl_2 (6 mL, 0,1 M) under argon atmosphere. The solution was cooled to 0°C and Fmoc-L-Gln(Trt)-OH (0,64 mmol, 1,04 eq), HOBt (0,918 mmol, 1.5 eq) and EDC (0,918 mmol, 1,5 eq) were added. The reaction was warmed to r.t. and stirred under argon. After 15 h, TLC (CH_2Cl_2/MeOH 9,5:0,5 and CH_2Cl_2/MeOH/NH_3 (aq) 9:1:0,1) indicated the absence of starting compound 17; the reaction was diluted with CH_2Cl_2 and washed with water (3x). The organic phase was gathered, dried over sodium sulphate, filtrated and evaporated. The crude was purified by FC on Biotage System (gradient CH_2Cl_2 100% – CH_2Cl_2/MeOH 9,3:0,7). 390 mg (0,48 mmol) of compound **18** were obtained (79% yield).

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.80 (d, *J* = 7.5 Hz, 2H, CH Ar Fmoc), 7.63 (d, *J* = 7.4 Hz, 2H CH Ar Fmoc), 7.43 (t, *J* = 7.4 Hz, 2H CH Ar Fmoc), 7.38 – 7.20 (m, 17H, CH Ar Fmoc, CH Ar Trt), 7.11 (s, 1H, CON*H*Trt), 6.84 (s, 1H, CON*H* PEG), 6.04 (d, *J* = 6.7 Hz, 1H, N*H* Fmoc), 5.83 (ddt, *J* = 16.9, 10.2, 6.6 Hz, 1H, H15), 5.04 (ddd, *J* = 17.1, 3.3, 1.5 Hz, 1H, H16a), 5.01 – 4.96 (m, 1H, H16b), 4.41 (d, *J* = 7.1 Hz, 2H, CH_2 Fmoc), 4.24 (t, *J* = 7.0 Hz, 1H, C*H* Fmoc), 4.19 – 4.10 (m, 1H, H4), 3.64 – 3.51 (m, 10H, CH_2 PEG), 3.49 – 3.40 (m, 4H, H12a,b, H6a,b), 2.58 – 2.49 (m, 1H, H2a), 2.49 – 2.39 (m, 1H, H2b), 2.18 – 2.07 (m, 3H, H14a,b, H3a), 2.03 – 1.95 (m, 1H, H3b), 1.73 – 1.65 (m, 2H, H13a,b).

^{13}C NMR (101 MHz, $CDCl_3$) δ 171.80 (C5), 171.32 (C1), 156.23 (O*C*ONH Fmoc), 144.48 (Cq Ar Trt), 143.80 (Cq Ar Fmoc), 143.68 (Cq Ar Fmoc), 141.17 (Cq Ar Fmoc), 141.14 (Cq Ar Fmoc), 138.11 (C15), 128.64 (C Ar Trt, C Ar Fmoc), 127.82 (C Ar Trt, C Ar Fmoc), 127.62 (C Ar Trt, C Ar Fmoc), 127.01 (C Ar Trt, C Ar Fmoc), 126.88 (C Ar Trt, C Ar Fmoc), 125.10 (C Ar Fmoc), 119.87 (C Ar Fmoc), 114.73 (C16), 70.56, 70.48, 70.37, 70.34, 70.07, 69.90 (*C*H_2 PEG, Cq Trt), 69.30 (C12), 66.76 (*C*H_2 Fmoc), 54.06 (C4), 47.06 (*C*H Fmoc), 39.20 (C6), 33.17 (C2), 30.11 (C14), 29.33 (C3), 28.59 (C13).

$C_{50}H_{55}N_3O_7$; calcd. mass: 810,00; MS-ESI: m/z 810,9 [$M+H^+$].



tert-butyl (S)-2-((tert-butoxycarbonyl)amino)-5-oxo-9,12,15-trioxa-6-azaicos-19-enoate **22**

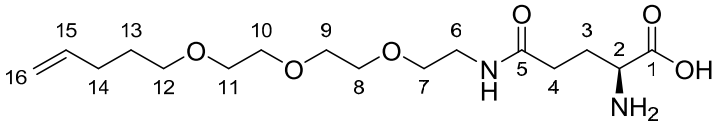
Boc-L-glutamic acid 1-tert-butyl ester (149 mg, 0,49 mmol) and compound **17** (0,58 mmol, 1,2 eq) were dissolved in CH_2Cl_2 (4,9 mL, 0,1 M) under argon and 0,98 mmol (2 eq) of Et_3N were added and the

solution was stirred at r.t. for 15 min. Then the temperature was set to 0°C and HOBt (0,73 mmol, 1,5 eq) and DCC (0,73 mmol, 1,5 eq) were added. The reaction was stirred at r.t. until TLC (CH_2Cl_2/MeOH 9,5:0,5) indicated the end of the coupling reaction. The reaction was recovered as described for compound **18** and the product was purified by FC (eluent CH_2Cl_2/MeOH 9,8:0,2) affording compound **22** (207 mg, 0,47 mmol, 95% yield)

^{1}H NMR (400 MHz, $CDCl_3$) δ 6.44 (s, 1H, CON*H*-PEG), 5.80 (td, *J* = 16.9, 6.7 Hz, 1H, H15), 5.26 (d, *J* = 7.6 Hz, 1H, N*H*Boc), 5.01 (d, *J* = 17.2 Hz, 1H, H16a), 4.95 (d, *J* = 10.2 Hz, 1H, H16b), 4.14 (s, 1H, H2), 3.67 – 3.52 (m, 10H, C*H*$_2$ PEG), 3.51 – 3.36 (m, 4H, H12a,b, H6a,b), 2.30 – 2.20 (m, 2H, H4a,b), 2.18 – 2.00 (m, 3H, H14a,b, H3a), 1.95 – 1.81 (m, 1H, H3b), 1.74 – 1.62 (m, 2H, H13a,b), 1.50 – 1.37 (m, 18H, (C*H*$_3$)$_3$ Boc, tBu).

^{13}C NMR (101 MHz, $CDCl_3$) δ 172.20 (C5), 171.57 (C1), 155.81 (O*C*ONH Boc), 138.25 (C15), 114.87 (C16), 82.14 (Cq OtBu), 79.78 (Cq Boc), 70.77, 70.59, 70.33, 70.12, 69.87 (C7-C12), 53.71 (C2), 39.37 (C6), 32.70 (C4), 30.27 (C14), 29.14 (C3), 28.78 (C13), 28.40, 28.05 ((*C*H$_3$)$_3$ OtBu, Boc).

$C_{25}H_{46}N_2O_8$; calcd. mass: 502,65; MS-ESI: m/z 503,4 [M+H$^+$].



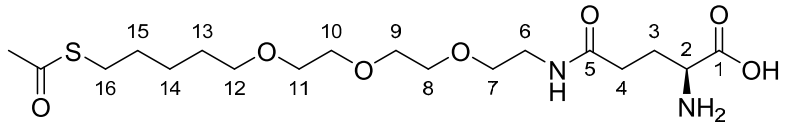
(S)-2-amino-5-oxo-9,12,15-trioxa-6-azaicos-19-enoic acid **23**

Protected L-glutamine derivative **22** (246 mg, 0,49 mmol) was dissolved in a mixture of CH_2Cl_2/TFA (12 + 3 mL) and stirred at r.t.. The reaction was followed by TLC (CH_2Cl_2/MeOH/NH_3 (aq) 7:3:0,3) and after 2 h the solution was neutralized with NH_3 (aq) 25%. The crude was concentrated and the product purified by FC (CH_2Cl_2/MeOH/NH_3 (aq) 7:3:0,2) obtaining 52 mg (0,15 mmol, 30% yield) of compound **23**.

^{1}H NMR (400 MHz, CD_3OD) δ 5.83 (ddt, *J* = 17.0, 10.1, 6.8 Hz, 1H, H15), 5.07 – 4.96 (m, 2H, H16a,b), 3.72 (t, *J* = 5.9 Hz, 1H, H2), 3.66 – 3.57 (m, 8H, H8a,b-H11a,b), 3.55 (t, *J* = 5.4 Hz, 2H, H7a,b), 3.49 (t, *J* = 6.5 Hz, 2H, H12a,b), 3.37 (t, *J* = 5.3 Hz, 2H, H6a,b), 2.46 (t, *J* = 7.1 Hz, 2H, H4a,b), 2.20 – 2.03 (m, 4H, H14a,b, H3a,b), 1.72 – 1.62 (m, 2H, H13a,b).

^{13}C NMR (101 MHz, CD_3OD) δ 174.87 (C1), 173.28 (C5), 139.43 (C15), 115.24 (C16), 71.57, 71.51, 71.24, 71.08, 70.47 (C7-C12), 55.10 (C2), 40.47 (C6), 32.91 (C4), 31.34 (C14), 29.98 (C13), 27.79 (C3).

$C_{16}H_{30}N_2O_6$; calcd. mass: 346,42; MS-ESI: m/z 347,3 [$M+H^+$].



(S)-22-amino-2,19-dioxo-9,12,15-trioxa-3-thia-18-azatricosan-23-oic acid

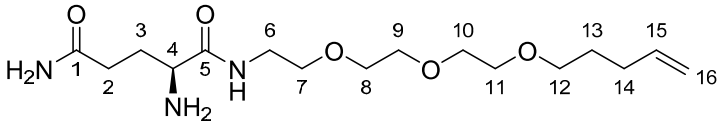
24

50 mg of compound **23** (0,144 mmol) were dissolved in MeOH (1,5 mL, 0,1 M).Thiolacetic acid (0,19 mmol, 1,3 eq) and DPAP (0,03 mmol, 0,2 eq) were added to this solution and the reaction was stirred at r.t. under UV light exposition at 365 nm. After 2 h TLC (CH_2Cl_2/MeOH/NH_3 (aq) 7:3:0,2) indicated the complete conversion of the starting compound; the reaction was concentrated and the residue was purified by FC (eluent CH_2Cl_2/MeOH/NH_3 (aq) 7:3:0,2). 32 mg (0,075 mmol) of thioester **24** were obtained (53% yield).

^{1}H NMR (500 MHz, CD_3OD) δ 3.66 – 3.60 (m, 7H, CH_2 PEG x 6, H2), 3.60 – 3.57 (m, 2H, CH_2 PEG x 2), 3.55 (t, *J* = 5.4 Hz, 2H, H7), 3.48 (t, *J* = 6.5 Hz, 2H, H12), 3.37 (t, *J* = 5.5 Hz, 2H, H6), 2.87 (t, *J* = 7.3 Hz, 2H, H16a,b), 2.45 (t, *J* = 7.1 Hz, 2H, H4a,b), 2.30 (s, 3H, SCOCH_3), 2.18 – 2.04 (m, 2H, H3a,b), 1.64 – 1.55 (m, 4H, H15a,b, H13a,b), 1.48 – 1.40 (m, 2H, H14a,b).

^{13}C NMR (101 MHz, CD_3OD) δ 197.51 (CH_3COS), 190.57, 175.09 (C1), 173.52 (C5), 72.02, 71.57, 71.52, 71.25, 71.09, 70.50 (CH_2 PEG), 55.63 (C2) 40.47 (C6), 33.07 (C4), 30.60 (C13), 30.53 ($SCOCH_3$), 30.14 (C15), 29.76 (C16), 27.97 (C3), 26.36 (C14).

$C_{18}H_{34}N_7O_7S$; calcd. mass: 422,54; MS-ESI: m/z 423,3 $[M+H^+]$.

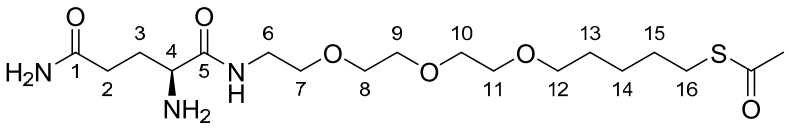


(S)-2-amino-N1-(2-(2-(2-(pent-4-en-1-yloxy)ethoxy)ethoxy)ethyl)pentanediamide **19**

Compound **18** (148 mg, 0,18 mmol) was dissolved in 1,8 mL of CH_2Cl_2 (0,1 M) and 0,4 mL of TFA were added. After 6 h, TLC (CH_2Cl_2/MeOH 9,5:0,5) showed no more starting compound (Trt deprotection), and the reaction was neutralized to pH 7 adding NH_3 25%. Then the solvent was evaporated and the residue was redissolved in DMF (8 mL). Piperidine was added (2 mL) and the reaction was stirred at r.t.. After 12 h TLC confirmed the Fmoc deprotection; the reaction was concentrated and the residue purified by FC (eluent CH_2Cl_2/MeOH/NH_3 (aq) 8,5:1,5:0,1). 60 mg (0,17 mmol) of deprotected compound **19** were gathered (96% yield).

^{1}H NMR (500 MHz, CD_3OD) δ 5.83 (ddt, *J* = 17.0, 10.2, 6.7 Hz, 1H, H15), 5.02 (ddd, *J* = 17.1, 3.5, 1.6 Hz, 1H, H16a), 4.95 (ddt, *J* = 10.2, 2.2, 1.2 Hz, 1H, H16b), 3.67 – 3.60 (m, 6H, CH_2 PEG), 3.60 – 3.54 (m, 4H, CH_2 PEG), 3.48 (t, *J* = 6.5 Hz, 2H, H12a,b), 3.46 – 3.33 (m, 3H, H4, H6a,b), 2.28 (t, *J* = 7.8 Hz, 2H, H2a,b), 2.15 – 2.09 (m, 2H, H14a,b), 1.99 – 1.90 (m, 1H, H3a), 1.88 – 1.79 (m, 1H, H3b), 1.70 – 1.62 (m, 2H, H13a,b).

$C_{16}H_{31}N_3O_5$; calcd. mass: 345,44; MS-ESI: m/z 346,4 $[M+H^+]$, 368,4 $[M+Na]^+$.



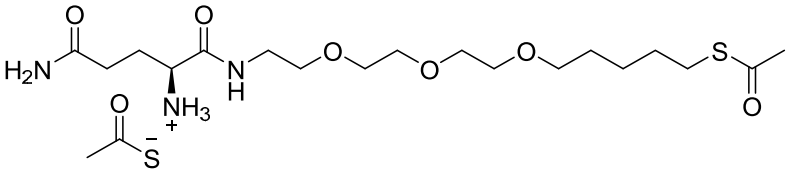
(S)-S-(1,4-diamino-1,5-dioxo-9,12,15-trioxa-6-azaicosan-20-yl)-ethanethioate **20**

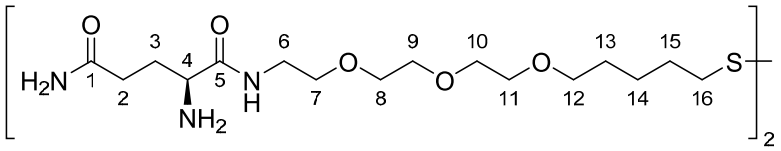
Compound **19** (60 mg, 0,174 mmol) was converted in the thioester derivative **20** as described for compound **24**. After FC of the crude reaction (eluent CH_2Cl_2/MeOH/NH_3 (aq) 8,5:1,5:0,1) 29 mg of compound **20** (0,07 mmol, 47% yield) were obtained. The ^{1}H-NMR analysis suggests that the product is as thioacetate salt.

^{1}H NMR (500 MHz, MeOD) δ 4.32 (dd, *J* = 8.4, 5.7 Hz, 1H, H4), 3.66 – 3.60 (m, 6H, C*H*$_2$ PEG), 3.60 – 3.53 (m, 4H, C*H*$_2$ PEG), 3.47 (t, *J* = 6.5 Hz, 2H, H12a,b), 3.40 (t, *J* = 5.4 Hz, 1H, H6a), 3.36 (t, *J* = 5.5 Hz, 1H, H6b), 2.87 (t, *J* = 7.3 Hz, 2H, H16a,b), 2.34 – 2.26 (m, 5H, C*H*$_3$COSCH$_2$, H2a,b), 2.11 – 2.01 (m, 1H, H3a), 1.99 (s, 3H, C*H*$_3$COS$^-$), 1.95 – 1.86 (m, 1H, H3b), 1.63 – 1.55 (m, 4H, H13a,b, H15a,b), 1.47 – 1.39 (m, 2H, H14a,b).

^{13}C NMR (101 MHz, CD_3OD) δ 197.53 (CH_3*C*OSCH$_2$), 177.69 (CH_3*C*OS$^-$), 173.91 (C5), 173.30 (C1), 72.04, 71.62, 71.56, 71.28, 71.14, 70.44 (*C*H$_2$ PEG), 54.38 (C2), 40.39 (C6), 32.56 (C4), 30.62 (C13), 30.52 (SCO*C*H$_3$), 30.17 (C15), 29.77 (C16), 29.10 (C3), 26.39 (C14), 22.52 (*C*H$_3$COS$^-$).

structure proposed from spectra





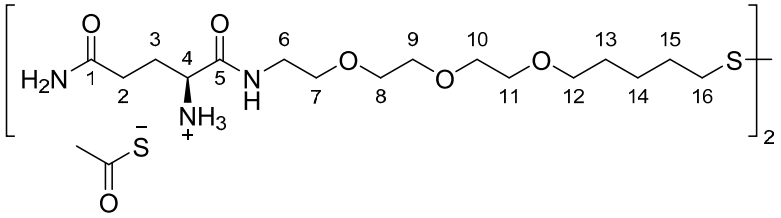
(S)-2-amino-N1-(2-(2-(2-((5-mercaptopentyl)oxy)ethoxy)ethoxy)ethyl) pentanediamide (dimer) **21**

Compound **20** (10 mg, 0,0237 mmol) was dissolved in 0,6 mL of MeOH (previously degased under a stream of argon) and 0,06 mmol (2,5 eq) of sodium methoxide were added. The reaction was stirred at r.t. and after 24 h a sample of the crude was analysed by ^{1}H-NMR to check the complete thioester deprotection. The spectrum clearly indicated the disappereance of the triplet at 2,9 ppm (-CH_2-SAc) and of the singlet at 2,3 ppm (CH_3COS-). A new triplet at 2,65 ppm was present, indicating the formation of a disulfide bond. The reaction was then acidified with HCl 1 N (final pH = 6) and the solvent evaporated. The crude was directly used for the preparation of the nanoparticle. The ^{1}H-NMR analysis suggests that the product is as thioacetate salt.

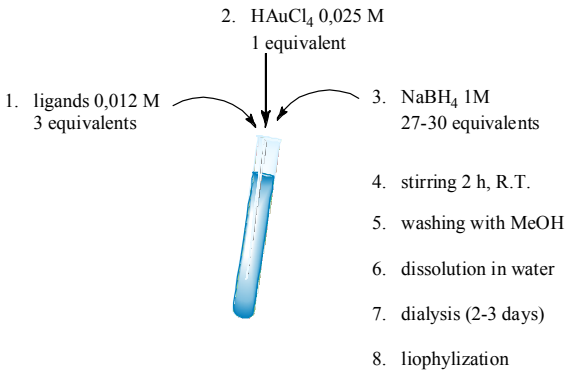
^{1}H NMR (500 MHz, MeOD) δ 4.31 (dd, *J* = 8.6, 5.6 Hz, 1H, H4), 3.68 – 3.58 (m, 8H, C*H*$_2$ PEG), 3.56 (t, *J* = 5.5 Hz, 2H, C*H*$_2$ PEG), 3.49 (t, *J* = 6.5 Hz, 2H, H12a,b), 3.44 – 3.33 (m, 2H, H6a,b), 2.70 (t, *J* = 7.2 Hz, 2H, H16a,b), 2.30 (t, *J* = 7.6 Hz, 2H, H2a,b), 2.11 – 2.03 (m, 1H, H3a), 2.00 (s, 3H, C*H*$_3$COS$^-$), 1.96 – 1.90 (m, 1H, 3b), 1.71 (dt, *J* = 14.7, 7.3 Hz, 2H, H15a,b), 1.65 – 1.56 (m, 2H, H13a,b), 1.54 – 1.43 (m, 2H, H14a,b).

^{13}C NMR (101 MHz, $CDCl_3$) δ 177.65 (C*H*$_3$COS$^-$), 174.01 (C5), 173.33 (C1), 72.13, 71.50, 71.47, 71.24, 71.08, 70.55 (*C*H$_2$ PEG), 54.49 (C4), 40.33 (C6), 39.54 (C16), 32.55 (C2), 30.26 (C13), 30.00 (C15), 29.01 (C3), 26.03 (C14), 22.58 (*C*H$_3$COS$^-$).

structure proposed from spectra



General procedure for gold nanoparticle preparation



The general procedure for the preparation of the gold nanoparticles (Au-NPs), loaded with the desired ligands, is the follow.

The ligand(s) (3 eq) to load on the Au-NPs are dissolved in MeOH, at a final concentration of 0,012 M. If the ligands are more than one, the final concentration must be always 0,012 M. The ratio between the ligands is checked with a ^{1}H NMR spectra, dissolving the ligands in deuterated methanol. To this solution, 1 eq of a solution of $HAuCl_4$ (hydrogen tetrachloroaurate) in water is addes. The concentration of this solution must be 0,025 M. After the addition, the reaction is stirred for few minutes. In this step, Au(I) polymers are formed and the reaction becomes turbid, with a pale yellow colour. A solution of $NaBH_4$ (27-30 eq) in water (1 M) is added in four portion; between the addition, the vessel containing the reaction is vigorously shaken. After the addition of $NaBH_4$, a black precipitate is formed, since gold nanoparticles are

forming. The reaction is left for 2 h at room temperature with 180 r.p.m. shaking. At the end of the 2 hours, the black precipitate (Au-NPs) is washed several times with MeOH, with cycles of vortexing and centrifugation, removing the surnatant and adding fresh MeOH. In this step, the excess of ligands which are still dissolved and were not attached to the NPs are removed. The washing solutions are gathered and combined, in order to verifiy, with ^{1}H NMR analysis, the quantity of the ligands attached to Au-NPs and the ratio (if more ligands are used in the initial MeOH solution. The final washed precipitate is dissolved in water. The water-dissolved nanoparticles are then purified with a dialysis process, for 2 – 3 days. This process allows the removal of boron salts and any other impurity. The final water dispersion of the GNPs is freeze-dried, to obtain a black powder.

References

[1] a)M. Palazzo, S. Gariboldi, L. Zanobbio, S. Selleri, G. F. Dusio, V. Mauro, A. Rossini, A. Balsari, C. Rumio, Journal of Immunology 2008, 181, 3126-3136; b)L. Zanobbio, M. Palazzo, S. Gariboldi, G. F. Dusio, D. Cardani, V. Mauro, F. Marcucci, A. Balsari, C. Rumio, Am. J. Pathol. 2009, 175, 1066-1076.

[2] B. La Ferla, V. Spinosa, G. D'Orazio, M. Palazzo, A. Balsari, A. A. Foppoli, C. Rumio, F. Nicotra, ChemMedChem 2010, 5, 1677-1680.

[3] a)H. Fillmann, N. A. Kretzmann, B. San-Miguel, S. Llesuy, N. Marroni, J. Gonzalez-Gallego, M. J. Tunon, Toxicology 2007, 236, 217-226; b)Y. Huang, N. Li, K. Liboni, J. Neu, Cytokine 2003, 22, 77-83.

[4] a)X. H. Weng, K. W. Beyenbach, A. Quaroni, American Journal of Physiology-Gastrointestinal and Liver Physiology 2005, 288, G705-G717; b)A. Ikari, M. Nakano, Y. Suketa, H. Harada, K. Takagi, Journal of Cellular Physiology 2005, 203, 471-478; c)J. R. Turner, B. K. Rill, S. L. Carlson, D. Carnes, R. Kerner, R. J. Mrsny, J. L. Madara, The American journal of physiology 1997, 273, C1378-1385; d)N. Li, J. Neu, Journal of Nutrition 2009, 139, 710-714; e)N. Li, P. Lewis, D. Samuelson, K. Liboni, J. Neu, American Journal of Physiology-Gastrointestinal and Liver Physiology 2004, 287, G726-G733.

[5] M. Mammen, S. K. Choi, G. M. Whitesides, Angewandte Chemie-International Edition 1998, 37, 2755-2794.

[6] a)J. Giudicelli, M. F. Bertrand, S. Bilski, T. T. Tran, J. C. Poiree, Biochemical Journal 1998, 330, 733-736; b)B. R. Stevens, A. Fernandez, B. Hirayama, E. M. Wright, E. S. Kempner, Proceedings of the National Academy of Sciences of the United States of America 1990, 87, 1456-1460; c)M. Takahashi, P. Malathi, H. Preiser, C. Y. Jung, Journal of Biological Chemistry 1985, 260, 551-556; d)R. J. Turner, E. S. Kempner, Journal of Biological Chemistry 1982, 257, 794-797.

[7] E. Romeo, M. H. Dave, D. Bacic, Z. Ristic, S. M. R. Camargo, J. Loffing, C. A. Wagner, F. Verrey, American Journal of Physiology-Renal Physiology 2006, 290, F376-F383.

[8] a)J. M. de la Fuente, A. G. Barrientos, T. C. Rojas, J. Rojo, J. Canada, A. Fernandez, S. Penades, Angewandte Chemie-International Edition 2001, 40, 2258-+; b)O. Martinez-Avila, K. Hijazi, M. Marradi, C. Clavel, C. Campion, C. Kelly, S. Penades, Chemistry-a European Journal 2009, 15, 9874-9888.

[9] S. Zalipsky, Advanced Drug Delivery Reviews 1995, 16, 157-182.

[10] F. M. Veronese, G. Pasut, Drug Discovery Today 2005, 10, 1451-1458.

[11] A. Dondoni, Angewandte Chemie International Edition 2008, 47, 8995-8997.

[12] F. Cardona, B. La Ferla, Journal of Carbohydrate Chemistry 2008, 27, 203-213.

[13] G. J. McGarvey, C. A. LeClair, B. A. Schmidtmann, Organic Letters 2008, 10, 4727-4730.

[14] A. Wang, J. Hendel, F.-I. Auzanneau, Beilstein Journal of Organic Chemistry 2010, 6.

[15] T. Buskas, E. Söderberg, P. Konradsson, B. Fraser-Reid, The Journal of Organic Chemistry 2000, 65, 958-963.

[16] Á. G. Barrientos, J. M. de la Fuente, T. C. Rojas, A. Fernández, S. Penadés, Chemistry – A European Journal 2003, 9, 1909-1921.

[17] S. A. Svarovsky, Z. Szekely, J. J. Barchi, Tetrahedron: Asymmetry 2005, 16, 587-598.

[18] R. Dettmann, T. Ziegler, Carbohydr. Res. 2011, 346, 2348-2361.

Chapter 5. Iminosugar-decorated calix[4]arenes: chemical tools for the investigation of multivalent inhibition of glycosidases.

Abstract

In this paper report the synthesis of a multivalent calix[4]arenes scaffold decorated with iminosugars, in particular N-derivatives of deoxynojirimycin and of L-idonojirimycin. The decoration of the polyvalent structure was performed exploiting a click chemistry approach, by the copper(I) catalyzed azide alkyne cycloaddition, which provides fast and clean reactions. The generated calix[4]arene will serve as molecular probes for the study of the multivalency effect associated with glycosidase, recently observed.

Introduction

Multivalency represents a fundamental principle used by Nature to achieve strong, but reversible, interactions between biological entities, such as biomacromolecules, which possess low affinity[1]. The presence of multiple copies of a biomacromolecule on the surface of a cell, for example, allows to increase dramatically the binding strength of the ligand with its receptor (or a cluster of them), compared to a single ligand-protein interaction. These events occur in a plethora of biological processes, in particular in cell-cell communications, as in the immune system. The mechanisms and the rationale of the multivalency principle in biology are multiple (see Introduction – Multivalency and Glycobiology): if a ligand interacts with its target protein, it is possible to observe an enhancement in the binding of a second one, due to particular structural reinforcement in the same protein. The multivalency effect can be observed also on the side of the biological effects that can result when a cluster of ligands, disposed close together, binds a particular molecular target or a set of them: in this case it is possible to

obtain an improvement of the biological response. This is partially explained by allosteric effect on the same protein, or by events of dimerization/oligomerization of the target protein, causing an enhancement of the intracellular biological signal. These kind of phenomena are well-documented between carbohydrate ligands and lectins, and in general in cellular recognition events, but less for enzymes, where other mechanisms could be taken into account, as statistical rebinding and an effective increase of local concentration of the ligand loaded in a multivalent structure[2]. Although the multivalency effect of a polyvalently-decorated skeleton is difficult to rationalize in the case of enzymes, especially those bearing a single binding and catalytic site, an increase number of scientific works describe the development of multivalent scaffolds decorated with suitable ligands or inhibitors, that show a high degree of multivalency effect against the respective enzymes[2] [3] [4] [5].

In the last decades, many research group have focused their attention in the preparation of synthetic polyvalent structure, decorated with glycoderivatives or glycoconjugates, as tools to study the multivalency effect exerted by a carbohydrate-based ligand toward their respective binding proteins (lectins) or processing enzymes.

Some of these papers outline the multivalent properties of particular sugar (or glycoderivative)-decorated structures against a specific class of hydrolytic enzymes, called glycosidases. These enzymes cleave the O-glycosidic bond between two sugar units constituting di-, oligo- and polysaccharides, and their inhibitors show a great clinical relevance, for many pathologies, as diabetes or obesity. The developed polyvalent structures, dendrimers[2], fullerenes[3], and cyclodextrins[4], grafted with iminosugars, known glycosidase inhibitors, display a moderate to

remarkable effects onthe enhancement of the inhibition activity, that could not simply be compared to the sum of the single inhibitors.

Starting from these literature data, our work is focused on the development of new structures presenting multiple copies of iminosugars derivatives, to achieve new and complementary information about the multivalency effect exerted by this class of inhibitors against glycosidases. As multivalent scaffold we chose to use a calix[4]arene, in order to prepare a polydecorated structure made of multiple copies of 1-deoxynojirimicin (1-DNJ) derivatives . Calixarenes are structures made of several aromatic rings linked together by methylene bridges, and have been extensively used in the research field of host-guest chemistry and, more recently, as multivalent scaffold for biological relevant molecules and macromolecules. The use of glyco-decorated calixarenes has already been reported to shows enhanced binding affinities towards sugar lectins [6] [7]. We prepared a set of calix[4]arenes functionalized with 1-DNJ derivatives, in which the iminosugar entity is connected to the calixarene core with a spacer bearing a triazole group, as consequence of the copper catalyzed azide-alkyne cycloaddition (CuAAC)[8] approach used for the calixarene functionalization. The obtained polyvalent calixarenes will be tested for their multivalency inhibitory effect against a panel of commercially available glycosidases.

Results and Discussion

Initially, we decided to follow two different approach for the generation of the iminosugar-decorated calix[4]arenes, which were chosen among the family of this scaffold for their easier synthesis and their straightforward functionalization. Starting from a panel of already synthesized calix[4]arenes [6, 8-9], we planned the synthesis of suitably

functionalized 1-DNJ derivatives in order to attach them to the skeleton by the thiol-ene coupling (TEC) approach[6], and the CuAAC method (Figure 49).

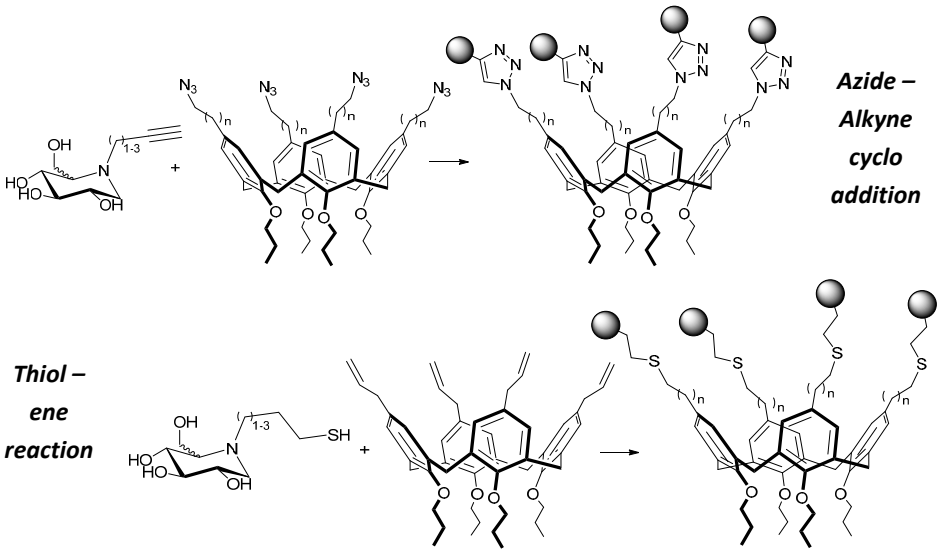
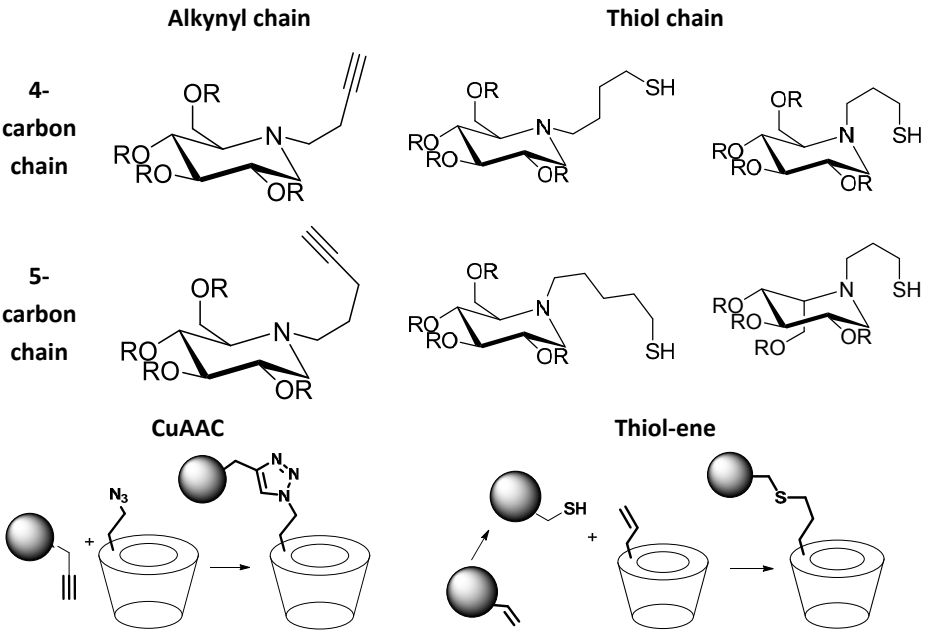


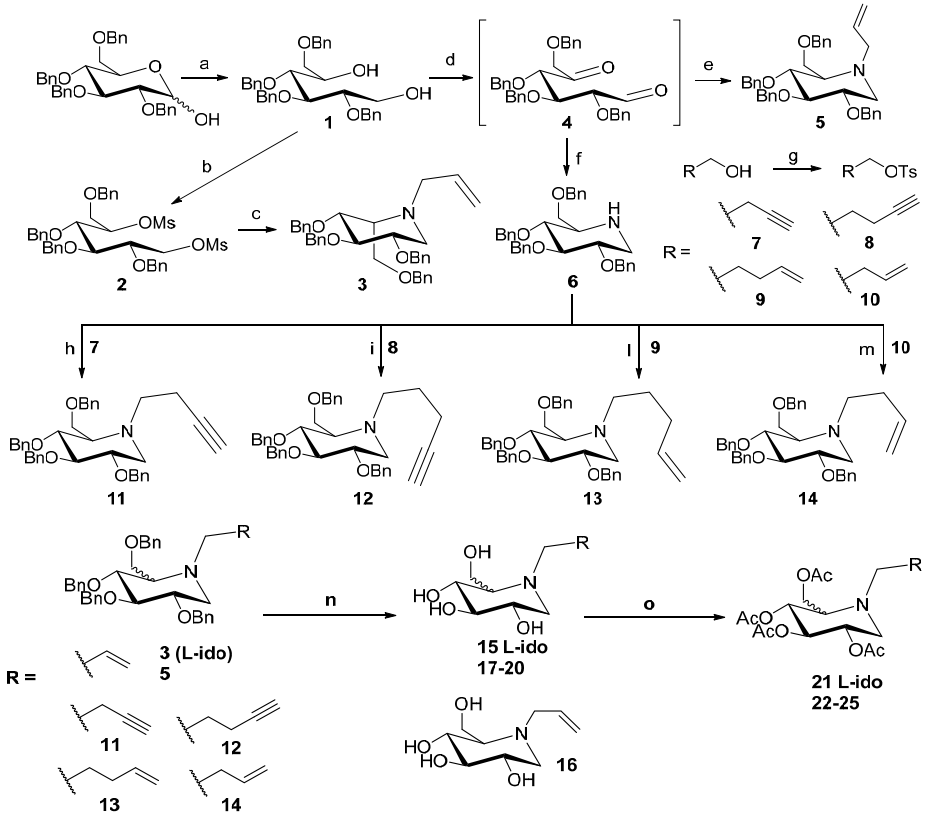
Figure 49 - The two "click reaction" approaches for the chemical functionalization of the calixarene scaffold.

We designed a pool of iminosugars, derivatives of 1-DNJ and L-ido-1-DNJ, bearing a N-alkyl chain with a terminal triple bond, to exploit in the CuAAC strategy, or with a terminal thiol group, to be used for a thiol-ene coupling reaction. These derivatives will then be anchored to n the calix[4]arene bearing respectively an azido or allylic appendages (Table 3).

Table 3 – Set of designed and prepared iminosugars derivatives for the calix[4]arenes decoration. R = H or Protecting groups (Ac or Bn).



As previously said, the thiol derivative was to be prepared from the alkene precursor, so we decided to prepare a small library containing both iminosugars functionalized with a 4- and with a 5-carbon chain, bearing a double or triple bond exploiting a common synthetic strategy, and the N-allyl derivatives were synthesized, for the generation of the 3 carbon chain derivatives. The overall synthetic plan used is presented below:



Scheme 13 - Synthesis of the library of iminosugars. Reagents and conditions: a) $NaBH_4$, CH_2Cl_2, EtOH, r.t., 24 h[10]; b) See Ref.[10] [11]; c) Allylamine (neat), reflux, 36 h; d) i. $(COCl)_2$, DMSO, CH_2Cl_2, -78°C, 2 h; ii. Et_3N, CH_2Cl_2, -78°C to r.t., 2 h; e) allylamine, molecular sieves 3 Å, $NaCNBH_3$, MeOH, 0°C to r.t., 12 h; f) $NH_4^+HCOO^-$, molecular sieves 3 Å, $NaCNBH_3$, MeOH, 0°C to r.t.., 12 h; g) TsCl, DMAP, CH_2Cl_2, Et_3N, r.t., 24 h; h,I,l,m) R-OTs, CH_3CN dry, K_2CO_3, reflux, 48 h; n) BCl_3 1M in CH_2Cl_2, CH_2Cl_2, -60°C to 0°C, 4 h; o) Ac_2O, Py, r.t., 24 h.

Commercially available 2,3,4,6-tetra-*O*-benzyl-D-glucopyranose was reduced with $NaBH_4$ affording glucitol **1**, which was used as starting material for the preparation of all compounds. A double mesylation (**2**) and the reaction with allylamine led to the obtainment of N-allyl-iminosugar with an inversion of configuration at C-5 position, thus giving

the L-ido configuration. For the synthesis of N-allyl-1-deoxynojirimicyn (**5**) glucitol **1** was oxidized through a Swern reaction into the instable intermediate **4**[10] and subjected to a double reductive amination using allylamine. The same protocol was repeated with ammonium formate affording the tetra-*O*-benzyl-1-DNJ (**6**), that was used for the preparation of derivatives **11**-**14,** obtained by reaction with tosylates (**7**-**10**) .The successive debenzylation with BCl_3 led to the obtainment of the deprotected iminosugars (**15**-**20**), which were acetylated without further purification steps, to generate the final compounds **21**-**25**. The L-ido derivative **16** was not subjected to acetylation and used after a purification step.

Alongside the preparation of the library, our work focused on the synthesis of the calix[4]arene scaffold. We prepared various calix[4]arenes which differs for the type and the position of the inserted functionalizations (Figure 50). The first ones (**27**-**30**) were prepared according to published procedures[6] and bear from four to eight allyl groups, on each side of the structure or only on one. The presence of the double bonds permit the exploitation of the TEC strategy for the decoration of the calixarene. The second ones, already known and synthesized structures[6, 8], present four azido functionalities which can be exploited for the conjugation of compounds with a terminal alkynyl group by the CuAAC method.

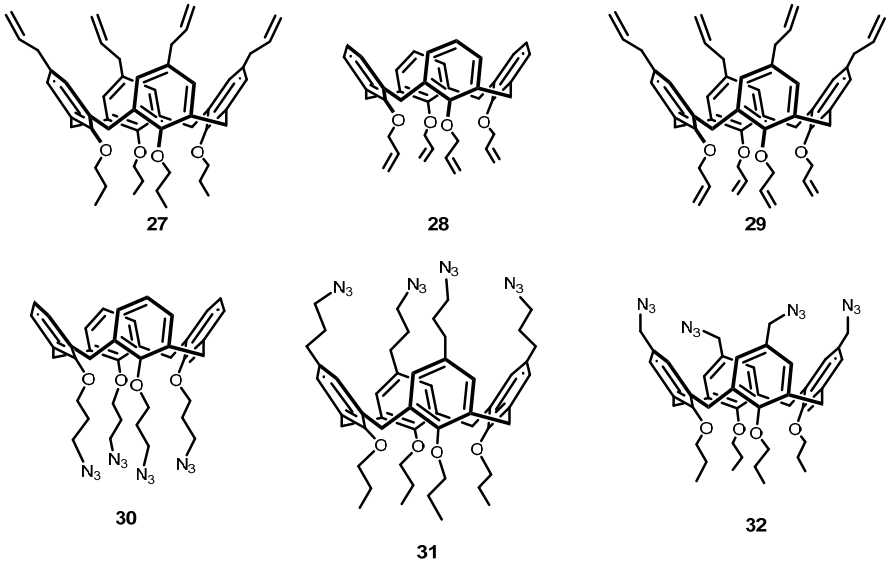
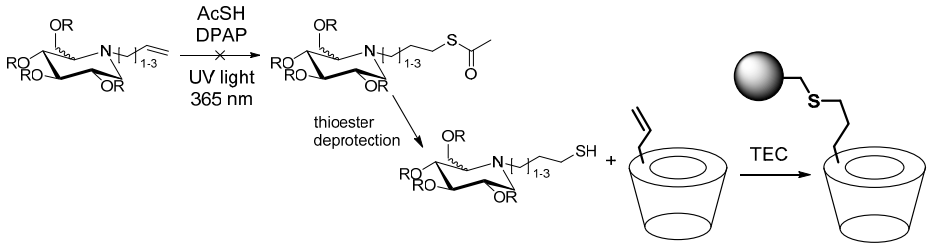


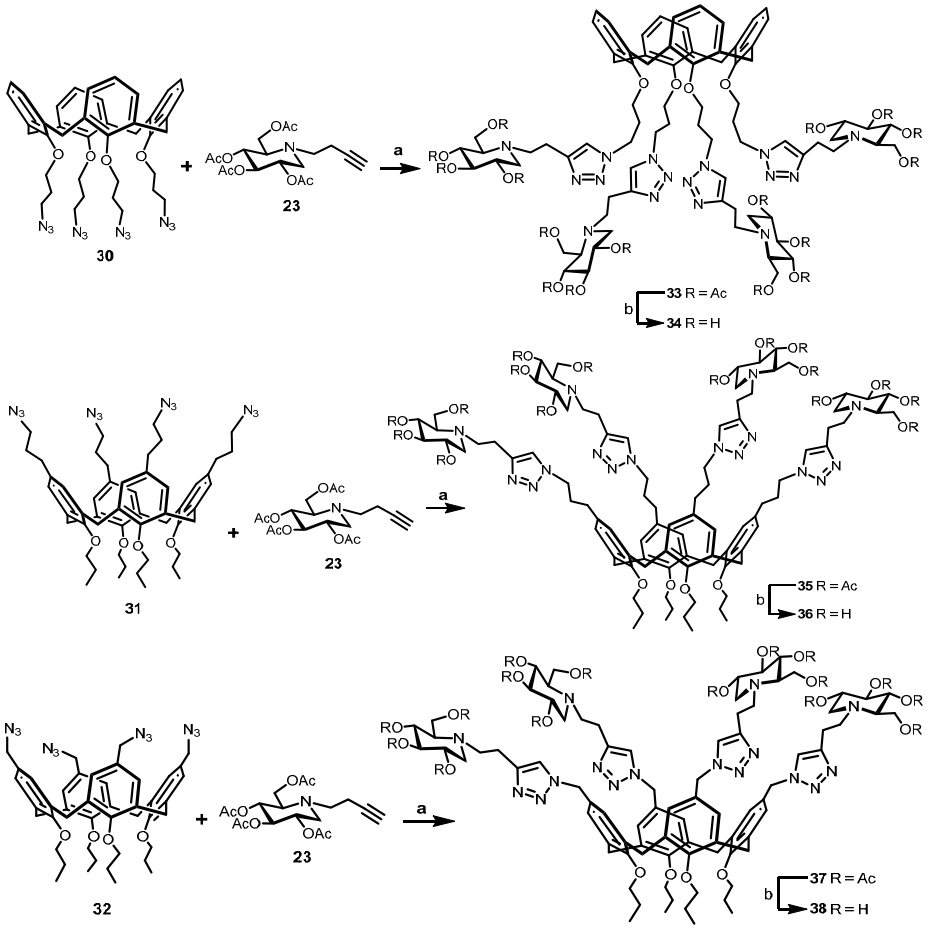
Figure 50 - Structures of the different calix[4]arenes synthesized.

Initially, our efforts focused on the transformation of the alkenyl appendage on the iminosugars **16** and **21**-**25** into a thioester functionality by a thiol-ene reaction with thiolacetic acid, in presence of the radical photoinitiator DPAP upon UV-light exposure (365 nm), but unfortunately no satisfying results were obtained with this approach (Scheme 14). The reaction conditions were identical to those reported in Fiore, et al. [6] and Fiore, et al. [12].



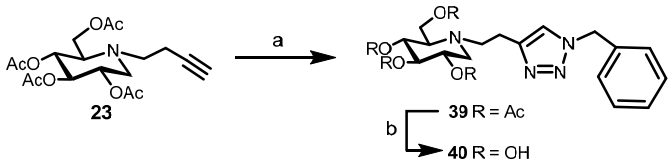
Scheme 14 – Thiol-ene coupling strategy for the decoration of calix[4]arene with N-alkenyl iminosugars.

We decide to move our attention for the development of the desired calix[4]arenes using the Cu(I)-azide alkyne cycloaddition (CuAAC), using the scaffolds functionalized with azido groups and iminosugars substituted with an alkynyl chain. In this case we exploited the reaction conditions and procedures of CuAAC already drescribed in Dondoni and Marra [8], Marra, et al. [9]. We were able to obtain three different iminosugar-decorated calix[4]arenes (**34**, **36** and **38**), performing the cycloaddition between the N-alkynyl iminosugar **23** with the three azido-scaffolds **30**, **31** and **32** (Scheme 15). Calix[4]arene **40** contains a propylene azido chain on the lower rim, and the performed cycloaddition with **23** generated the decorated scaffold **33**, containing four 1-DNJ units linked to the tetra-aromatic core by a triazole-bearing tether, which is directly connected to the calixarene by a propilenoxy chain and to the iminosugar by an ethylene bridge. The complete deacetylation with ammonia in methanol afforded the final compound **34**. Starting calixarene **31** bears propyl appendages on the phenolic oxygen atoms located on the lower rim, while propylene azido function on the upper rim, which were transformed into triazole entities by the CuAAC with **23**. The last calixarene was generated by the starting compound **32**, in which the azido groups are separated from the aromatic core of the scaffold by a single methylene bridge. Finally calix[4]arenes **37** and **38** show a similarity to **35** and **36**, but the spacer between the iminosugar and the aromatic ring is shorter.



Scheme 15 – Preparation of the three iminosugar-decorated calix[4]arenes by the CuAAC approach. Reagents and conditions: a) Cu(I), iPr_2EtN, Toluene, 80°C, 2 h; b) NH_3, MeOH, r.t., 18 h.

Furthermore, a cycloaddition reaction between the same iminosugar and benzyl azide afforded compound **40** (Scheme 16), which represents the ligand constituting the calixarene tetramer. The monomer will be used as a control during the enzymatic assays to determine the multivalency effect of the calixarene polyvalent structures.



Scheme 16 – Synthesis of 40. Reagents and conditions: a) Benzyl azide, Cu(I), Toluene, 80°C, 20 h; b) NH_3, MeOH, r.t., 18 h.

The obtained final calixarenes **34**, **36** and **38**, and the monovalent ligand **40** will be tested for their inhibitory activity against a panel of commercially available glycosidases, exploting a simple and fast colorimetric assay. Suitable p- or o-nitrophenol glycosides will be used as chromogenic substrates of several glycosidases, which catalytic activity allows the cleavage of the *O*-glycosidic bond, thus liberating the respective p- or o-nitrophenolate. The inhibition activity exerted by the tested compound will cause a decrease of absorbance of the reaction solution, due to less amount of cleaved chromogenic susbtrate. Using different concentration of tested compounds and performing the reaction with several substrate concentration it will be possible to determine the inhibition constant for each compound, thus allowing a comparison of the multivalency efficacy among the prepared calixarenes and against the monovalent ligand.

Conclusions

Recent encouraging results on glycosidases inhibition using polyvalent structure decorated with iminosugars have prompted us in the development of a new class of structures bearing multiple copies of of 1-deoxy-L-idonojirimycin and 1-DNJ derivatives, for which a remarkable inhibitory activity against several glycosidases is reported. Among the big number of available scaffolds with multivalent functionalities, we chose the calixarene skeleton for several reasons. Firstly, calixarenes possess a precise shape and geometry which can be exploited for the multiple functionalization with many bioorganic compounds, as carbohydrates or peptides. The calixarene scaffold is relatively simple, made of phenolic rings connected together in a conical structure by methylene bridges: the acidic oxygen can be easily alkylated, allowing the insertion of many appendages bearing different chemical functionalities. It permits not only a chemical modulation of the calixarene reactivity, but even the exploitation of chemoselective methods for the decoration of such scaffolds. In our case, we selected a calix[4]arene skeleton bearing allylic or azidoalkyl appendage, thus allowing the use of two chemical approach: the thiol-ene coupling and the copper(I)-catalyzed azide alkyne cycloaddition. While the first strategy did not give the expected results, the CuAAC reactions allowed the generation of three calix[4]arenes with different geometry and shape. The same reaction performed using benzyl azide and iminosugar **23** afforded a monovalent ligand, compound **40**, which can be used as monovalent control during the enzymatic assays for the inhibitory activity of the synthesized compounds. The execution of the enzymatic test will allow to understand if the obtained calix[4]arenes show a multivalent inhibitory effect against a pool of commercial glycosidases. The results can be used as base for the generation of a second library of structures both with a

better multivalency or, moreover, for the creation of hybrid structures: the typical geometries and shapes of the calixarenes, made generally of two sides or rims, can be exploited for a a dual ligand type functionalization, such as iminosugars and fluorescent probes, to achieve also trafficking studies performing *in vivo* experiments. The synthesis of these types of calix[4]arenes and the execution of these experiments will allow to gather new and more detailed information about the multivalent effect recently observed on glycosidases.

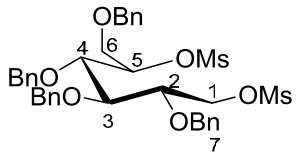
Experimental Section

General remarks

All commercial chemicals were purchased from Sigma-Aldrich and used without further purification. All moisture-sensitive reactions were performed under a nitrogen atmosphere using oven-dried glassware. Anhydrous solvents were dried over standard drying agents and freshly distilled prior to use or dried with molecular sieves for at least 24h prior to use. Thin layer chromatography (TLC) was performed on silica gel 60 F_{254} plates (Merck) with detection under UV light when possible, or by charring with a solution of $(NH_4)_6Mo_7O_{24}$ (21g), $Ce(SO_4)_2$ (1g), concentrated H_2SO_4 (31 mL) in water (500 mL) or with an ethanol solution of ninhydrin or with Dragendorff' spray reagent. Flash-column chromatography was performed on silica gel 230–400 mesh (Merck). ^1H and ^{13}C NMR spectra were recorded at 25°C, unless otherwise stated, with a Varian Mercury 400-MHz instrument. Chemical shift assignments, reported in parts per million, were referenced to the corresponding solvent peaks. Mass spectra were recorded on a QTRAP system with ESI source, while HRMS were registered on a QSTAR elite system with a nanospray ion source.

Synthesis of iminosugars

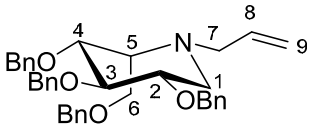
Synthesis of compound **1**, intermediate **4** and tetra-O-benzyl-1-DNJ **6** were already described in literature. All characterizations are consistent to those reported in literature[10] [11].



2,3,4,6-tetra-O-benzylglucitol dimesylate **2**

The synthesis if compound 2 was carried out according to a procedure described in Jiang, et al. [13] and Rao and Perlin [14]. All NMR data obtained were consistent to those reported in literature.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.40 – 7.16 (m, 20H, C*H* Ar), 5.03 (dt, *J* = 7.7, 2.7 Hz, 1H, H5), 4.76 (d, *J* = 11.1 Hz, 1H, CH_2Ph), 4.67 (d, *J* = 11.6 Hz, 1H, CH_2Ph), 4.62 (d, *J* = 11.1 Hz, 1H, CH_2Ph), 4.60 – 4.54 (m, 3H, CH_2Ph), 4.50 – 4.44 (m, 3H, CH_2Ph, H1a), 4.25 (dd, *J* = 11.2, 7.2 Hz, 1H, H1b), 4.05 (dd, *J* = 4.8, 2.8 Hz, 1H, H4), 3.95 – 3.89 (m, 1H, H2), 3.86 (dd, *J* = 11.1, 2.6 Hz, 1H, H6a), 3.77 (dd, *J* = 11.1, 7.9 Hz, 1H, H6b), 3.71 (t, *J* = 5.2 Hz, 1H, H3), 2.95 (s, 3H, OSO_2CH_3), 2.83 (s, 3H, OSO_2CH_3).



2,3,4,6-tetra-O-benzyl-L-ido-N-allyl-1-deoxynojirimycin **3**

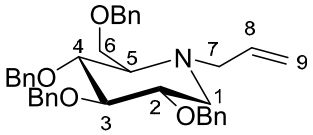
Dimesylate **2** (2042 mg, 2,926 mmol) was dissolved in allylamine (15 mL) under argon atmosphere. The reaction was heated and stirred at reflux. After 36 h, TLC (PE/AcOEt 8:2) showed no more starting compound, and the solvent was evaporated. The residue was purified by FC (eluent

PE/AcOEt 9:1 to 8:2), affording compound **3** (1,238 g, 2,196 mmol, 75% yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.44 – 7.28 (m, 20H, C*H* Ar), 5.91 – 5.74 (m, 1H, H8), 5.18 (dd, *J* = 23.5, 5.9 Hz, 2H, H9a,b), 4.90 (d, *J* = 10.9 Hz, 1H, CH_2Ph), 4.85 (d, *J* = 10.9 Hz, 1H, CH_2Ph), 4.75 (d, *J* = 11.5 Hz, 1H, CH_2Ph), 4.72 – 4.64 (m, 3H, CH_2Ph), 4.62 – 4.53 (m, 2H, CH_2Ph), 3.90 (dd, J = 10.2, 6.8 Hz, 1H, H6a), 3.81 – 3.69 (m, 2H, H6b,H4), 3.67 – 3.54 (m, 2H, H3,H2), 3.51 – 3.39 (m, 2H, H5, H7a), 3.24 (dd, J = 14.0, 6.9 Hz, 1H, H7b), 2.99 (dd, *J* = 11.8, 4.8 Hz, 1H, H1a), 2.59 (dd, *J* = 11.4, 10.2 Hz, 1H, H1b).

^{13}C NMR (101 MHz, $CDCl_3$) δ 139.18, 138.69, 138.64, 138.55 (Cq Ar), 136.30 (C8), 128.46, 128.44, 128.40, 128.36, 128.07, 127.86, 127.85, 127.69, 127.62, 127.59, 127.52 (C Ar), 117.28 (C9), 83.01 (C3), 80.17 (C4), 78.85 (C2), 75.49, 73.37, 73.04, 72.72 (CH_2Ph), 64.61 (C6), 60.09 (C5), 58.09 (C7), 49.05 (C1), 29.82.

$C_{37}H_{41}NO_4$;; calcd. mass 563,74; MS-ESI: m/z 564,5 $[M+H]^+$.



2,3,4,6-tetra-O-benzyl-N-allyl-1-deoxynojirimycin **5**

The synthesis of **5** was carried out according to the procedures described in Matos, et al. [10] and Wennekes, et al. [11], with some modifications. DMSO (1,3 mL, 18,52 mmol, 10 eq.) was added at -78°C to a round-bottom flask containing 6 mL of dry CH_2Cl_2, under argon atmosphere. A solution of trifluoroacetic anhydride (1,3 mL, 9,26 mmol, 5 eq.) in dry CH_2Cl_2 was successively added dropwise to the flask and the reaction was stirred for 30 min at -78°C. Then, a CH_2Cl_2 solution (8 mL) of tetra-*O*-benzyl-glucitol **1** (1 g, 1,852 mmol, 1 eq.) in dry CH_2Cl_2 was added dropwise; the reaction was stirred at -78°C. After 2 h, a solution of Et_3N

(1,65 mL) in CH_2Cl_2 (6 mL) was added to the reaction that was successively warmed to r.t. and concentrated. The residue, containing the intermediate **4**, was diluted in MeOH (12 mL) under argon atmosphere; allylamine (0,29 mL, 3,9 mmol, 2,1 eq. respect to starting material **1**), molecular sieves 3 Å and $NaCNBH_3$ (4,63 mmol, 2,5 eq.) were added. The reaction was stirred at r.t. and followed by TLC (Toluene/AcOEt 9:1). After 24 h, the reaction was filtered on a celite pad to remove molecular sieves; the filtrate was evaporated, the residue was dissolved in AcOEt and washed three times with $NaHCO_3$ satd. solution. The organic phase was filtered, dried over Na_2SO_4 and concentrated. The crude was subjected to flash chromatography (eluent Toluene/AcOEt 9:1) affording 232 mg (0,416 mmol) of **5**, with a yield of 22%. The L-ido epimer was also isolated (30 mg, 0,053 mmol, 3% yield).

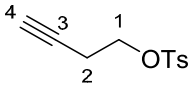
^{1}H NMR (400 MHz, CDCl3) δ 7.42 – 7.04 (m, 20H), 5.95 – 5.75 (m, 1H), 5.11 (dd, J = 25.3, 13.6 Hz, 2H), 4.95 (d, J = 11.0 Hz, 1H), 4.87 (d, J = 10.8 Hz, 1H), 4.81 (d, J = 11.0 Hz, 1H), 4.72 – 4.60 (m, 2H), 4.48 (s, 2H), 4.38 (d, J = 10.8 Hz, 1H), 3.72 – 3.53 (m, 4H), 3.45 (t, J = 9.1 Hz, 1H), 3.38 (dd, J = 14.3, 5.5 Hz, 1H), 3.22 – 3.06 (m, 2H), 2.29 (d, J = 9.6 Hz, 1H), 2.19 (t, J = 10.9 Hz, 1H).

$C_{37}H_{41}NO_4$; calcd. mass 563,74; MS-ESI: m/z 564,6 $[M+H]^+$.

*General procedure for the synthesis of compounds **7-10***

Alkenyl or alkynyl alcohol (1 eq) and DMAP (cat.) were dissolved in CH_2Cl_2 under argon atmosphere. TsCl (1,1 eq) and Et_3N (1,1 eq) were added to the solution. The reaction was stirred at r.t. and followed by TLC (PE/AcOEt 9:1). After 24 h, water was added to the reaction, the phases was separated and the aqueous layer was extracted with CH_2Cl_2 (3x). The combined organic phases were combined, dried over sodium

sulphate, filtrated and concentrated. The crude was purified by FC (eluent PE/AcOEt 9,5:0,5).



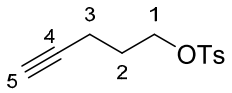
but-3-yn-1-yl-p-toluensulfonate **7**

300 mg of but-3-in-1-ol (4,28 mmol) were converted into the tosylate **7** (3,81 mmol) with a yield of 89%.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.79 (d, *J* = 8.3 Hz, 2H, H Ar), 7.35 (d, *J* = 8.5 Hz, 2H, H Ar), 4.09 (t, *J* = 7.0 Hz, 2H, H1a,b), 2.55 (td, *J* = 7.0, 2.7 Hz, 2H, H2a,b), 2.45 (s, 3H, C*H*$_3$Ph), 1.97 (t, *J* = 2.7 Hz, 1H, H4).

^{13}C NMR (101 MHz, $CDCl_3$) δ 145.15, 132.79 (Cq Ar), 130.01, 128.06 (C Ar), 78.48 (C3), 70.88 (C4), 67.55 (C1), 21.77 (*C*H_3Ph), 19.52 (C2).

$C_{11}H_{12}O_3S$; calcd. mass 224,27; MS-ESI: m/z 225,3 $[M+H]^+$.



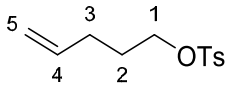
pent-4-yn-1-yl-p-toluensulphonate **8**

Pent-4-in-1-ol (300 mg, 3,57 mmol) were converted into compound **8** (3,3 mmol) with a yield of 93%.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.78 (d, *J* = 8.3 Hz, 2H, C*H* Ar), 7.34 (d, *J* = 8.0 Hz, 2H, C*H* Ar), 4.13 (t, *J* = 6.1 Hz, 2H, H1a,b), 2.44 (s, 3H, C*H*$_3$Ph), 2.25 (td, *J* = 6.9, 2.6 Hz, 2H, H3a,b), 1.88 (t, *J* = 2.6 Hz, 1H, H5), 1.87 – 1.80 (m, 2H, H2a,b).

^{13}C NMR (101 MHz, $CDCl_3$) δ 144.93, 132.91 (Cq Ar), 129.96, 128.01 (C Ar), 82.19 (C4), 69.55 (C5), 68.83 (C1), 27.76 (C3), 21.74 (*C*H_3Ph), 14.77 (C2).

$C_{12}H_{14}O_3S$; calcd. mass 238,30; MS-ESI: m/z 239,3 $[M+H]^+$.



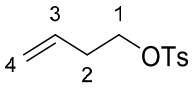
pent-4-en-1-yl-p-toluensulfonate **9**

Pent-4-en-1-ol (300 mg, 3,48 mmol) was transformed into the respective tosylate **9** with a yield of 93% (3,23 mmol).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.77 (d, J = 8.3 Hz, 2H, C*H* Ar), 7.34 (d, J = 8.0 Hz, 2H, C*H* Ar), 5.68 (ddt, J = 16.4, 9.6, 6.7 Hz, 1H, H4), 5.03 – 4.85 (m, 2H, H5a,b), 4.02 (t, J = 6.4 Hz, 2H, H1a,b), 2.44 (s, 3H, CH_3Ph), 2.14 – 1.94 (m, 2H, H3a,b), 1.81 – 1.60 (m, 2H, H2a,b).

^{13}C NMR (101 MHz, $CDCl_3$) δ 144.84 (Cq Ar), 136.69 (C4), 133.11 (Cq Ar), 129.93 (C Ar), 127.96 (C Ar), 115.94 (C5), 69.91 (C1), 29.46 (C3), 28.03 (C2), 21.74 (*C*H_3Ph).

$C_{12}H_{16}O_3S$; calcd. mass 240,32; MS-ESI: m/z 241,2 $[M+H]^+$.



but-3-en-1-yl-p-toluensulfonate **10**

But-3-en-1-ol (202 mg, 2,8 mmol) was transformed into the respective tosylate **10** with a yield of 93% (2,52 mmol).

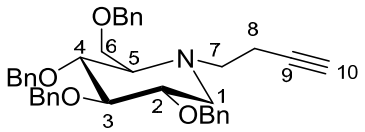
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.78 (d, *J* = 8.3 Hz, 2H, H Ar), 7.34 (d, *J* = 8.0 Hz, 2H, H Ar), 5.66 (ddt, *J* = 17.1, 10.4, 6.7 Hz, 1H, H3), 5.07 (ddd, *J* = 10.3, 2.7, 1.4 Hz, 2H, H4), 4.05 (t, *J* = 6.7 Hz, 2H, H1), 2.45 (s, 3H, -CH_3Ph), 2.43 – 2.34 (m, 2H, H2).

^{13}C NMR (101 MHz, $CDCl_3$) δ 144.88, 133.07 (Cq Ar), 132.48 (C3), 129.92, 127.97 (C Ar), 118.30 (C4), 69.52 (C1), 33.21 (C2), 21.74 (*C*H_3Ph).

$C_{11}H_{14}O_3S$; calcd. mass 226,29; MS-ESI: m/z 227,3 $[M+H]^+$.

General procedure for the synthesis of N-substituted 1-DNJ ***11-14***

2,3,4,6-tetra-O-benzyl-1-deoxynojirimycin (1 DNJ-OBn) (1 eq) and tosylate **7**-**10** (2 eq) were dissolved in dry CH_3CN under argon atmosphere. K_2CO_3 (3 eq) was added to the solution and the reaction was heated and stirred at reflux for 24 h. The reaction was followed by TLC (PE/AcOEt 8:2). At the end of the reaction, the reaction was concentrated under vacuum; the residue was resuspended with water and chloroform and stirred at r.t. for 10 min. The two phases were separated and the aqueous layer washed three time with chloroform. The combined organic phases were dried, filtrated and concentrated. The product was purified from the crude by FC (PE/AcOEt 9:1) to afford compounds **11**-**14**.

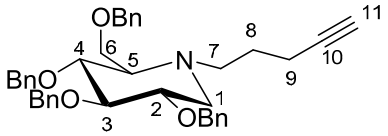


2,3,4,6-tetra-O-benzyl-N-(but-3-yn-1-yl)-1-deoxynojirimycin **11**

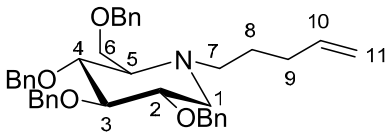
From 569 mg (1,09 mmol) of 1-DNJ-OBn, 572 mg (0,99 mmol) of compound **11** were obtained, with a yield of 91%.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.47 – 7.19 (m, 18H, C*H* Ar), 7.16 – 7.06 (m, 2H, C*H* Ar), 4.96 (d, *J* = 11.0 Hz, 1H, CH_2Ph), 4.87 (d, *J* = 10.8 Hz, 1H, CH_2Ph), 4.81 (d, *J* = 11.0 Hz, 1H, CH_2Ph), 4.70 (d, *J* = 11.6 Hz, 1H, CH_2Ph), 4.65 (d, *J* = 11.6 Hz, 1H, CH_2Ph), 4.54 (d, *J* = 12.2 Hz, 1H, CH_2Ph), 4.46 (d, *J* = 12.2 Hz, 1H, CH_2Ph), 4.37 (d, *J* = 10.7 Hz, 1H, CH_2Ph), 3.74 – 3.58 (m, 3H, H2, H6a,b), 3.55 (t, *J* = 9.2 Hz, 1H, H4), 3.46 (t, *J* = 9.0 Hz, 1H, H3), 3.08 (dd, *J* = 11.1, 4.8 Hz, 1H, H1a), 3.05 – 2.91 (m, 2H, H1b, H7a), 2.49 – 2.35 (m, 2H, H7b, H5), 2.34 – 2.24 (m, 2H, H8a,b), 1.97 (t, *J* = 2.6 Hz, 1H, H10).

^{13}C NMR (101 MHz, $CDCl_3$) δ 139.00, 138.50, 138.49, 137.75 (Cq Ar), 128.65 – 127.55 (C Ar x 20), 87.31 (C3), 82.72 (C6), 78.52 (C4), 78.41 (C2), 75.52 (*C*H_2Ph), 75.35 (*C*H_2Ph), 73.62 (*C*H_2Ph), 72.91 (*C*H_2Ph), 69.63 (C9), 65.68 (C10), 62.98 (C5), 54.51 (C7), 51.19 (C1), 13.94 (C8).
$C_{38}H_{41}NO_4$; calcd. mass 575,75; MS-ESI: m/z 576,8 $[M+H]^+$.



2,3,4,6-tetra-O-benzyl-N-(pent-4-yn-1-yl)-1-deoxynojirimycin **12**
From 590 mg (1,128 mmol) of 1-DNJ-OBn, 580 mg (0,97 mmol) of compound **12** were obtained, with a yield of 86%.
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.43 – 7.05 (m, 20H, C*H* Ar), 4.97 (d, *J* = 11.0 Hz, 1H, CH_2Ph), 4.88 (d, *J* = 10.8 Hz, 1H, CH_2Ph), 4.83 (d, *J* = 11.1 Hz, 1H, CH_2Ph), 4.74 – 4.62 (m, 2H, CH_2Ph), 4.54 – 4.45 (m, 2H, CH_2Ph), 4.42 (d, *J* = 10.8 Hz, 1H, CH_2Ph), 3.73 – 3.64 (m, 2H, H2, H6a), 3.64 – 3.55 (m, 2H, H4, H6b), 3.47 (t, *J* = 9.0 Hz, 1H, H3), 3.09 (dd, *J* = 11.2, 4.8 Hz, 1H, H1a), 2.89 – 2.76 (m, 1H, H7a), 2.76 – 2.65 (m, 1H, H7b), 2.38 – 2.30 (m, 1H, H5), 2.25 (t, *J* = 10.8 Hz, 1H, H1b), 2.20 – 2.02 (m, 2H, H9a,b), 1.93 (t, *J* = 2.6 Hz, 1H, H11), 1.73 – 1.54 (m, 2H, H8a,b).
^{13}C NMR (101 MHz, $CDCl_3$) δ 139.07, 138.57, 137.88 (Cq Ar x 4), 128.53, 128.49, 128.43, 127.98, 127.94, 127.76, 127.65, 127.55 (C Ar x 20), 87.40 (C3), 84.06 (C10), 78.66 (C4), 78.53 (C2), 75.44, 75.32, 73.54, 72.87 (*C*H_2Ph), 68.69 (C11), 65.63 (C6), 63.87 (C5), 54.66 (C1), 51.27 (C7), 23.00 (C8), 16.40 (C9).
$C_{39}H_{43}NO_4$; calcd. mass 589,78; MS-ESI: m/z 590,9 $[M+H]^+$.



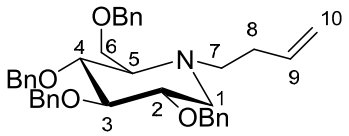
2,3,4,6-tetra-O-benzyl-N-(pent-4-en-1-yl)-1-deoxynojirimycin **13**

Compound **13** was synthesized according to the general procedure. 597 mg (1,141 mmol) of 1-DNJ-OBn were converted into 586 mg (0,99 mmol) of compound **13**, with a yield of 87%.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.44 – 7.04 (m, 20H, C*H* Ar), 5.76 (ddt, *J* = 16.9, 10.2, 6.6 Hz, 1H, H10), 5.04 – 4.92 (m, 3H, H11a,b, CH_2Ph), 4.89 (d, *J* = 10.8 Hz, 1H, CH_2Ph), 4.83 (d, *J* = 11.1 Hz, 1H, CH_2Ph), 4.74 – 4.63 (m, 2H, CH_2Ph), 4.55 – 4.45 (m, 2H, CH_2Ph), 4.43 (d, *J* = 10.8 Hz, 1H, CH_2Ph), 3.72 – 3.64 (m, 2H, H2, H6a), 3.61 (t, *J* = 9.3 Hz, 1H, H4), 3.55 (dd, *J* = 10.4, 2.1 Hz, 1H, H6b), 3.47 (t, *J* = 9.1 Hz, 1H, H3), 3.10 (dd, *J* = 11.1, 4.8 Hz, 1H, H1a), 2.77 – 2.65 (m, 1H, H7a), 2.65 – 2.55 (m, 1H, H7b), 2.32 (d, *J* = 9.5 Hz, 1H, H5), 2.26 (t, *J* = 10.8 Hz, 1H, H1b), 2.04 – 1.85 (m, 2H, H9a,b), 1.61 – 1.39 (m, 2H, H8a,b).

^{13}C NMR (101 MHz, $CDCl_3$) δ 139.11, 138.63, 138.62 (Cq Ar), 138.32 (C10), 137.86 Cq Ar), 128.60 - 127.55 (C Ar x 20), 114.92 (C11), 87.47 (C3), 78.67 (C4), 78.63 (C2), 75.45 (*C*H_2Ph), 75.32 (*C*H_2Ph), 73.57 (*C*H_2Ph), 72.86 (*C*H_2Ph), 65.31 (C6), 63.70 (C5), 54.60 (C7), 51.90 (C1), 31.67 (C9), 22.87 (C8).

$C_{39}H_{45}NO_4$; calcd. mass 591,79; MS-ESI: m/z 592,8 $[M+H]^+$.



2,3,4,6-tetra-O-benzyl-N-(but-3-en-1-yl)-1-deoxynojirimycin **14**

The general procedure applied on 667 mg (1,275 mmol) of tetra-*O*-benzyl-1-DNJ afforded 653 mg (1,13 mmol, 89% yield) of compound **14**.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.58 – 6.92 (m, 20H, H Ar), 5.68 (td, *J* = 17.2, 6.9 Hz, 1H, H9), 5.03 – 4.92 (m, 3H, H10a,b, C*H*$_2$Ph (x1)), 4.88 (d, *J* = 10.8 Hz, 1H, C*H*$_2$Ph), 4.82 (d, *J* = 11.1 Hz, 1H, C*H*$_2$Ph), 4.73 – 4.62 (m, 2H, C*H*$_2$Ph), 4.49 (s, 2H, C*H*$_2$Ph), 4.40 (d, *J* = 10.8 Hz, 1H, C*H*$_2$Ph), 3.71 – 3.63 (m, 2H, H2, H3), 3.63 – 3.52 (m, 2H, H6a, H6b), 3.46 (t, *J* = 9.1 Hz, 1H, H4), 3.10 (dd, *J* = 11.0, 4.6 Hz, 1H, H1a), 2.79 – 2.69 (m, 2H, H7a,b), 2.39 – 2.25 (m, 2H, H1b, H5), 2.25 – 2.03 (m, 2H, H8a,b).

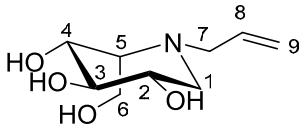
^{13}C NMR (101 MHz, $CDCl_3$) δ 139.08, 138.58, 137.80 (Cq Ar x 4), 136.32 (C9), 129.08 – 127.17 (C Ar x 20), 116.08 (C 10), 87.45 (C3), 78.62 (C4), 78.57 (C2), 75.48 (*C*H_2Ph), 75.35 (*C*H_2Ph), 73.59 (*C*H_2Ph), 72.88 (*C*H_2Ph), 65.21 (C6), 63.35 (C5), 54.55 (C7), 51.83 (C1), 28.19 (C8).

$C_{38}H_{43}NO_4$; calcd. mass 577,76; MS-ESI: m/z 578,8 $[M+H]^+$.

*General procedure for benzyl deprotection – synthesis of compounds **15-20***

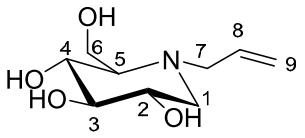
The debenzylation reaction was performed according to a procedure described in Compain, et al. [3], with slight modifications.

N-substituted-tetra-*O*-benzyl-1-DNJ derivatives (**5**, **11**-**14**) or L-ido compound **3** (1 eq) were dissolved in dry CH_2Cl_2 (0,1 M) under argon atmosphere and the solution was cooled to -60°C. 4 eq of BCl_3 (1 M solution in CH_2Cl_2) were added dropwise to this solution. The reaction was stirred under argon and allowed to warm to 0°C over 3 h. TLC (PE/AcOEt 7:3 and AcOEt/MeOH 7:3) confirmed the exhaustive debenzylation. A 20:1 MeOH/H_2O mixture (20 mL) was added and the resulting mixture concentrated under vacuum; this step was repeated twice. The obtained residue was directy used for the next reaction without further purification.



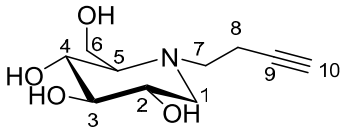
L-ido-N-allyl-1-deoxynojirimycin **15**

Compound **3** (1196 mg, 2,12 mmol) was subjected to complete debenzylation according to the general procedure. The obtain crude residue was used for the successive acetylation without further purification. $C_9H_{17}NO_4$; calcd. mass 203,24; MS-ESI: m/z 204,1 $[M+H]^+$.



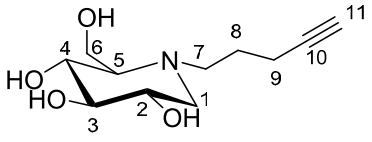
N-allyl-1-deoxynojirimycin **16**

Compound **5** (134 mg, 0,238 mmol) was subjected to complete debenzylation according to the general procedure. The crude was purified by FC (eluent AcOEt/MeOH 8,5:1,5) to afford compound **16** (34 mg, 0,17 mmol, 70% yield). ^{1}H NMR (400 MHz, cd_3od) δ 5.94 (td, *J* = 17.2, 8.0 Hz, 1H, H8), 5.29 – 5.18 (m, 2H, H9a,b), 3.89 (qd, *J* = 12.0, 2.5 Hz, 2H, H6a,b), 3.58 – 3.41 (m, 2H, H7a, H2), 3.37 (t, *J* = 9.3 Hz, 1H, H4), 3.22 – 3.08 (m, 2H, H7b, H3), 3.00 (dd, *J* = 11.3, 4.9 Hz, 1H, H1a), 2.19 – 2.06 (m, 2H, H5, H1b). ^{13}C NMR (101 MHz, $cdcl_3$) δ 134.70 (C8), 119.43 (C9), 80.51 (C3), 71.76 (C4), 70.59 (C2), 67.53 (C5), 58.94 (C7), 57.65 (C6), 56.88 (C1). $C_9H_{17}NO_4$; calcd. mass 203,24; MS-ESI: m/z 204,2 $[M+H]^+$.



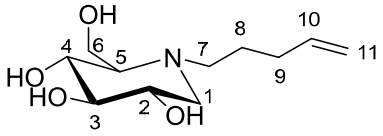
N-(but-3-yn-1-yl)-1-deoxynojirimycin **17**

$C_{10}H_{17}NO_4$; calcd. mass 215,25; MS-ESI: m/z 216,2 $[M+H]^+$.



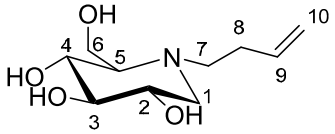
N-(pent-4-yn-1-yl)-1-deoxynojirimycin **18**

$C_{11}H_{19}NO_4$; calcd. mass 229,28; MS-ESI: m/z 230,2 $[M+H]^+$.



N-(pent-4-en-1-yl)-1-deoxynojirimycin **19**

$C_{11}H_{21}NO_4$; calcd. mass 231,29; MS-ESI: m/z 232,2 $[M+H]^+$.

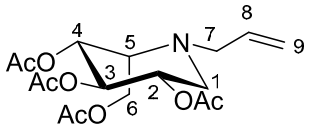


N-(but-3-en-1-yl)-1-deoxynojirimycin **20**

$C_{10}H_{19}NO_4$; calcd. mass 217,27; MS-ESI: m/z 218,2 $[M+H]^+$.

*General procedure for the synthesis of O-Acetylated iminosugars **21**-**25***

Crude compound **15** or **17-20** was dissolved in dry Pyridine (0,2 M) under argon. Ac_2O (6 eq relative to benzylated compound **5** or **11-14**) was added to the solution and the reaction was stirred at r.t.. The reaxction was followed by TLC (AcOEt/MeOH 7:3 and PE/AcOEt 7:3). After 24 h, some drops of MeOH were added and the solvent evaporated to dryness. The residue was purified by FC (PE/AcOEt 7:3) to afford the final O-Acetylated iminosugar derivative.



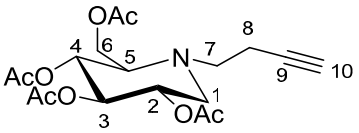
2,3,4,6-tetra-O-acetyl-L-ido-N-allyl-1-deoxynojirimycin **21**

485 mg (1,3 mmol, 45% yield over two steps) of compound **21** were obtained.

^{1}H NMR (400 MHz, CDCl3) δ 5.74 (ddt, J = 16.8, 10.1, 6.3 Hz, 1H, H8), 5.28 – 5.10 (m, 3H, H9a,b, H3), 5.04 (dd, J = 10.2, 6.0 Hz, 1H, H4), 4.93 (td, J = 10.0, 5.8 Hz, 1H, H2), 4.40 (dd, J = 12.1, 6.6 Hz, 1H, H6a), 4.25 (dd, J = 12.1, 3.4 Hz, 1H, H6b), 3.51 – 3.41 (m, 1H, H5), 3.35 (dd, J = 13.9, 5.8 Hz, 1H, H7a), 3.22 (dd, J = 13.9, 6.7 Hz, 1H, H7b), 3.00 (dd, J = 12.3, 5.8 Hz, 1H, H1a), 2.62 (dd, J = 12.0, 10.5 Hz, 1H, H1b), 2.07 (s, 3H), 2.04 (s, 3H, CH3CO), 2.02 – 1.98 (m, 6H, CH3CO).

^{13}C NMR (101 MHz, CDCl3) δ 170.67, 170.29, 170.23, 170.12 (COCH3), 135.12 (C8), 118.33 (C9), 71.16 (C3), 70.57 (C4), 69.98 (C2), 58.81 (C6), 57.69 (C5), 57.48 (C7), 47.51 (C1), 21.17, 21.14, 21.02, 20.99, 20.96, 20.93, 20.90.

$C_{17}H_{25}NO_8$; calcd. mass 371,39; MS-ESI: m/z 372,2 $[M+H]^+$.



2,3,4,6-tetra-O-acetyl-N-(but-3-yn-1-yl)-1-deoxynojirimycin **22**

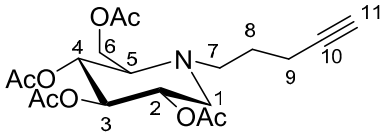
Crude **17** was converted in peracetylated iminosugars **22**, obtaining 168 mg (0,44 mmol) of compound, with a yield of 46% over two steps.

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.08 – 5.00 (m, 2H, H3,4), 5.00 – 4.90 (m, 1H, H2), 4.25 – 4.14 (m, 2H, H6a,b), 3.19 (dd, *J* = 11.4, 5.1 Hz, 1H, H1a), 2.97 (t, *J* = 7.4 Hz, 2H, H7a,b), 2.80 (dd, *J* = 6.0, 2.6 Hz, 1H, H5), 2.48 (dd, *J* =

11.3, 10.3 Hz, 1H, H1b), 2.32 (dd, *J* = 8.6, 5.9 Hz, 2H, H8a,b), 2.09 (s, 3H, C*H*$_3$CO), 2.04 – 2.00 (m, 9H, C*H*$_3$CO), 1.99 (t, *J* = 2.6 Hz, 1H, H10).

^{13}C NMR (101 MHz, $CDCl_3$) δ 170.97, 170.48, 170.16, 169.87 (*C*OCH$_3$), 81.99 (C9), 74.59 (C10), 70.16 (C3), 69.46 (C4), 69.33 (C2), 60.78 (C5), 60.00 (C6), 53.03 (C1), 50.52 (C7), 21.02, 20.98, 20.90, 20.83 (CO*C*H$_3$), 15.31 (C8).

$C_{18}H_{25}NO_8$; calcd. mass 383,40; MS-ESI: m/z 384,2 $[M+H]^+$.



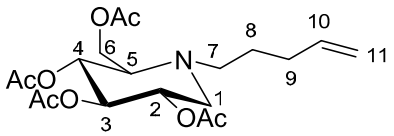
2,3,4,6-tetra-O-acetyl-N-(pent-4-yn-1-yl)-1-deoxynojirimycin **23**

Crude **18** was converted in peracetylated iminosugars **23** (128 mg, 0,322 mmol, 36% yield over two steps).

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.06 – 4.95 (m, 2H, H3, H4), 4.95 – 4.85 (m, 1H, H2), 4.19 – 4.06 (m, 2H, H6a,b), 3.15 (dd, *J* = 11.5, 5.0 Hz, 1H, H1a), 2.93 – 2.79 (m, 1H, H7b), 2.66 – 2.48 (m, 2H, H5, H7a), 2.27 (t, *J* = 10.8 Hz, 1H, H1b), 2.17 (td, *J* = 6.8, 2.4 Hz, 2H, H9a,b), 2.03 (s, 3H, C*H*$_3$CO), 2.01 – 1.93 (m, 9H, C*H*$_3$CO), 1.91 (t, *J* = 2.4 Hz, 1H, H11), 1.71 – 1.49 (m, 2H, H8a,b).

^{13}C NMR (101 MHz, $CDCl_3$) δ 170.92, 170.37, 170.05, 169.77 (*C*OCH$_3$), 83.54 (C10), 74.59 (C3), 69.42 (C4), 69.25 (C2), 68.99 (C11), 61.70 (C5), 59.66 (C6), 52.93 (C1), 50.14 (C7), 24.30 (C8), 20.80, 20.78, 20.74, 20.72 (CO*C*H$_3$), 15.91 (C9).

$C_{19}H_{27}NO_8$; calcd. mass 397,42; MS-ESI: m/z 398,3 $[M+H]^+$.



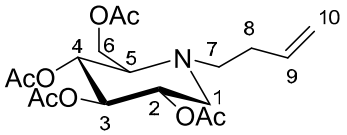
2,3,4,6-tetra-O-acetyl-N-(pent-4-en-1-yl)-1-deoxynojirimycin **24**

Crude **19** was converted in peracetylated iminosugars **24** (149 mg, 0,37 mmol, 40% yield over two steps).

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.77 (ddt, *J* = 16.9, 10.2, 6.6 Hz, 1H, H10), 5.16 – 4.84 (m, 5H, H11a,b, H2,H3,H4), 4.15 (d, *J* = 2.4 Hz, 2H, H6a,b), 3.19 (dd, *J* = 11.4, 5.0 Hz, 1H, H1a), 2.81 – 2.70 (m, 1H, H7a), 2.64 (d, *J* = 9.0 Hz, 1H, H5), 2.62 – 2.52 (m, 1H, H7b), 2.33 (t, *J* = 10.7 Hz, 1H, H1b), 2.07 (s, 3H, C*H*$_3$CO), 2.06 – 1.97 (m, 11H, C*H*$_3$CO x 9, H9a,b), 1.61 – 1.42 (m, 2H, H8a,b).

^{13}C NMR (101 MHz, $CDCl_3$) δ 171.07, 170.52, 170.17, 169.88 (*C*OCH_3), 137.92 (C10), 115.30 (C11), 74.75 (C2), 69.56 (C3), 69.49 (C4), 61.53 (C5), 59.63 (C6), 53.03 (C1), 51.28 (C7), 31.28 (C9), 24.01 (C8), 21.02, 20.98, 20.90, 20.83 (CO*C*H_3).

$C_{19}H_{29}NO_8$; calcd. mass 399,44; MS-ESI: m/z 400,4 $[M+H]^+$.



2,3,4,6-tetra-O-acetyl-N-(but-3-en-1-yl)-1-deoxynojirimycin **25**

326 mg (0,85 mmol) of compound **25** were obtained, with a yield of 56% yield over two steps.

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.72 (ddt, *J* = 17.0, 10.1, 6.8 Hz, 1H, H9), 5.11 – 4.99 (m, 4H, H10a,b, H3,4), 4.95 (td, *J* = 9.8, 5.1 Hz, 1H, H2), 4.16 (d, *J* = 2.6 Hz, 2H, H6a,b), 3.20 (dd, *J* = 11.4, 5.0 Hz, 1H, H1a), 2.89 – 2.75 (m, 1H, H7a), 2.75 – 2.61 (m, 2H, H7b, H5), 2.39 (t, *J* = 10.9 Hz, 1H, H1b), 2.28 – 2.12 (m, 2H, H8a,b), 2.07 (s, 3H, C*H*$_3$CO), 2.05 – 1.96 (m, 9H, C*H*$_3$CO).

^{13}C NMR (101 MHz, $CDCl_3$) δ 171.06, 170.52, 170.16, 169.87 ($COCH_3$), 135.61 (C9), 116.66 (C10), 74.71 (C2), 69.53 (C3), 69.48 (C4), 61.17 (C5), 59.62 (C6), 52.99 (C1), 51.30 (C7), 29.26 (C8), 21.02, 20.98, 20.90, 20.83 ($COCH_3$).

$C_{18}H_{27}NO_8$; calcd. mass 385,41; MS-ESI: m/z 386,2 $[M+H]^+$.

*Synthesis of calix[4]arenes structures **27-32***

The synthesis, purification and characterization of the presented calix[4]arenes are already described in literature in Dondoni and Marra [8], Marra, et al. [9], Fiore, et al. [6]. All characterization data are consistent to those published.

*Synthesis of calix[4]arenes **33**-**34**, **35**-**36**, **37**-**38** and monovalent ligands **39**-**40**.*

The synthesis of the final calix[4]arene was performed according to the same procedure described in the works mentioned above. NMR and mass characterization has to be completed.

References

[1] C. Fasting, C. A. Schalley, M. Weber, O. Seitz, S. Hecht, B. Koksch, J. Dernedde, C. Graf, E.-W. Knapp, R. Haag, *Angewandte Chemie-International Edition* **2012**, *51*, 10472-10498.

[2] J. Diot, M. I. Garcia-Moreno, S. G. Gouin, C. O. Mellet, K. Haupt, J. Kovensky, *Organic & Biomolecular Chemistry* **2009**, *7*, 357-363.

[3] P. Compain, C. Decroocq, J. Iehl, M. Holler, D. Hazelard, T. Mena Barragan, C. Ortiz Mellet, J.-F. Nierengarten, *Angewandte Chemie-International Edition* **2010**, *49*, 5753-5756.

[4] C. Decroocq, D. Rodriguez-Lucena, V. Russo, T. Mena Barragan, C. Ortiz Mellet, P. Compain, *Chemistry-a European Journal* **2011**, *17*, 13825-13831.

[5] S. Cecioni, O.-A. Argintaru, T. Docsa, P. Gergely, J.-P. Praly, S. Vidal, *New Journal of Chemistry* **2009**, *33*, 148-156.

[6] M. Fiore, A. Chambery, A. Marra, A. Dondoni, *Organic & Biomolecular Chemistry* **2009**, *7*, 3910-3913.

[7] M. Marradi, S. Cicchi, F. Sansone, A. Casnati, A. Goti, *Beilstein Journal of Organic Chemistry* **2012**, *8*, 951-957.

[8] A. Dondoni, A. Marra, *Journal of Organic Chemistry* **2006**, *71*, 7546-7557.

[9] A. Marra, L. Moni, D. Pazzi, A. Corallini, D. Bridi, A. Dondoni, *Organic & Biomolecular Chemistry* **2008**, *6*, 1396-1409.

[10] C. R. R. Matos, R. S. C. Lopes, C. C. Lopes, *Synthesis-Stuttgart* **1999**, 571-573.

[11] T. Wennekes, R. J. B. H. N. van den Berg, W. Donker, G. A. van der Marel, A. Strijland, J. M. F. G. Aerts, H. S. Overkleeft, *The Journal of Organic Chemistry* **2007**, *72*, 1088-1097.

[12] M. Fiore, M. Lo Conte, S. Pacifico, A. Marra, A. Dondoni, *Tetrahedron Letters* **2011**, *52*, 444-447.

[13] Y. Jiang, Z. Fang, Q. Zheng, H. Jia, J. Cheng, B. Zheng, *Synthesis-Stuttgart* **2009**, 2756-2760.

[14] V. S. Rao, A. S. Perlin, *Canadian Journal of Chemistry-Revue Canadienne De Chimie* **1984**, *62*, 886-890.

Chapter 6. Antiproliferative activity of Arsenical C-glucoside derivative on neuroblastoma cell line SN-K-BE

Abstract

C-glucoside derivatives covalently linked to the arsenic atom in different oxidation states have been synthesized. While compound **5**, bearing an As^{V} atom, did not show any significant biological activity, compound **6**, in which the glucose moiety is linked to the arsenic atom as dithioarsenal (As^{III}), showed promising antiproliferative activity on cell line of neuroblastoma (SK-N-BE). In conclusion, compound **6**, synthesised avoiding protection-deprotection steps, represents a promising lead compound for the development of selective antitumor agents.

Introduction

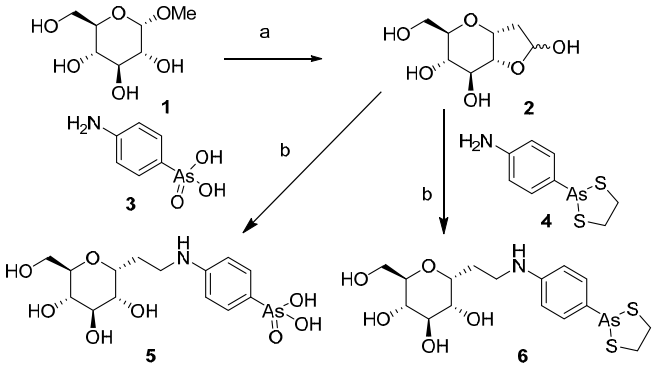
Since Hippocrates and then through Chinese alchemy, Fowler's solution, Erlich's Salvarsan, until the recent use of arsenic trioxide as a therapy for promyelocytic leukemia, arsenic has represented a Janus bifrons of medicine, being remedy and poison, healer and killer. This double-edged sword properties stem for a large part from arsenic physicochemical similarity with phosphorus as metabolic pathways intended for PO_4^{3-} may not distinguish it from AsO_4^{3-}[1]. This means that arsenic is a serious threat for health but also an ideal bullet to be delivered to a therapeutic target in a masked way, provided it is conjugated with a highly selective carrier. Aiming at the design of an arsenic-based anti-cancer drug, we envision to exploit the preferential uptake of glucose by cancer cells, known as Warburg effect. This effect starts in the early stages of carcinogenesis and seems to increase with tumor aggressiveness[2]. Despite this process being functional for tumor growth, it is absolutely unfavorable from an energetic point of view, thus forcing tumor cells to import a great quantity of glucose. This results in an increase in the

expression of membrane proteins devoted to the uptake of glucose molecules, in particular the GLUT transporters[3]. The use of 2-(^{18}F)-fluoro-2-deoxy-D-glucose (FDG) as tracer in the positron emission tomography (PET) confirms the *in vivo* accumulation of glucose inside tumor cells, and represents the main application of the Warburg effect. The success in this diagnostic application has suggested the therapeutic exploitation of the preferential uptake of glucose, through its conjugation to cytotoxic drugs, such as cyclophosphamide. Unfortunately these bulky conjugates do not penetrate inside the cells through the GLUT transporters, but through the co-trasporter sodium-glucose (SAAT1)[4], thus showing a reduced selectivity towards tumor cells. Being devoid of significant steric hindrance, arsenic is in principle an ideal conjugate for glucose. Furthermore, by matching the selectivity of glucose to the widespread biological impact of arsenic, we meant to obtain compounds able to specifically deliver to cancer cells a pharmacological agent capable of simultaneously targeting different metabolic pathways. This approach may be particularly useful in cancer since it may limit the development of resistance by targeting different key points of tumor metabolism.

In this work we generated arsenical *C*-glucoside derivatives, in which the arsenic group is covalently linked to the sugar moiety through the alkyl appendage at the anomeric centre, thus maintaining the pyranosidic structure of the glucose moiety whilst generating a metabolically stable derivative. The arsenal compound should therefore accumulate and exert its toxic effect inside tumor cell.

Results and Discussion

We designed and synthesized two *C*-glucoside derivatives covalently linked to the arsenic atom in different oxidation states as arsenic acid (As^{V}) and dithioarsenal (As^{III}) (**5** and **6**, respectively) and carried out preliminary *in vitro* biological evaluation on a tumoral cell line of human neuroblastoma (SK-N-BE).



Scheme 1: Reagents and conditions: a) Ref.[6]; b) MeOH/H_2O, $NaCNBH_3$, 3 or 4, yield 5 (35%) and 6 (44 %).

The synthesis (scheme 1) was designed in order to avoid the use of protecting groups and to insert the arsenic containing moiety at the end of the synthetic route. As arsenic moiety we chose the two readily available phenyl arsine derivatives **3** and **4**, that can be chemoselectively attached to the hemiacetalic group of the glucose derivative **2** by reductive amination. Thus the key C-glycosyl intermediate **2** was generated by stereoselective *C*-allylation[5] of methyl α-D-glucopyranoside, followed by ozonolysis and reductive amination of the obtained aldehyde (present as hemiacetal **2**)[6].

The effects of the two compounds were tested on human neuroblastoma cells (SK-N-BE) and compared with that of As_2O_3. While

compound **5** did not show any significant biological activity at the concentration tested (10, 30, 50, 100 µM), **6** and As_2O_3 led to a concentration-dependent increase of cells with morphological signs of apoptosis (irregular outline, picnotic nuclei, spherical conformation and loss of adherence). The data showed that As_2O_3 significantly reduces the proliferative activity of SK-N-BE cells in a dose- and time-dependent manner, while the activity of **6** was dose-dependent but it was not influenced by time (figure 1). The antiproliferative effect at different time points, expressed as IC_{50}, is shown in table 1.

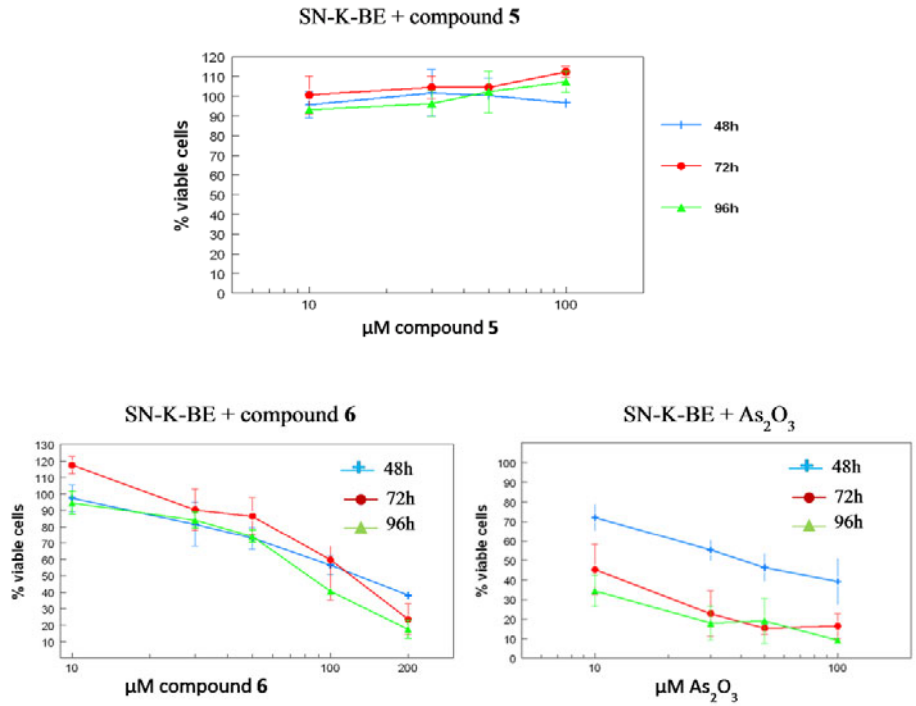


Figure 1. Cell viability analysis of SK-N-BE cells treated with compound 5, 6 or As_2O_3.

Table 1: The in vitro antiproliferative activity of As_2O_3 and compound 6.

Test compound	IC_{50} (48h)	IC_{50} (72h)	IC_{50} (96h)
As_2O_3	28.18 μM	7.10 μM	3.31 μM
Compound **6**	110.5 μM	91.93 μM	88.43 μM

In conclusion, we have demonstrated the antiproliferative effect of an As^{III} derivative covalently linked to glucose, that, combined with the Warburg effect, opens the way to new selective anticancer agents. Interestingly, the antiproliferative effect of the arsenal *C*-glucoside is not time-dependent like that of As_2O_3. Compound **6**, synthesized avoiding protection-deprotection steps, represents therefore a potential lead compound for the development of a "magic bullet" aiming at cancer cells.

Experimental Section

General Remarks. All solvents, when necessary, were dried with molecular sieves for at least 24 h prior to use. 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 $(NH_4)_6Mo_7O_{24}$ (21 g), $Ce(SO_4)_2$ (1 g), concd. H_2SO_4 (31 mL) in water (500 mL). Flash column chromatography was performed on silica gel 230-400 mesh (Merck). ^1H and ^{13}C NMR spectra were recorded at 25 °C unless otherwise stated, with a Varian Mercury 400 MHz instrument. Chemical shift assignments, reported in ppm, are referenced to the corresponding solvent peaks. HRMS were recorded on a ABSciex 2000 QTRAP LC/MS/MS System with an ESI source. Optical rotations were measured at room temperature

using an Atago Polax-2L polarimeter and are reported in units of 10-1 deg·cm^3·g^{-1}.

4-[2-(α-D-glucopyranosyl)ethylamino]phenyl)arsonic acid 5: crude compound **2** (1.46 mmol, 1 equiv.) was treated with an HCl 0.01 M solution (7 mL) for 1 h at 70 °C. The pH of the solution was adjusted to pH > 7 and compound 3 (1.75 mmol, 1.2 equiv.), suspended in MeOH/H_2O (30 mL), was added. Then, the pH of the solution was decreased to pH = 4 with AcOH; after 20 min, $NaCNBH_3$ (2.04 mmol, 1.4 equiv.) was added. After 30 min the solvent was removed under reduced pressure and the crude is purified by flash chromatography (AcOEt/MeOH/H_2O/AcOH 70/30/5/5) affording compound **5** with an overall yield of 35% from 2.: $[\alpha]^D_{25}$ = + 61.7 (*c* = 1, H_2O), ^{1}H NMR (400 MHz, D_2O) δ 7.53 (d, J = 8.4 Hz, 2H), 6.84 (d, J = 8.4 Hz, 2H), 4.18 – 4.03 (m, 1H), 3.90 – 3.74 (m, 1H), 3.73 – 3.61 (m, 2H), 3.61 – 3.54 (m, 1H), 3.54 – 3.48 (m, 1H), 3.40 – 3.13 (m, 3H), 2.05 – 1.89 (m, 2H). ^{13}C NMR (100 MHz, D_2O) δ ppm 154.6, 134.1, 123.7, 116.3, 76.83, 75.93, 75.45, 73.664, 72.99, 63.74, 42.32, 25.64; MS calcd for $C_{14}H_{22}AsNO_8$ [M + H]+ 408.1: found 408.1.

4-(1,3,2-Dithiarsolan-2-yl)-*N*-[2-(α-D-glucopyranosyl)ethyl]aniline 6: crude compound **2** (0.97 mmol, 1 equiv.) was treated with an HCl 0.01 M solution (7 mL) for 1 h at 70 °C. The pH of the solution was adjusted to pH > 7 and compound **4** (1.16 mmol, 1.2 equiv.), suspended in CH_2Cl_2/MeOH 3:4 (70 mL), was added. Then, the pH of the solution was decreased to pH = 4 with AcOH; after 20 min, $NaCNBH_3$ (1.36 mmol, 1.4 equiv.) was added. After 30 min the solvent is removed under reduced pressure and the crude was purified by flash chromatography (AcOEt/MeOH 95/5) affording compound **6** with an overall yield of 44 %

from **2**.: $[\alpha]^{D}_{25}$ = + 31.7 (c = 0.5, DMSO); ^{1}H NMR (400 MHz, CD_3OD) ppm 7.29-7.21 (m, 1H), 7.04-6.94 (m, 1H), 6.60-6.46 (m, 2H), 4.00-3.90 (m, 1H), 3.75 (bd, J = 11.78 Hz, 1H), 3.56-3.47 (m, 2H), 3.44-3.35 (m, 2H), 3.23-3.06 (m, 7H), 1.90-1.81 (m, 2H). ^{13}C NMR (100 MHz, CD_3OD) 166.7, 153.0, 135.8, 132.8, 117.0, 116.2, 78.79, 78.03, 77.60, 75.67, 75.27, 66.08, 44.94, 44.80, 27.77; MS calcd for $C_{16}H_{24}AsNO_5S_2$ [M + H]+ 450: found 450.

Drugs and chemicals: Arsenic trioxide (As_2O_3) was purchased from Sigma Aldrich Company (MI). Arsenic trioxide was dissolved in 0,1 M NaOH and kept as a stock of 1 mM As_2O_3 dissolved in distilled water. Growth medium DMEM High glucose, fetal bovine serum (FBS), antibiotics (penicillin and streptomycin), L-Glutamine and L-phosphate buffered saline (PBS) were obtained from Euroclone. 3-(4, 5-dimethylthiasol-2-yl)-2, 4,-diphenyltetrazolium bromide (MTT) solution and Dimethyl sulfoxide (DMSO) were purchased from Sigma Aldrich Company (MI).

Cell lines and culture conditions. Human neuroblastoma cell line SK-N-BE was cultured in DMEM High glucose supplemented with 10% fetal bovine serum (FBS), 1% (w/v) penicillin/streptomycin and 2 mM L-glutamine in a humidified incubator containing 95% air + 5% CO_2 atmosphere.

MTT assay: Cell proliferation was evaluated by MTT metabolism as previously described[7]. Briefly, SK-N-BE were plated in 96-well polystyrene tissue culture plates (Falcon) at a density of 5×10^3/well in fresh medium. After 48h of incubation As_2O_3 or compound **6** solutions (range 10-200 µM) were added to each well. Cells incubated in culture medium alone served as untreated control. Cells were cultured for 48, 72 and 96 hours respectively. After incubation, 20 µl aliquots of MTT

solution (5 mg/mL in PBS) were added to each well and re-incubated for 4 hours at 37°C. Then, the supernatant culture medium were carefully aspirated and 200 µl aliquots of dimethylsulfoxide (DMSO) were added to each well to dissolve the formazan crystals, followed by incubation for 10 minutes to dissolve air bubbles. The culture plate was placed on a micro-plate reader and the absorbance was measured at 540 nm. The amount of color produced is directly proportional to the number of viable cell. All assays were performed in six replicates for each concentration and means ± SD values were used to estimate the cell viability. Cell proliferation rate was calculated as the percentage of MTT absorption as follows: % survival = (mean experimental absorbance/mean control absorbance) × 100. Morphological changes were followed by phase contrast microscopy.

Statistical analysis: Student *t* test was used for all comparison, all tests were two-tailed. The IC_{50} value were calculated using GraphPad Prism 5.

References

[1] Rosen, B. P. FEBS Lett., 2002, 529(1), 86-92.

[2] Cairns, R. A.; Harris, I. S.; Mak, T. W. *Nat Rev Cancer*, **2011**, *11(2)*, 85-95.

[3] Zhao, F. Q.; Keating, A. F. *Current Genomics*, **2007**, *8*, 113-128.

[4] Veyhl, M.; Wagner, K.; Volk, C.; Gorboulev, V.; Baumgarten, K.; Weber, W. M.; Schaper, M.; Bertram, B.; Wiessler, M.; Koepsell, H. *Proc Natl Acad Sci USA*,**1998**, *17;95(6)*, 2914-9.

[5] Bennek, J. A.; Gray, G. A. *J. Org. Chem*., **1987**, *52*, 892-897.

[6] Cardona, F.; La Ferla, B. *J. Carbohydr. Chem*., **2008**, *27(4)*, 203-213.

[7] Mosmann, T. *J Immunol Methods*, **1983**, *65(1-2)*, 55-63.

Chapter 7. Conclusions

Carbohydrates represent not only an important class of macromolecules, widely spread in Nature and involved in practically all the biological process of living organism, but also tools that chemists and biochemists can use for the generation of potential drug candidates, and conjugates or derivatives, for the investigation and the study of the processes in which carbohydrates are involved.

In the first case, we started from an anti-inflammatory effect observed with high glucose doses administered to *in vitro* and *in vivo* inflammatory models, that suggested the search of glucose derivatives bearing the same function. This prompted us to generate a library of metabolically inert C-glycoside derivatives; these bioactive molecules represent both promising drug candidates, for the treatment of severe inflammatory disease, as sepsis, endotoxic shock, mucositis or Chron disease, and chemical tools for the comprehension of the mechanism at the base of this observed phenomenon. The use of these carbohydrate-based structure allowed to make more evident a new biological role, a sort of immunological protection, for a transport protein (SGLT1), for which the physiological function has been well established.

The metabolic behavior of sugars, and in particular of glucose, in tumor cells, represent the fundament of our idea of the development of new carbohydrate-based antiproliferative agents, containing the cytotoxic element Arsenic. The synthesized molecule represents a lead compound, a starting point for the development of more and more selective and effective anticancer agents, for which there is a high demand.

Iminosugars are a class of glycoderivatives widely studied both by chemists and biochemists; their inhibitory and chaperon-like activity on glycosidases prompted the researchers towards the synthesis of more and more analogues, with increasing efficacy; nowadays, two iminosugars are commercially available for the treatment of diabetes

mellitus type II (Miglitol) and for Gaucher disease (Miglustat). Beside their clinical importance, iminosugars have been used and developed as models and tools for the study of phenomena, like multivalent inhibiton of glycosidases, which could lead to new details and knowledge about how these critical enzymes work in living organisms.

Pubblications and communications

Papers

La Ferla, B., Spinosa, V., D'Orazio, G., Palazzo, M., Balsari, A., Foppoli, A. A., Rumio, C. and Nicotra, F. (2010), "Dansyl C-Glucoside as a Novel Agent Against Endotoxic Shock". *ChemMedChem*, 5: 1677–1680.

Barbara La Ferla, Cristina Airoldi, Cristiano Zona, Alexandre Orsato, Francisco Cardona, Silvia Merlo, Erika Sironi, Giuseppe D'Orazio and Francesco Nicotra, *Natural glycoconjugates with antitumor activity*, *Natural Product Report*, **2011**, 28, 630-648

Oral communications

G. D'Orazio, B. La Ferla, M. Coppola, C. Magni, F. Nicotra, "Sodium-Glucose Cotransporter 1 (SGLT1) target receptor for carbohydrate based anti-inflammatory agents". *Italian-Spanish Joint Workshop CARIPLO project 2008-3175, Chemical tools for molecular recognition studies: Synthesis and NMR characterization of bioactive molecules*, Università degli Studi di Milano – Bicocca, 22/4/2010.

Giuseppe D'Orazio, Diego Cardani, Cristiano Rumio, Francesco Nicotra, Barbara La Ferla, "SGLT1-MEDIATED ANTIINFLAMMATORY AND PROTECTIVE EFFECT OF DANSYL-GLYCODERIVATIVES. SYNTHESIS, BIOLOGICAL EVALUATION AND STUDIES ON THE MECHANISM OF ACTION", *XIII Convegno-Scuola sulla Chimica dei Carboidrati*, Pontignano (Si), June 24th – 27th, 2012.

Other communications

G. D'Orazio, B. La Ferla, M. Coppola, C. Rumio, F. Nicotra, "Design and synthesis of 2nd generation carbohydrate-based antiinflammatory agents" - *XXXV Corso Estivo "A. Corbella - Seminari di Chimica Organica" della Società Chimica Italiana*, Gargnano (Palazzo Feltrinelli), 14 – 18/06/2010.

Giuseppe D'Orazio, Barbara La Ferla, Marta Coppola, Cristiano Rumio, Francesco Nicotra "DANSYLATED C-GLYCOSIDES: DRUG CANDIDATES AS ANTIIFLAMMATORY AGENTS AND MOLECULAR TOOLS FOR BIOLOGICAL STUDIES", *16th European Carbohydrate Symposium (Eurocarb 16)* – Sorrento, 03-07/07/2011.

D'Orazio, G.; Zona, C.; Parisi, G.; Policano, C.; Mechelli, R.; Codacci-Pisanelli, G.; Pitaro, M.; Ristori, G.; Salvetti, M.; Nicotra, F.; La Ferla, B.; "ARSENIC C-GLYCOSIDES: APPLYING THE PAUL ERLICH' "MAGIC BULLET" CONCEPT IN THE TREATMENT OF CANCER", *26th International Carbohydrate Symposium (ICS2012)*, Madrid, July 22nd to 27th, 2012.

Giuseppe D'Orazio, Diego Cardani, Cristiano Rumio, Francesco Nicotra, Barbara La Ferla, "SGLT1-MEDIATED ANTIINFLAMMATORY AND PROTECTIVE EFFECT OF DANSYL-GLYCODERIVATIVES. SYNTHESIS, BIOLOGICAL EVALUATION AND STUDIES ON THE MECHANISM OF ACTION", *XIII Convegno-Scuola sulla Chimica dei Carboidrati*, Pontignano (Si), June 24th – 27th, 2012.

Barbara La Ferla, Cristiano Zona, Giuseppe D'Orazio, Francisco Cardona, Crisitina Airoldi, Alexandre Orsato, Francesco Nicotra, "Glycomimetics

conjugated Nanoparticles with therapeutic activity and glycidic scaffold based targeting agents", *MultiGlycoNano - COST-Meeting 2012*, Berne (Switzerland), 2nd-4th February 2012.

List of abbreviations

aa	Aminoacid
Ac	Acetyl
Ac_2O	Acetic anhydride
AcOEt	Ethyl acetate
AcOH	Acetic acid
Ag_2O	Silver oxide
AllylMgBr	Allylmagnesium bromide
AllylTMS	Allyltrimethylsilane
Asn	Asparagine
ATP	Adenosine triphosphate
B^0AT1	Sodium-glutamine cotransporter 1
BBM	Brush border membrane
BCl_3	Boron trichloride
$BF_3{\cdot}OEt_2$	Boron trifluoride diethyl etherate
Bn	Benzyl
BnBr	Benzyl bromide
BSTFA	N,O-Bis(trimethylsilyl) trifluoroacetamide
Caco-2	Human epithelial colorectal adenocarcinoma cell
CD14	Cluster of differentiation 14
CH_2Cl_2	Dichloromethane
CH_3CN	Acetonitrile
$(COCl)_2$	Oxalyl chloride
CpG-ODN	CpG-Oligodeoxynucleotide
CSA	Camphorsulphonic acid
CSGAGs	Chondroitin/dermatan
CuAAC	Copper(I)-catalyzed azide alkyne cycloaddition
DansylCl	Dansyl chloride
DCC	*N*-*N*′-dicyclohexyl carbodiimide
DIAD	Diisopropyl azodicarboxylate
DIPEA	Diisopropylethylamine
DMAP	Dimethylaminopyridine
DMEM	Dulbecco's Modified Eagle Medium
DMF	Dimethylformamide
DMSO	Dimethylsulfoxide
DNA	Deoxy-ribonucleic acid
1-DNJ	1-Deoxynojirimycin
DPAP	2,2-Dimethoxy-2-acetophenone
ECM	Extracellular matrix
EDC	N-(3-Dimethylamino propyl)-N′-ethyl carbodiimide hydrochloride
EGF	Epidermal Growth Factor
ELISA	Enzyme-Linked ImmunoSorbed Assay
ER	Endoplasmic Reticulum
ESI-MS	ElectroSpray Ionization – Mass Spectroscopy
Et_2O	Diethylether
EtOH	Ethanol
Et_3N	Triethylamine
Et_3SiH	Triethylsilane

FBS	Fetal bovine serum
FC	Flash chromatography
FDG	2-(^{18}F)-fluoro-2-deoxy-D-glucose
GAGs	Glycosaminoglycans
Gal	Galactose
GGM	Glucose-galactose malabsorption
Glc	Glucose
L-Gln	L-Glutamine
GLUT	Glucose transporter
GSLs	Glycosphingolipids
HCl	Hydrochloric acid
H_2O	Water
HOBt	Hydroxybenzotriazole
HSGAGs	Heparin/Heparan
HT29	Human colon adenocarcinoma cell line
IBDs	Intestinal Bowel Disease
IC_{50}	Half maximal inhibitory concentration
IECs	Intestinal Epithelial Cells
IGF	Insulin Growith Factor
IL	Interleukin
iPr_2EtN	Diisopropylethylamine
K_2CO_3	Potassium carbonate
LPS	Lipopolysaccharide
LSDs	Lysosomal storage disease
MD2	Lymphocyte antigen 96
MeOH	Methanol
MeONa	Sodium methylate
MHC-II	Class II Major Histocompatibility Complex
MsCl	methanesulfonyl chloride
MTT	3-(4,5-dimethylthiasol-2-yl)-2,4-Diphenyl tetrazolium bromide
Na	Sodium
$NaBD_4$	Sodium borodeuteride
$NaBH_4$	sodium borohydride
$NaBT_4$	Sodium borotritide
$NaCNBH_3$	Sodium cyanoborohydride
NaH	Sodium hydride
$NaIO_4$	Sodium periodate
Na^+/K^+-ATPase	Sodium-potassium ATP-synthase
NaN_3	Sodium azide
NANA	N-Acetyl Neuraminic Acid (Sialic Acid)
NaOH	Sodium hydroxide
NH_3	Ammonia
NH_4Cl	Ammonium chloride
$NH_4^+HCOO^-$	Ammonium formate
NIS	*N*-iodo-succinimide
NMR	Nuclear Magnetic Resonance
NOESY	Nuclear Overhauser Effect SpectroscopY
3OMG	3-*O*-Methyl-D-glucopyranose
OMIM	Online Mendelian Inheritance in Man
ORT	Oral Rehydration Therapy
OsO_4	Osmium tetroxide
PAMPs	Pathogen-Associated Molecular Patterns
PBS	Phosphate Buffer

	Saline
$Pd(OH)_2/C$	Palladium hydroxide on activated charcoal
PE	Petroleum ether
PEG	Polyethyleneglycol
PGs	Proteoglycans
$(PhO)_2PON_3$	Diphenyl phosphoryl azide
Ph_3P	Triphenylphosphine
PRRs	Pattern Recognition Receptors
Py	Pyridine
QC	Quality Control
ROS	Reactive Oxygen Species
Satd.	Saturated
Ser	Serine
SGLT1	Sodium Glucose coTransporter 1
SIRS	Systemic Inflammatory Response Syndrome
SK-N-BE	Human neuroblastoma cell line
SMF	Sodium Motive Force
SPR	Surface Plasmon Resonance
SSSFs	Sodium substrate symporter family
tBuOH	Tert-butanol
TEC	Thiol-ene coupling
TEG	Triethyleneglycol
TEM	Transmission Electron Microscopy
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
Thr	Threonine
TLC	Thin Layer Chromatography
TLR	Toll-Like Receptor
TM	Transmembrane
TMSOTF	Trimethylsilil trifluoromethane sulphonate
TNFα	Tumor Necrosis Factor α
TS	Transition State
TsCl	Toluensulfonyl chloride