Università degli studi di Parma

Dottorato di Ricerca in Scienze Chimiche

Ph.D. in Chemical Sciences

XXIV ciclo

Multivalent Glycocalixarenes for Lectin Inhibition

Tutors:

Prof. Alessandro Casnati

Dr. Francesco Sansone

Dottoranda:

Silvia Bernardi

Coordinatore:

Prof. Giovanni Predieri

2009-2011

Alla mia famiglia,

con immensa riconoscenza

Contents

Chapter 3

Studies of the lectin-glycocalixarene interactions by NMR techniques and computational methods

Chapter 4

New glycocalix[4]arenes via Cu-catalyzed azide-alkyne cycloaddition

Chapter 5

Pentavalent glycocalix[5]arenes for cholera toxin inhibition

Abstract of the thesis

The growing knowledge of the biological role played by the glycoside cluster effect prompted us to design and synthesize compounds able to inhibit target lectins with high selectivity and efficiency by combining a multivalent presentation of ligands with carbohydrates tailored for the best-possible contact complementarity with the proteins. To this aim calixarenes were chosen as multivalent scaffolds because they offer the unique opportunity to easily change the valency, the molecular shape, the conformation and the symmetry of the glycocluster. The work reported in this Ph.D. thesis concerns the synthesis and properties of a series of calix[n]arenes containing units of galactose, lactose, LacNAc or LacNAc modified at 2-N or 3' positions and connected to the macrocyclic structure via thioureido or triazole groups. These compounds showed high affinity for different types of galectins. Quite interestingly, the selectivity for the different lectins was confirmed to be also dependent on the calixarene conformation, pointing out the importance of the multivalent presentation in space of the glycosyl units. Enzymatic galactosylation and 2,3- and 2,6-sialylation reactions were also explored using glycosylcalixarenes as substrate. Moreover, NMR and computational studies were performed on selected glycocalixarenes in order to better understand the nature of the carbohydrate-lectin interactions in solution on an atomic level. At last, pentavalent glycocalix[5]arenes functionalized with carbohydrates of increasing complexity, including GM2 and GM1 oligosaccharides, were synthesized to allow a perfect match in valency with the target lectin, the pentavalent B subunit of cholera toxin. Preliminary ELISA tests remarkably showed an IC_{50} value in low nanomolar range indicating the GM1-derivatized glycocalix[5]arene to be one of the most efficient cholera toxin inhibitors known so far.

Chapter 1

Glycobiology and multivalency

1.1 Carbohydrates: the sweet molecules of life

Together with amino acids and nucleotides, carbohydrates constitute one of the most important classes of biomolecules in nature. It is well known that they not only have important structural functions (like in the case of cellulose and chitin) and are one of the main energy reserves of cells (as glycogen and starch), but they also play a fundamental role in molecular recognition events of many biological processes. Carbohydrates are in fact involved in a wide variety of interactions with different biological species such as proteins, hormones, toxins, viruses, bacteria, also acting as mediator of cell-cell communication (figure **1.1**)[1]. The incredibly high structural diversity which can be generated by these oligosaccharides, the so called Sugar Code[2], stores rich and detailed levels of biochemical information which can be decrypted by suitable sugar binding proteins.

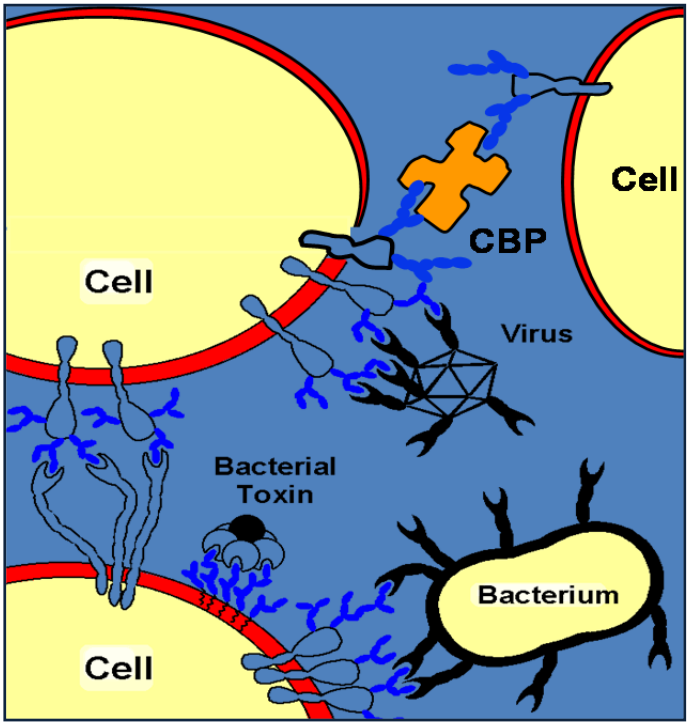


Figure 1.1 Schematic representation of possible interactions between carbohydrates and other biological receptors[3].

The complexity of these oligosaccharides derives from the presence on the single carbohydrate of many attachment points, that allow their connection via a great variety of linkages and can lead to the formation of highly branched and stereochemically different

compounds[4]. A single carbohydrate can in fact be connected to another sugar moiety through any of the hydroxyl groups and, in addition, the stereochemistry of each glycoside bond can be either α or β. On the contrary, amino acids and nucleotides can form only linear oligo- and polymers, since there is just one possible way of connection (figure **1.2**).

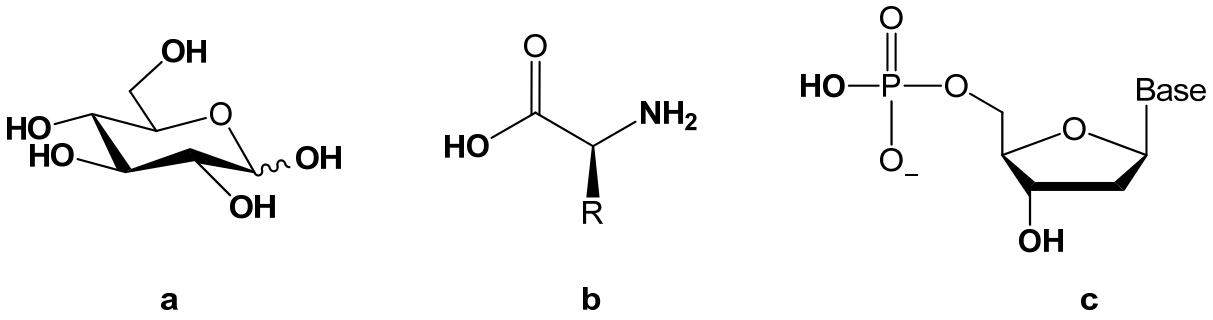


Figure 1.2 Monomeric building blocks for oligosaccharides, proteins and nucleic acids (**a**, **b**, **c**, respectively). In **bold** are underlined the groups potentially involved in oligomerization processes.

It can been calculated that for a hexanucleotide 4^6 different structures may be found (using the four types of basis), while for a hexapeptide 6 x 10^7 isomers can be formed. If similar calculations are made for a hexasaccharide, more than 10^{12} different structures might be obtained[5]. This structural complexity is also increased by the number of possible modifications that can be introduced. For example the hydroxyl functionalities can be acetylated, oxidized, phosphorylated, sulforylated and replaced by other groups, especially containing nitrogen. It is not surprising therefore that nature uses carbohydrates to mediate and encode the information for many biochemical processes, exploiting the high specificity in the ligand-receptor interactions. Immunological response[6,7,8], tumour metastasis[9], fertilization[10] and inflammation[11] are just some examples of the processes in which carbohydrates take part. Besides they act as mediators in diverse cellular activities, such as recognition, growth and apoptosis[12,13].

1.2 Lectins: the carbohydrate binding proteins

The initial contact between two cells or a cell and a pathogenic agent, often starts with a carbohydrate-protein recognition process. The outer surface of cells shows a complex set of carbohydrates, usually belonging to oligo-, poly-saccharides, glycolipids and glycoproteins. The proteins that interact with carbohydrates are called in general carbohydrate-binding proteins (CBP). In nature, there are several examples of these kind of proteins. Some examples are carbohydrate-specific enzymes and anti-carbohydrate antibodies. A very important class of CBP that has recently raised the interest of many research groups, are lectins[14,15,16]. They represent models to study the carbohydrate-protein interactions, and valuable tools for the analysis of carbohydrates, either in the free or glycoconjugate form. Moreover, and more importantly, for their involvement in medically relevant recognition events, lectins appear as more and more attractive targets for new diagnostic and therapeutic strategies. Although they were first described at the end of the 19th century, they became of great interest only starting from the 1960s, when it was demonstrated their involvement in physiological and pathological phenomena, from cell differentiation to cancer[17]. Lectins are proteins that bind to mono- and oligo-saccharides with high specificity, reversibly and without modifying them. Therefore they do not show catalytic activity and, unlike antibodies, they are not the product of an immune response, but they decrypt the information hold in the saccharide structure to convert them in signals and give rise to biochemical processes. Lectins are ubiquitously spread in nature. They have been identified in animals, plants, and in most microorganisms, ranging from viruses to bacteria. Their localization in many cellular layers reflects the number of processes they mediate. For example membrane-bound lectins are involved in microbial adhesion and cell-cell recognition[18,19], while intracellular lectins mediate protein trafficking[20].

According to their origin, lectins are divided in microbial, plant or animal lectins[15].

One of the most studied plant lectins is viscumin (*Viscum Album Agglutinin* or VAA), extracted from mistletoe[21]. Together with, for instance, the infamous biohazard ricin, viscumin belongs to the AB-type group of ribosome inactivating proteins (RIPs)[22,23]. It presents a heterodimeric structure, consisting in two subunits, A and B, connected via a disulphide bond. The toxic A chain, formed by 254 amino acids and with a molecular weight of approximately 29 kDa, is a site-specific rRNA N-glycosidase that shuts off protein

synthesis[23]. The B chain, formed by 264 amino acids and again with a molecular mass of 29 kDa, represents indeed the proper lectin part and is responsible for binding to glycans on cell surfaces enabling the internalization of the toxic unit. Over a broad range of protein concentrations (0.01-25 mg/mL), viscumin occurs like a dimer $[AB]_2$, where the two units are held together via noncovalent interactions between the two B chains[21] (figure **1.3**). As a dimer it displays strong agglutinin activity and cytotoxicity[24,25]. Viscumin recognizes galactose-ending oligosaccharides, in both α and β stereochemistry[26]. Also the residues possibly bound to galactose can affect the binding affinity, suggesting that viscumin presents an extended binding site[27]. For example, it was found to have high affinity towards the β-D-Gal(1-2)β-D-Gal disaccharide[27] and for oligosaccharides in which the galactose moiety is α(2,6)-sialylated, in particular for glycoproteins and gangliosides with terminal sequence Neu5Ac-α(2,6)-Galβ(1-4)-GlcNAc[28].

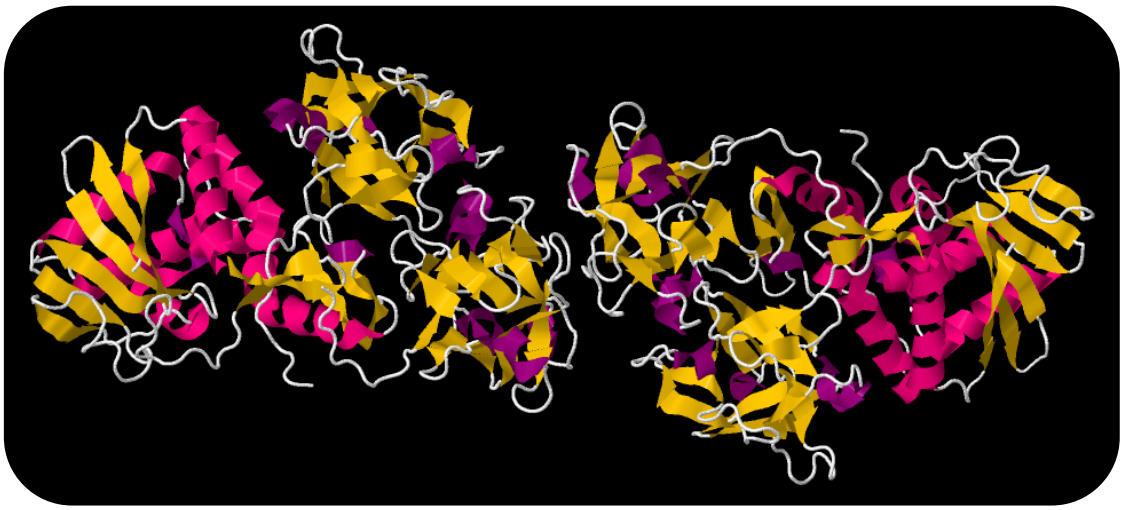


Figure 1.3 The $[AB]_2$ dimeric structure of viscumin (pdb: 1SZ6).

Animal lectins constitute a heterogeneous group of oligoproteins that differ for structure and dimensions. Initially they were classified in five groups, according to monosaccharide for which they exhibit the highest affinity: i) D-mannose, ii) D-galactose/N-acetyl-D-galactosamine, iii) N-acetyl-D-glucosamine, iv) D-glucose, v) N-acetylneuraminic acid. This classification based on monosaccharide specificity however can result misleading, since several lectins exhibit high specificity only to di-, tri- and tetrasaccharides, with binding constants up to 1000-fold higher compared to the single glycosyl moiety. Moreover some

lectins only show affinity for oligosaccharides, and their specificity can be markedly different from one oligosaccharide to another. For these reasons, nowadays their classification is preferably based on sequence homologies and evolutionary relatedness. X-ray crystallography at high resolution of lectin-carbohydrate complexes has allowed the determination of the protein Carbohydrate Recognition Domain (CRD). The current most accepted classification sees most of the lectins divided into six families: C-type, S-type, F-type, I-type, P-type and R-type lectins (table **1.1**). Some families also present a cation-dependent binding capacity, like C-type, for example, that are calcium dependent lectins.

Lectin Family	Cation dependency	Carbohydrate specificity
C-type	yes	variable (Man, Gal, Fuc)
I-type	no	variable (Man, GlcNAc, α(2,3)/α(2,6)-Sia)
S-type (Galectins)	no	β-Gal
F-type	yes	Fuc
P-type	variable	Man-6-P
R-type	no	β-Gal

Table 1.1 Lectin families.

Other families of animal lectins, which include pentraxines (pentameric subunit arrangement), ganglioside binding proteins, sulfoglucoronosyl lipid-binding proteins and others, resist to this classification because of differences in their CRD that do not allow to find sequence homologies or evolutionary relation.

A very important class of animal lectins are galectins. They recognize glycoconjugates presenting a terminal β-galactose residue, and show a conserved amino acid sequence consisting of approximately 130 residues for the carbohydrate binding domain[29]. Galectins are widely distributed in the animal kingdom, and are present in a large variety of organs and tissues, like liver, lungs and small intestine. Galectins are mainly located in the nucleus and cytoplasm, although they can be found also at cell surface or within the extracellular

matrix where they can interact with glycosylated compounds[30,31]. Considering their importance in a number of biological processes, it is not surprising that galectins have recently been under intense investigations. They take part in the cell-cell communication, adhesion and cell migration processes. They also have a role in embryonic development, differentiation and RNA splicing[32]. Moreover, galectins are involved in inflammatory processes (for example galectin-1 acts as anti-inflammatory agent[33], oppositely to galectin-3 which instead shows pro-inflammatory activity[34]), apoptosis (for example galectin-1 induces apoptosis of activated human T cells and human T leukemia cell lines[35], while galectin-3 contributes in preventing this mechanism[36]), tumor growth and metastasis[37, 38, 39]. Currently, fifteen members of the family have been identified in mammals[40]. Because of the lack of human counterparts of galectins -5, -6, -11, -14 and -15, there are only 10 types of human galectins. Based on their structure, they are divided into three subgroups: proto-type, tandem-repeat-type and chimera-type (figure **1.4**).

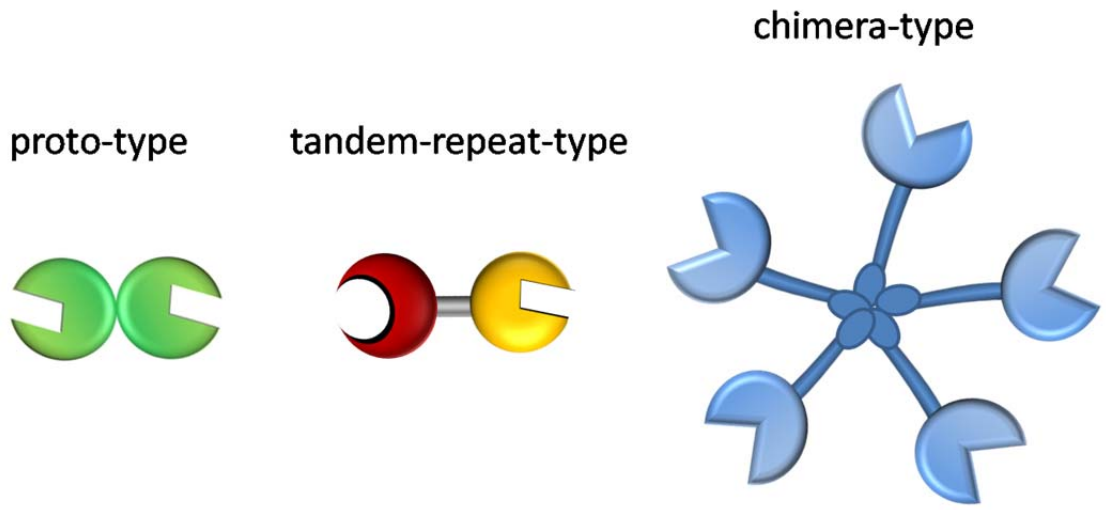


Figure 1.4 Schematic representation of the three different subgroups of galectins: proto-type, tandem-repeat-type and chimera-type.

Proto-type galectins (-1, -2, -5, -7, -10, -11, -13, -14 and -15) contain one carbohydrate recognition domain (about 15 kDa) and they can be in the form of monomers or homodimers; tandem-repeat-type galectins (-4, -6, -8, -9 and -12) present two distinct CRDs in a single peptide chain; chimera-type, whose only member is galectin-3 (about 30 kDa), is constituted by one CRD with a non lectin part directly connected to it, which is responsible

for the oligomerization[41], usually pentamerization. Due to the simultaneous presence of several CRDs, when binding to multivalent glycosylated substrates galectins can form lattices and undergo agglutination.

1.3 Multivalency

Although galectins bind saccharides with high specificity, a single carbohydrate-protein interaction shows rather weak binding affinities, with dissociation constants in the milli-micromolar range[42]. In order to enhance the binding affinity and overcome this limitation, nature exploits a powerful tool, the multivalency[43], that in case of proteins binding to carbohydrates is also named "glycoside cluster effect"[42,44]. It consists in the ability of a multivalent receptor to interact with a multimeric ligand via several, simultaneous, noncovalent recognition events (figure **1.5**) giving rise to an overall binging constant that is higher than the sum of the binding constants of the corresponding monovalent events.

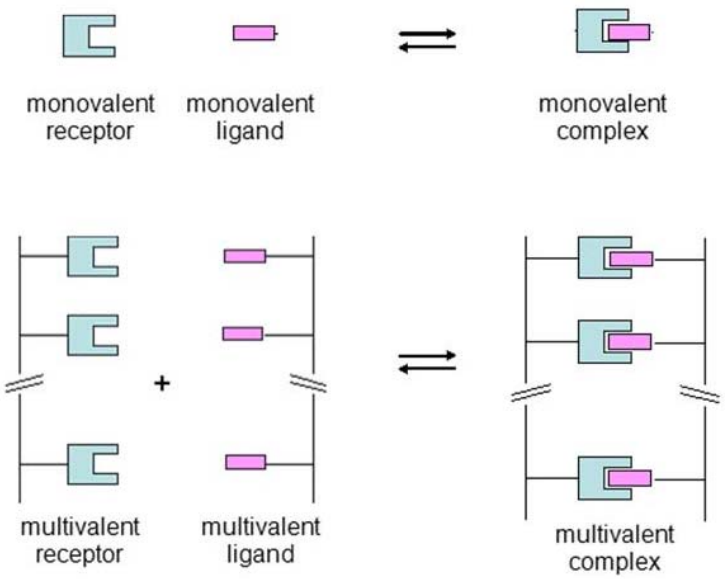


Figure 1.5 Example of monovalent *vs* multivalent complexes. Figure adapted from ref. 45.

In this way, relatively weak recognition events, as in the case of carbohydrate-protein interactions, are converted into strong and specific interactions, that also result in high thermodynamic and kinetic stability[46].

Considering a multivalent receptor and a multivalent ligand, there are several possibility of binding between the two. The first interaction possible is obviously intermolecular, but then a competition between inter- and intra-molecular binding events can take place. If the former are predominant, intermolecular aggregates of variable stoichiometry are the result of the binding, otherwise, if the latter are preponderant, the formation of a 1:1 receptor-ligand multivalent complex will be obtained (figure **1.6**).

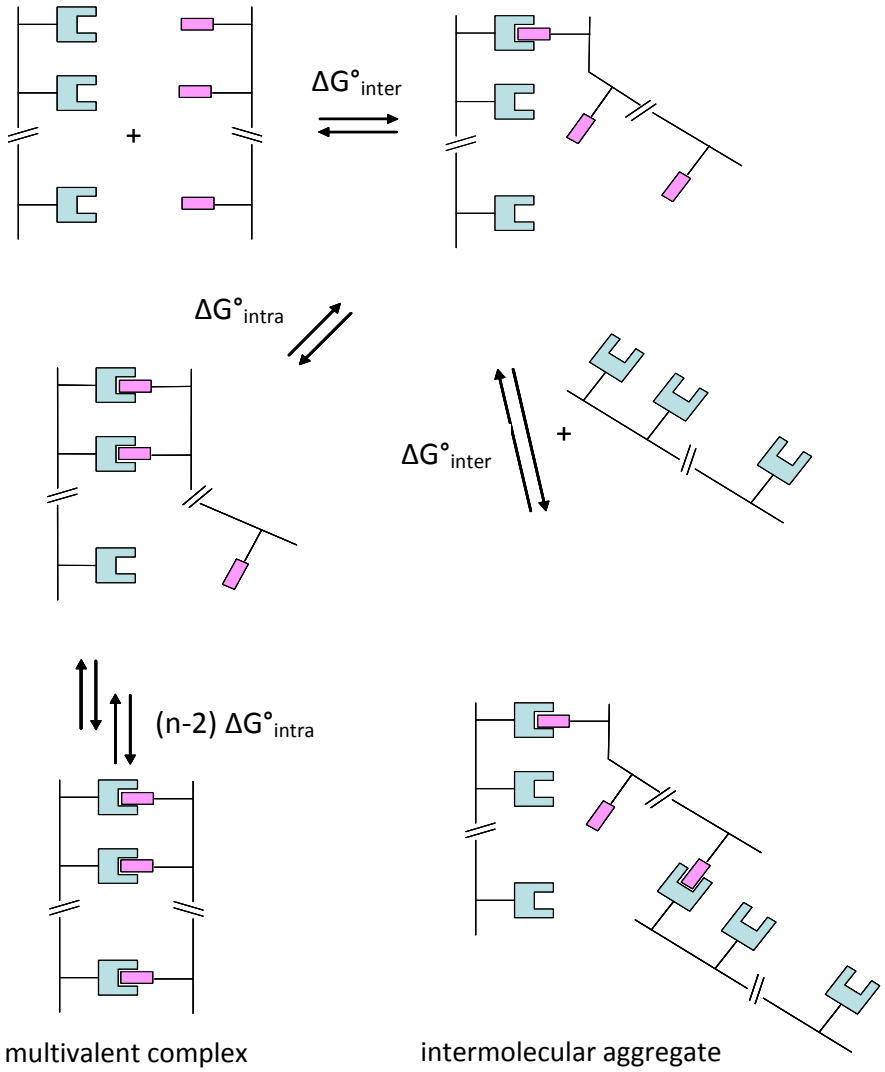


Figure 1.6 Intermolecular and intramolecular processes for the formation of a multivalent complex or of an intermolecular aggregate. Figure adapted from ref. 45.

The concept of multivalency is sometimes associated to the one of cooperativity[47], which instead describes how the binding of one ligand can influence the receptor affinity toward a further binding event. One of the most famous cases of cooperativity in biology is haemoglobin binding four oxygen molecules. In that case there is a synergistic effect, with the subsequent binding events showing greater affinity than the former ones. However, cooperativity can also be negative, when the binding is lower. As pointed out by Ercolani[48], cooperativity in multivalent interactions results extremely rare. While cooperative effects can be easily and rigorously evaluated with Hill or Scatchard coefficients, much more difficult is to prove the existence of multivalent effects.

In general, to describe multivalent binding, the approach based on the additivity of the free energies suggested by Jencks[49] can be used. The standard binding free energy of the multivalent interaction ΔG°_{multi} can be described with the following expression:

$$\Delta G^\circ_{multi} = n\Delta G^\circ_{mono} + \Delta G^\circ_{interaction}$$

where ΔG°_{mono} is the standard free energy for the corresponding monovalent interaction, *n* is the valency of the formed complex and $\Delta G^\circ_{interaction}$ is the avidity, that is to say an additional free energy component taking into account and balancing the favourable and unfavourable effects generated by the multivalent bond formation.

Whitesides et *al.*[43] have introduced the β parameter, also named "enhancement factor", which is the ratio between the two association constants K_{multi} and K_{mono} for the multivalent and monovalent binding, respectively.

$$\beta = K_{multi}/K_{mono}$$

This parameter represents an useful tool to compare in an immediate way the efficiency of ligands having different valencies and topologies. Molecules with high β values are efficient ligands, since they tend to bind the receptor in a multivalent fashion. When the number of tethering units is known, the β parameter can be normalized for the single binding event, just by dividing it to the valency *n*, giving rise to the parameter β/n.

Another method often used to evaluate the presence of multivalent effects is based on the comparison of the IC_{50} values of the multivalent ($IC_{50\text{-}multi}$) and monovalent ($IC_{50\text{-}mono}$) binding. The ratio between the two values defines a parameter called relative potency (*rp*):

$$rp = IC_{50\text{-multi}}/IC_{50\text{-mono}}$$

IC_{50}, defined as the half maximal inhibitory concentration, gives a measure of the effectiveness of a compound in inhibiting a biological o biochemical function. Also in this case, it is possible to normalize the relative potency to the valency, by dividing *rp* for the numbers of ligating units *n*.

In a more rigorous way, Kitov and Bundle[50] have developed a thermodynamic model to connect the experimentally accessible overall binding constant with the microscopic thermodynamic parameters. They adapted the expression suggested by Jencks[49] in order to take into consideration the free energies contributes of partially bound species and a statistical factor that accounts for the several equivalent ways bound forms may be achieved. The standard free energy of the multivalent interaction, now defined as $\Delta G^{\circ}_{avidity}$, is described as a function of three terms: ΔG°_{inter} (similar to the free energy of the intrinsic monovalent interaction $\Delta G°_{mono}$), ΔG°_{intra} (binding free energies for the additional intramolecular interactions of the ligand branches to the remaining binding sites on the receptor surface) and a statistical term $\Delta S^{\circ}_{avidity}$, called “avidity entropy”, which represents the probability of the association and dissociation of individual arms and it is calculated on the basis of the topology of the complex. $\Delta S^{\circ}_{avidity}$ can grow rapidly with the valency of the complex, always favouring binding of a multivalent ligand to a multivalent receptor. This also explains why multivalency can overcome the loss of conformational entropy. By a nonlinear fitting of the measured binding energies for a series of multivalent ligands, ΔG°_{inter} and ΔG°_{intra} can be determined, thus allowing to design and maximize the avidity of multivalent ligands.

1.4 Multivalent polyglycosylated molecules

The discovery of functional lectin–carbohydrate interactions and the possibility to inhibit them by using multivalent polyglycosylated molecules, offer enormous potential to the design of new drugs which could interfere with several physiological and pathological events. To achieve the formation of a determined multivalent receptor-ligand assembly, a high degree of molecular design is essential. An appropriate topological and spatial display

of ligands is in fact the prerequisite to obtain a perfect match between the receptor domains and the binding functionalities.

Multivalent ligands can be of different types, but they are usually characterized by some common features. They present a central core (scaffold) bearing several covalent connections (linkers, spacers) to the peripheral binding units (figure **1.7**). The noncovalent interactions formed between these terminal ligating units and the binding sites of the macromolecular receptor are the ones responsible for the multivalent complex formation.

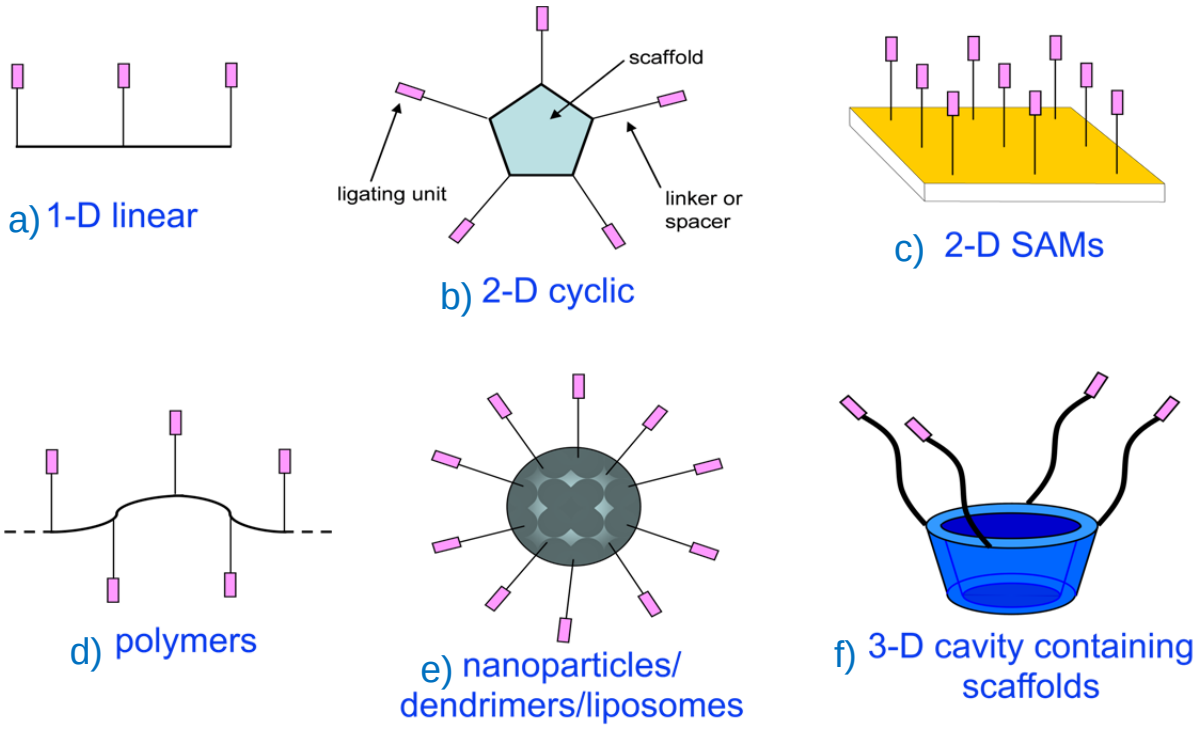


Figure 1.7 Different topologies of multivalent ligands. a) 1-D linear arrangement; b) 2-D cyclic/macrocyclic; c) 2-D self-assembled monolayers (SAM) on Au/quartz; d) polymers/peptoids; e) nanoparticles/dendrimers/ liposomes; f) 3-D cavity containing scaffolds (cyclodextrins/calixarenes). Adapted from ref. 45.

Any multivalent scaffold can in principle be used, from those having low valency such as benzene derivatives, monosaccharides, transition metal complexes, azamacrocycles, cyclodextrins or calixarenes; to high valency ones such as dendrimers, polymers, peptoids, proteins, micelles, liposomes, and self-assembled monolayers (SAMs) on nanoparticles or plane surfaces.

Once decided the core structure where to anchor the binding moieties, a fundamental parameter in the ligand design is the nature of the spacers. In fact, they can influence the type and strength of the resulting interaction with the multivalent receptor[43,51]. Since multivalency can be explained mainly through entropic effects (as described in paragraph **1.3**), a positive entropic factor, able to favor intramolecular binding, can be obtained through the use of spacers of proper length. Spacers that are too short cannot allow the simultaneous interaction of all the ligating units with the receptor sites, while spacers that are too long can lead to an useless loss of roto-translational degree of freedom. In both cases intermolecular binding would result favored instead of the multivalent complex.

1.5 Calix[n]arenes

Many polyglycosylated molecules have been prepared as multivalent ligands to study multivalent carbohydrate-protein interactions[52]. They can be classified into low valency systems, which present less than 20 carbohydrate ligands, and high valency systems, like glycoproteins, glycosylated monolayers and nanoparticles, which instead possess a larger number (even hundreds) of carbohydrate ligands.

In this contest, calixarene-based glycoclusters result very attractive[53,54]. Calix[n]arenes[55] are synthetic macrocycles derived from the condensation of para-substituted phenols and formaldehyde. The word calixarene derives from *calix* because of the resemblance of this type of molecules to a Greek vase called “calyx krater”, and from the word *arene* that refers to the aromatic building blocks forming the annulus. On these macrocycles it is possible to distinguish a lower rim, defined by the presence of the hydroxyl functionalities, and an upper rim, identified by the para-positions of the phenol units (figure **1.8**).

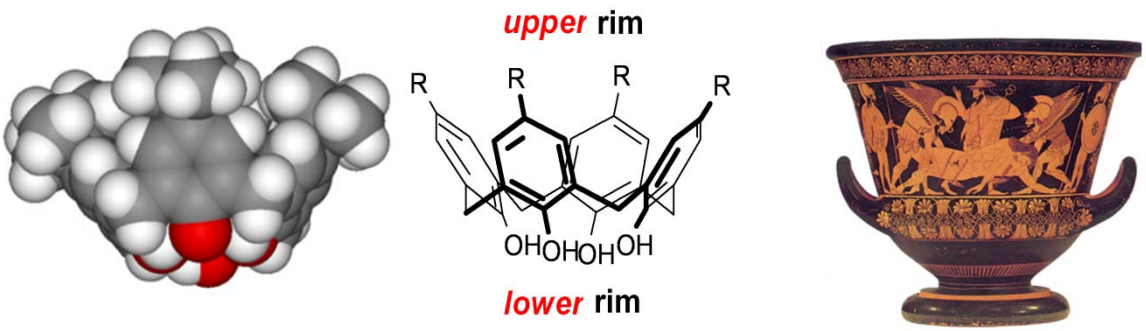


Figure 1.8 a) CPK model of a *cone*-calix[4]arene; b) molecular formula of a *cone*-calix[4]arene; c) "calix krater" Greek vase, whose shape has suggested the name "calix" for this class of macrocycles.

The shape, size (figure **1.9**) and conformational properties of these molecules can be fine-tuned by varying the procedures used for their synthesis and functionalization. For example, calix[4]arenes can be obtained at high temperature in presence of NaOH, while calix[8]arenes are formed in the same conditions but at lower temperature. These data suggested the hypothesis that the tetramer is the thermodynamic while the octamer is the kinetic product under these reaction conditions. On the contrary, for the synthesis of calix[6]arene the templating effect of the bigger potassium cation is exploited. While even numbered calix[n]arenes (n = 4, 6, 8) can be obtained with good yields, especially from phenols bearing bulky substituents at para-position, odd numbered calix[n]arenes (n = 5, 7, 9) are obtained in considerably lower yields.

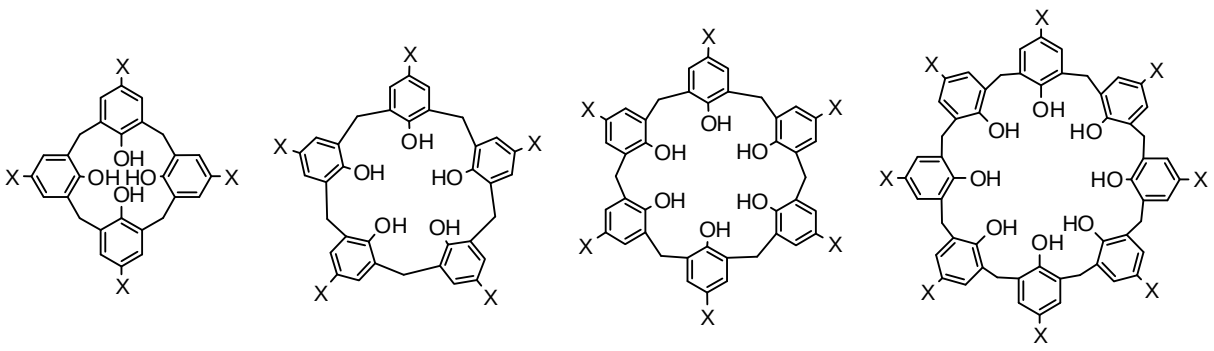


Figure 1.9 Different sizes of calix[n]arenes.

Calix[n]arenes with n > 4 show high conformation mobility in solutions. On the other side, calix[4]arenes can be fixed in one of the four different limit conformations by functionalizing the lower rim with bulky enough substituents, typically larger than the ethyl group. These conformations are named *cone, partial cone*, *1,3-alternate* and *1,2 alternate* (figure **1.10**), which show very different molecular recognition properties[56].

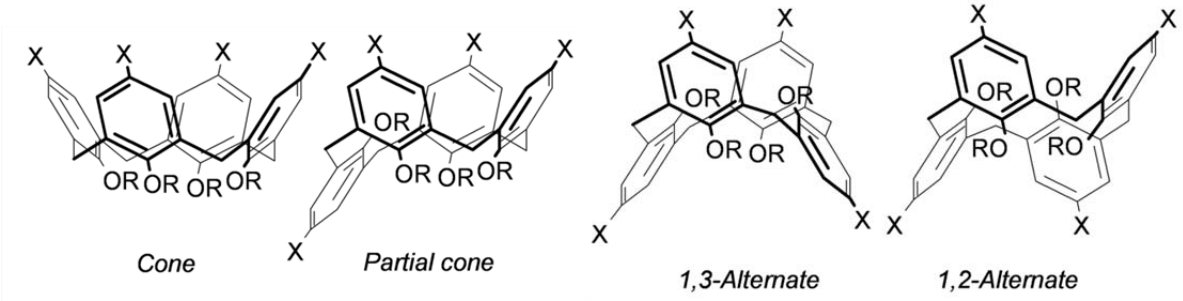


Figure 1.10 The four limiting conformations of calix[4]arenes, when R > Et.

Calixarenes can be functionalized both at the upper and at the lower rim, also with significant selectivity among the equivalent positions. For all these characteristics, this class of macrocycles results particularly attractive as scaffolds also for linking glycoside units and obtain multivalent systems. Their derivatives that are functionalized with carbohydrates at the upper or lower rim are named glycocalixarenes[45,53,54,57].

These neoglycoconjugates have attracted attention also because they can be potentially used for site-directed drug delivery (figure **1.11**), by combining the known inclusion properties of calixarenes with the targeting ability furnished by the saccharide units. For calix[4]arenes blocked in the cone conformation, for instance, the choice to anchor the carbohydrate units to the upper rim allows to project all the glycosylated residues in the same direction. In this way the interaction sites of the protein would result in close proximity to the hydrophobic calixarene cavity that can complex guests and therefore potentially give rise to site-specific delivery of drugs. For example, in the case of tumor cells that overexpress galectins, the complexed drug could be carried selectively only to the molecular target, reducing in this way the doses and thus the chemotherapeutic side-effects.

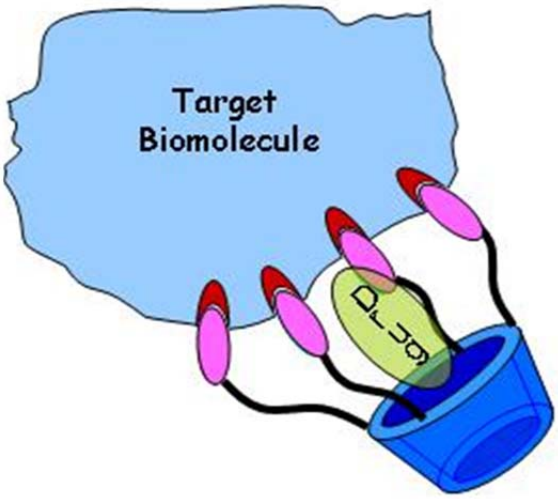


Figure 1.11 Site-specific delivery of drugs to protein by glycocalix[4]arenes that present *cone*-conformation.

Ungaro and co-workers have already prepared a series of glycosylthioureido calix[4]arenes which also shown to be able to efficiently and selectively bind the plant lectin Concavalin A and to complex small organic and anionic guests[58].

Moreover, calixarenes have found interesting applications in bioorganic chemistry. Upon proper derivatization, some of these compounds have been used, for example, as contrast agents in magnetic resonance imaging[59], as enzyme mimics[60], or as ligands able to interact with and recognize DNA[61] and perform gene transfection[62].

1.6 References

1 Varki A. *Glycobiology,* **1993**, *3*, 97-130.

2 *The Sugar Code. Fundamentals of glycosciences*, ed. H.-J. Gabius, Wiley-VCH, Weinheim, **2009**.

3 Adapted from F. Stoddart's web page http://www.chem.ucla.edu/dept/Faculty/stoddart/research/multivalency.htm

4 Gabius, H. J. *Naturwissenschaften* **2000**, *87*, 108-121.

5 Laine, R. A. *Glycobiology*, **1994**, *4*, 759-767.

6 Buzas, E. I.; Gyorgy, B.; Pasztoi, M.; Jelinek, I.; Falus, A.; Gabius, H. J. *Autoimmunity*, **2006**, *39* (8), 691-704.

7 Dwek, R. A. *Chem. Rev.*, **1996**, *96* (2), 683-720.

8 Calarese, D. A.; Scanlan, C. N.; Zwick, M. B.; Deechongkit, S.; Mimura, Y.; Kunert, R.; Zhu. P.; Wormald, M. R.; Stanfield, R. L.; Roux, K. H.; Kelly, J. W.; Rudd, P. M.; Dwek, R. A.; Katinger, H.; Burton, D. R.; Wilson, I. A. *Science*, **2003**, *300* (5628), 2065-2071.

9 Lau, K. S.; Dennis, J. W. *Glycobiology*, **2008**, *18* (10), 750-760.

10 Snell, W. J.; White, J. M. *Cell*, **1996**, *85* (5), 629-637.

11 Rudd, P. M.; Elliott, T.; Cresswll, P.; Wilson, I. A.; Dwek, R. A. *Science,* **2001**, *291* (5512), 2370-2376.

12 Lis, H.; Sharon, N. *Chem. Rev.*, **1998**, *98*, 637.

13 Simanek, E. E.; McGarvey, G. J.; Jablonowski, J. A.; Wong C. H. *Chem. Rev*., **1998**, *98*, 833.

14 Lis, H.; Sharon, N. *Acc. Chem. Res.* ***1998***, *98*, 637-674.

15 Gabius, H. J. *Eur. J. Biol.* **1997**, *243*, 543-576.

16 Gabius H. J. *Eukaryotic Glycosylation and Lectins: Hardware of the Sugar Code (Glycocode) in Biological Information Transfer*, Eurekah, **2001**.

17 Lis, H.; Sharon, N. *The Lectins: Properties, Functions and Applications in Biology and Medicine*; Liener, I. E.; Sharon, N.; Goldstein, I. J.; Eds. Academic Press, Inc. Orlando, **1986**, p. 293.

18 Gizken, K.; Cambi, A.; Torensma, R.; Frigdor, C. G.; *Curr. Prot. Pep. Sci.*, **2006**, *7*, 283.

19 Sharon, N.; *Biochimica et Biophysica Acta (BBA) – General Subjects*, **2006**, 1760, 527.

20 Delacour, D.; Koch, A.; Jacob, R. *Traffic*, **2009**, *10*, 1405-1413.

21 Olsnes, S.; Stirpe, F.; Sandvig, K.; Pihl, A.; *J. Biol. Chem.*, **1982**, *257*, 13263.

22 Olsnes, S.; Kozlov, J. V.; *Toxicon*, **2001**, *39*, 1723.

23 Endo, Y. *Adv. Lectin Res.*, **1989**, *2*, 60-73.

24 Andrè, S.; Pieters, R. J.; Vrasidas, I.; Kaltner, H.; Kuwabare, I.; Liu, F. T.; Liskamp, R. M.; Gabius, H. J. *Chembiochem*, **2001**, *2*, 822.

25 Gabius, H. J.; Darro, F.; Remmelink, M.; Andrè, S.; Kopitz, J.; Danguy, A.; Gabius, S.; Salmon, S.; Kiss, R.; *Cancer Investigation*, **2001**, *19* (2), 114-126.

26 Galanina, O. E.; Kaltner, H.; Khraltsova, L. S.; Bovin, N. V.; Gabius, H. J. *J. Mol. Recognit.*, **1997**, *10*, 139-147.

27 Lee, R. T.; Gabius, H. J.; Lee, Y. C. *Carbohydr. Res.*, **1994**, *254*, 269-276.

28 Muthing, J.; Meisen, I.; Bulau, P.; Langer, M.; Witthohn, K.; Lentzen, H.; Neumann, U.; Peter-Katalinic, J. *Biochemistry*, **2004**, *43*(11), 2996-3007.

29 Barondes, S. H.; Cooper, D. N. W.; Gitt, M. A.; Leffler, H. *Journal of Biological Chemistry*, **1994**, *269*(33), 20807-20810.

30 Hugues, R. C. *Biochem. Biophys. Acta*, **1999**, *1473*, 172-185.

31 Leffer, H. *Results. Probl. Cell. Differ.*, **2001**, *33*, 57-83.

32 Perillo, N. L.; Marcus, M. E.; Baum, L. G. *J. Mol. Med*., **1998**, *7*6, 402-412.

33 Rabinovich, G. A. *Scand. J. Immunol.* **2007**, *66*, 143.

34 Liu, F. T.; Hsu, D. K. *Drug News Perspect*, **2007**, *20*, 455.

35 Perillo, N. L.; Pace, K. E.; Seilhamer, J. J.; Baum, L. G. *Nature*, **1995**, *378*, 736-739.

36 Yang, R. Y.; Hsu, D. K.; Liu, F. T. *Proc. Natl. Acad. Sci. U.S.A.*, **1996**, *93*, 6737-6742.

37 Huflejt, M.; Leffler, H. *Glycoconjugate J.*, **2003**, *20*, 247-255.

38 Takenaka, Y.; Fukumori, T.; Raz, A. *Glycoconjugate J.,* **2002**, *19*, 543-549.

39 Lahm, H.; André, S.; Hoeflich, A.; Kaltner, H.; Siebert, H. C.; Sordat, B.; von der Lieth, C. W.; Wolf, E.; Gabius, H. J. *Glycoconjugate J.*, **2003**, *20*, 227-238.

40 Varki, A.; Cummings, R. D.; Esko, J. D.; et al. *Essentials of Glycobiology. 2nd edition*. **2009**, Eds. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press.

41 Leffler, H.; Carlsson, S.; Hedlund, M.; Qian, Y.; Poirier, F. *Glycoconjugate Journal*, **2004**, *19*, 433-440.

42 Lee, Y. C.; Lee R. T. *Acc. Chem. Res.* **1995**, *28*, 321-327.

43 Mammen, M.; Choi, S. K.; Whitesides, G. M. *Angew. Chem. Int. Ed. Engl.* **1998**, *37*, 2754-2794.

44 Lundquist, J. J.; Toone, E. J. *Chem. Rev.* **2002**, *102*, 555-578.

45 Baldini, L.; Casnati, A.; Sansone, F.; Ungaro, R. *Chem. Soc. Rev.*, **2007**, *36*, 254-266.

46 Mulder, A.; Huskens, J.; Reinhoudt, D. N. *Org. Biomol. Chem.*, **2004**, *2*, 3409-3424.

47 Badjić, J. D.; Nelson, A.; Cantrill, S.; Turnbull, B.; Stoddart, F. *Acc. Chem. Res.*, **2005**, *38*, 723-732.

48 Ercolani, G. *J. Am. Chem. Soc.,* **2003**, *125*, 16097-16103.

49 Jencks, W. P. *Proc. Natl. Acad. Sci. U. S. A.* **1981**, *78*, 4046-4050.

50 Kitov, P. I; Bundle, D. R. *J. Am. Chem. Soc.* **2003**, *125*, 16271-16284.

51 Rao, J.; Lahiri, J.; Weis, R. M.; Whitesides, G. M. *Angew. Chem. Int. Ed.*, **2000**, *122*, 2698-2710.

52 Roy, R. *Top Curr Chem* **1997**, *187*, 241-274.

53 Sansone, F.; Rispoli, G.; Casnati, A.; Ungaro, R. *Multivalent Glycocalixarenes, in "Synthesis and Biological Applications of Glycoconjugates"*, Eds. O. Renaudet and N. Spinelli, Bentham Science Publishers, **2011**, p. 36-63. doi: 10.2174/978160805277611101010036.

54 Dondoni, A.; Marra, A. *Chem. Rev.*, **2010**, *110* (9), 4949-4977.

55 a) Gutsche, C. D.; Stoddart J. F. *Calixarenes Revisited*, Ed. The Royal Society of Chemistry, Cambridge, UK, **1998**. b) Mandolini, L.; Ungaro, R. *Calixarenes in Action*, Eds. Imperial College Press, London, UK, **2000**. c) Asfari, Z.; Böhmer, V.; Harrowfield, J.; Vicens, J. *Calixarenes 2001*; Eds. Kluwer Academic Publishers, Dordrecht, The Netherlands, **2001.**

56 Andrè, S.; Sansone, F.; Kaltner, H.; Casnati, A.; Kopitz, J.; Gabius, H. J.; Ungaro, R. *ChemBioChem*, **2008**, *9*, 1649-1661.

57 Casnati, A.; Sansone, F.; Ungaro, R. *Acc. Chem. Res.* **2003**, *36*, 246-254.

58 a) Sansone, F.; Chierici, E.; Casnati, A.; Ungaro, R. *Org. Biomol. Chem.* **2003**, *1*, 1802-1809; b) Torvinen, M.; Neitola, R.; Sansone, F.; Baldini, L.; Ungaro, R.; Casnati, A.; Vainiotalo, P.; Kalenius, E. *Org. Biomol. Chem.* **2010**, *8*, 906-915; c) Torvinen, M.; Sansone, F.; Casnati, A.; Kalenius, E.; Jänis, J. *J. Mass Spectr.* **2011**, *46*, 787-793; d) Torvinen, M.; Kalenius, E.; Sansone, F.; Casnati, A.; Jänis, J. *J. Am. Soc. Mass Spectr.* **2012**, *23*, 359-365.

59 a) Aime, S.; Barge, A.; Botta, M.; Casnati, A.; Fragai, M.; Luchinat, C.; Ungaro, R. *Angew. Chem. Int. Ed*. **2001**, *40*, 4737-4739; b) Bryant, L. H.; Yordanov, A. T.; Linnoila, J. J.; Brechbiel, M. W.; Frank, J. A. *Angew. Chem. Int. Ed.* **2000**, *39*, 1641-1643.

60 a) Rondlez, Y.; Bertho, G.; Reinaud, O. *Angew. Chem., Int. Ed.* **2002**, *41*, 1044-1046; b) Izzet, G.; Douziech, E.; Prange, T.; Tomas, A.; Le, Mest, Y.; Reinaud, O. *Proc. Natl. Acad. Sci., U.S.A.* **2005**, *102*, 6831-6836; c) Cacciapaglia, R.; Casnati, A.; Mandolini, L.; Peracchi, A.; Reinhoudt, D. N.; Salvio, R.; Sartori, A.; Ungaro, R. *J. Am. Chem. Soc.* **2007**, *129*, 12512-12520.

61 a) Zadmard, R.; Schrader, T. *Angew. Chem. Int. Ed.* **2006**, *45*, 2703-2706; b) Schrader, T.; Zadmard, R.; Breitkreuz, C. *Supramolecular Chemistry* **2008**, *20*, *109-115*.

62 a) Lalor, R.; DiGesso, J. L.; Mueller, A.; Matthews, S. E. *Chem. Commun*. **2007**, 4907-4909; b) Horiuchi, S.; Aoyama, Y. *J. Contr. Release* **2006**, *116*, 107-114.

Chapter 2

Synthesis of new glycosyl-thioureido-calixarenes for galectin inhibition

Abstract

To achieve highly selective and strong interactions of a glycocluster for a given carbohydrate binding protein, it is important to combine a proper multivalent presentation of the single carbohydrates with a high affinity glycosyl structure. The work reported in this chapter concerns the synthesis of different glycocalixarenes containing units of LacNAc or LacNAc modified at 2-N or 3' positions. The tetra-NAc-lactosylamino-thioureidocalixarene (*cone*-4LacNAc[4]Prop) was synthesized through a classical chemical pathway and also by a chemo-enzymatic synthesis involving a galactosyl transferase and a glucosylcalixarene, thus demonstrating that also calixarenes can be used as substrates for enzymes. Enzymatic sialylation reactions were also explored by using the 2,3- and 2,6-sialyl transferases on *cone*-4Lac[4]Prop giving, however, mixtures of products of partial glycosylation. Some of these glycoclusters interact very efficiently with medically relevant lectins. The selectivity and efficiency in the inhibition of different human galectins is also discussed in terms of the valency and stereochemical presentation of the saccharide units in the glycocluster.

2.1 Introduction

The growing knowledge of the biological role of lectins and of the glycoside cluster effect allows to identify new attractive targets for drug design. Since glycan recognition is regulated by the structure of the sugar epitope and also by the topological aspects of its presentation, a suitable arrangement of the binding groups in synthetic glycoclusters can have the potential to enhance avidity and selectivity of these neoglycoconjugate ligands. The achievement of these results could disclose to such compounds the possibility of finding important medical applications. This is why our research group has been working for several years on the design and synthesis of glycocalixarenes potentially able to interact with lectins of clinical interest, such as the VAA (*Viscum Album Agglutinin*) plant toxin, that acts as potent biohazard, and human galectins, that act as factor in tumor progression.

An interesting class of glycoclusters, that showed to be very effective in the interaction with human galectins, is represented by the lactosyl-thioureido-calix[4]arenes represented in figure **2.1**[1].

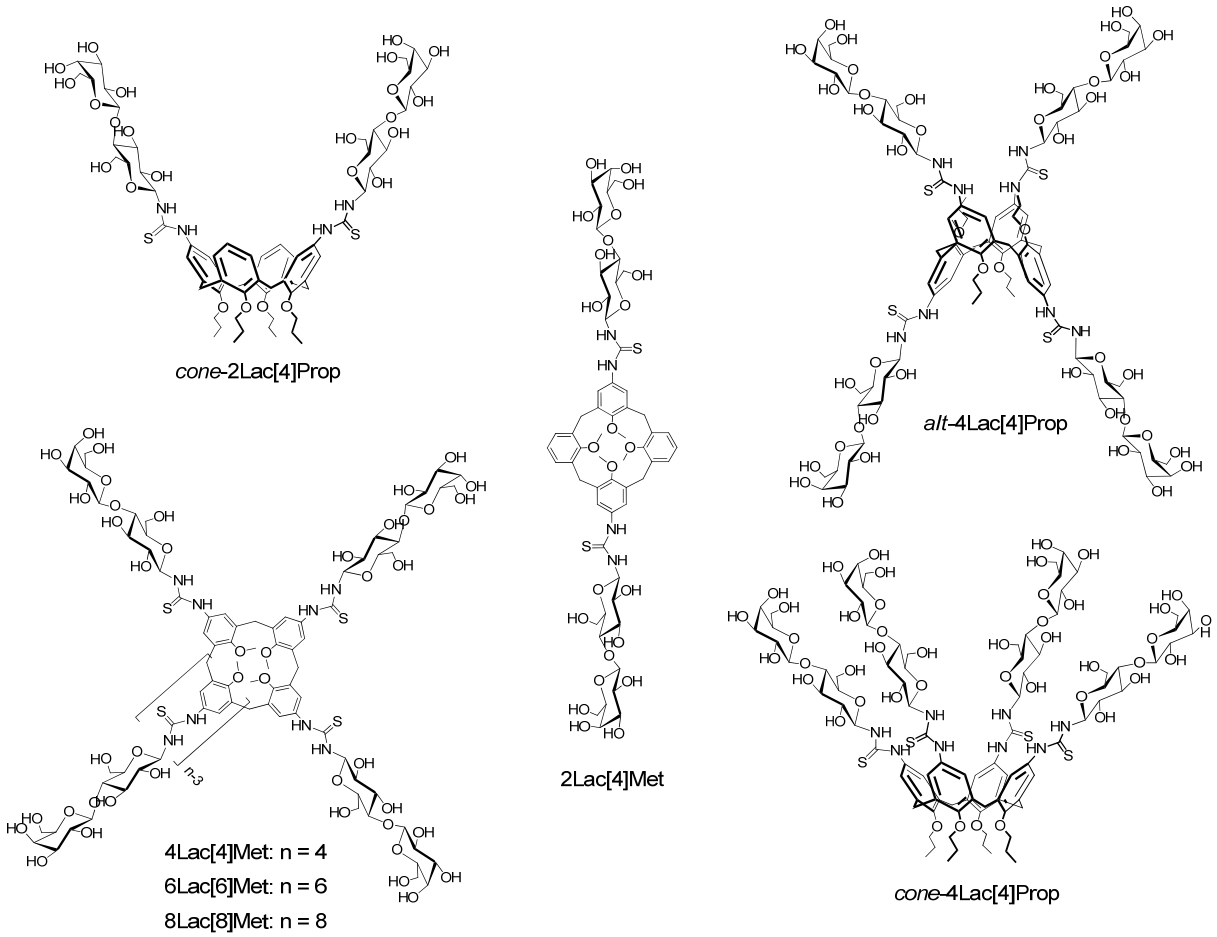


Figure 2.1 Examples of lactosyl-thioureido calix[4]arenes.

Cell-binding inhibition assays performed by Prof. Hans-Joachim Gabius and coworkers at Ludwig Maximilian University in Münich (Germany), demonstrated that these compounds present the capacity to inhibit the adhesion of galectins to the surface of tumor cells, and therefore they represent good candidates for antiadhesion therapies. For these studies selected human tumor lines were used. Binding of biotinylated galectins to these cell lines was determined by quantitative fluorescence detection through a Fluorescence Activated Cell Sorter by using streptavidin/R-phycoerythrin as indicator. Examples of the graphics obtained from these measurements are reported in figures **2.2** and **2.3**. The binding of labeled galectins to cells shifts the signal from the position of the control curve (gray area) to the right (solid black line) due to an increase of the fluorescence. An inhibitor able to interfere with galectins and to prevent their binding to cells moves the fluorescent profile back towards the curve relative to the control. For each curve the percentage of positive cells and the mean fluorescence intensity are displayed. Interestingly, these glycoclusters

present the ability to selectively recognize different types of galectins. For example, among all the tested lactosyl-thioureido-calix[4]arenes, the tetralactosylcalix[4]arene in the 1,3-alternate structure (*alt*-4-Lac[4]Prop in figure **2.1**) is a potent inhibitor for Gal-1 (figure **2.2**), while it is the worst inhibitor for Gal-3 (figure **2.3**). On the contrary, the *cone*-tetralactosylcalix[4]arene (*cone*-4Lac[4]Prop in figure **2.1**) is highly effective in the inhibition of Gal-3 (figure **2.3**), but shows very low affinity for Gal-1 (figure **2.2**).

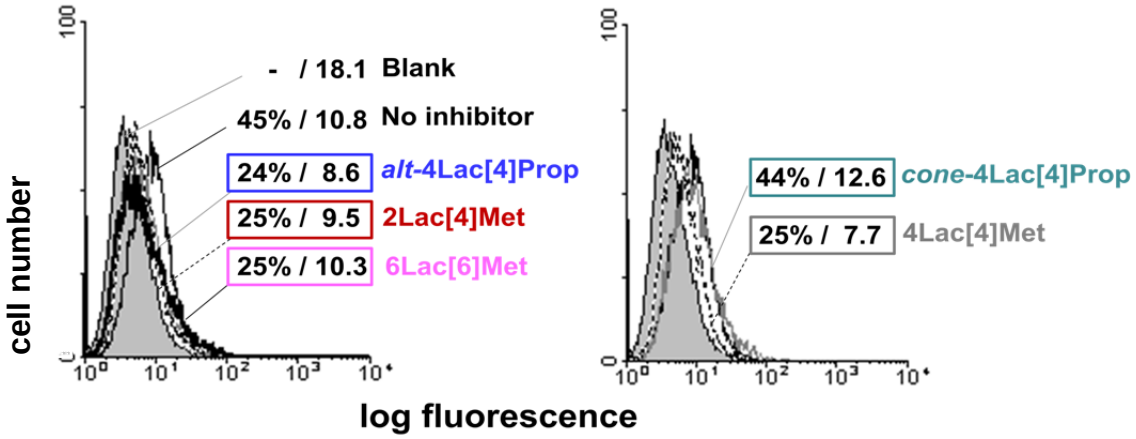


Figure 2.2 Inhibition on Gal-1 (10 µg/mL) binding to human pancreatic carcinoma cells (sugar concentration 2 mM) of some lactosyl-thioureido-calix[n]arenes. *Alt*-4Lac[4]Prop was found to be the best inhibitor for Gal-1.

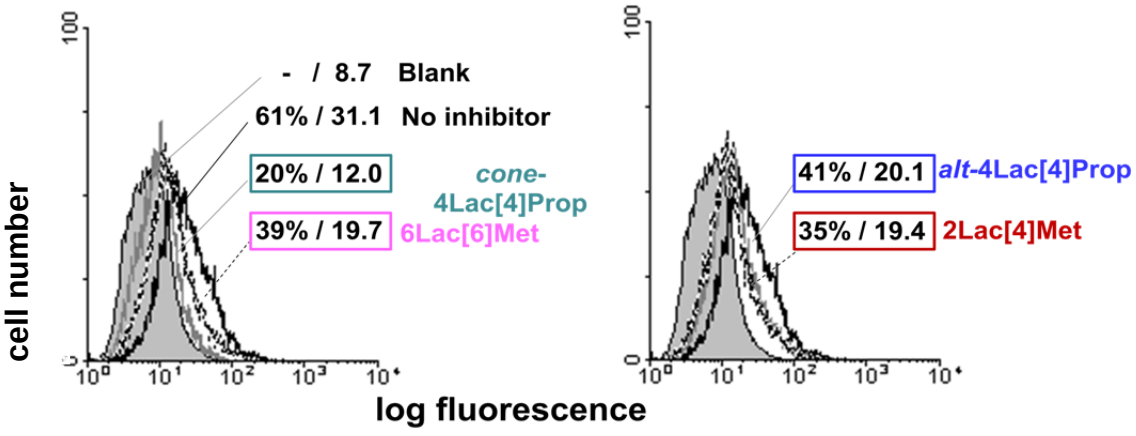


Figure 2.3 Inhibition on Gal-3 (10 µg/mL) binding to human colon adenocarcinoma cells (sugar concentration 0.1 mM) of some lactosyl-thioureido-calix[n]arenes. *Cone*-4Lac[4]Prop was found to be the best inhibitor for Gal-3.

Similar experiments were performed with the corresponding galactosyl-thioureido-calix[n]arenes, but these compounds showed to be less effective than the lacto-clusters. These bioassays pointed out the existence of a strict relationship between the conformational and structural features of the different calixarenes and their inhibitory activity. Therefore, these glycoclusters present the possibility to selectively target the different galectins. This is important because, depending on the pathology, different types of these lectins can be involved.

It is in this context that a relevant part of the present thesis work is placed. The aim is to obtain new glycoclusters based on a calixarene scaffolds and able to inhibit galectins with higher affinity and higher selectivity compared to the analogous lactosyl-calixarenes already tested and mentioned above. We planned to achieve this goal by using carbohydrates that, compared to lactose, possess a higher complementarity and therefore affinity with the binding site of the target lectins.

It is known for example that N-acetyl-lactosamine (Gal-β(1-4)GlcNAc or LacNAc) interacts with some lectins very efficiently[2,3] while a series of studies have been reported describing the important affinity of these lectins for lactosides and N-acetyl-lactosamines modified at 2-N and 3' positions with aromatic groups[4,5]. It was proposed that stabilizing cation-π, CH-π and hydrophobic interactions between these substituents and suitable residues in the recognition sites of the proteins improves the binding. On the other hand, also the presence of sialic acid at 3' or 6' position of lactose or N-acetyl-lactosamine residues should increase the efficiency and selectivity of the recognition process with (ga)lectins, being these trisaccharides the complete epitope that some lectins target on the cell surface.

However, the difficulty in the synthesis of these modified lactosides and trisaccharides together with the complexity of their conjugation to the calixarene scaffold make the preparation of the corresponding glycoclusters an important goal to be pursued in this thesis.

2.2 Results and discussion

In this chapter, the synthesis of calixarenes functionalized with LacNAc or modified LacNAc units is presented and discussed, including the first attempts of enzymatic sialylation of

lactosyl-calixarenes. Even if the sugar structure varies, some important features are kept constant in the design of the cluster molecules (figure **2.4**) with respect to the previously studied lactosyl-thioureido clusters. First of all, the calixarene scaffold constitutes again the cyclic platform where to covalently anchor the sugar moieties. The macrocyclic structures here reported have a ring size of 4 or 6 phenolic units. In the first case, calix[4]arenes in the cone conformation, blocked at the lower rim with propyl chains, were synthesized. Instead, for calix[6]arenes the lower rim was functionalized with methyl group, in order to favor a high conformational flexibility of the molecules. Again, the thiourea unit was chosen as proper spacer to connect the calixarene backbone to the carbohydrate unit, since it can be easily obtained by "click" reaction of an amine with an isothiocyanate giving very high yields and absence of by-products. Moreover this kind of conjugation does not involve bond formation at the anomeric center of the glycoside and thus preventing stereochemistry problems during the glycosylation reaction. In addition, the hydrogen atoms of the two urea NH groups allow further potential hydrogen bonding that could increase the binding with the lectins or help the inclusion of pharmacophore molecules[6,7]. As in the lactosyl-calixarenes, the sugar moieties connect the thiourea linkers through β anomeric bonds.

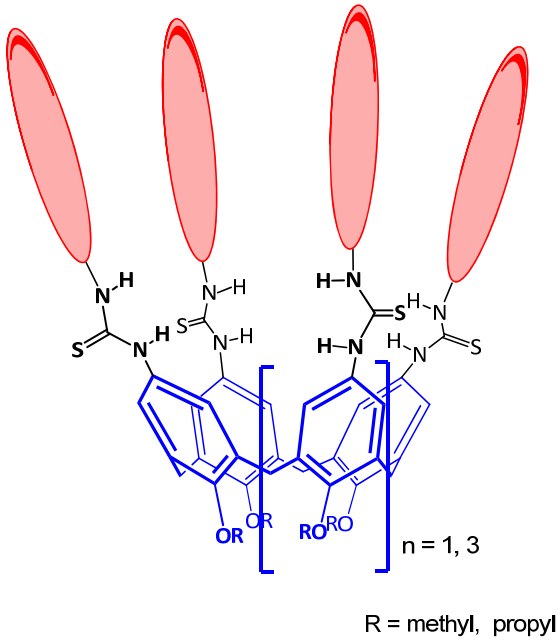


Figure 2.4 Schematic view of glycosy-thioureido-calixarenes.

In order to easily identify these compounds, whose IUPAC nomenclature results to be very long and complicated, in the thesis the running code *conf*-mS[n]Alk is used, as for lactosylcalixarenes in figure **2.1**. For each compound are indicated: the conformation *conf* of the calixarene when well defined (cone or 1,3-alternate), the number m and the type of saccharide S, the size [n] of the macrocycle and the alkyl chain Alk corresponding to the substituent at the phenolic oxygen.

2.2.1 Synthesis of *cone*-4LacNAc[4]Prop calixarene

In order to enhance the selectivity of the cluster-lectin recognition process, *cone*-4LacNAc[4]Prop calixarene **1** was synthesized (figure **2.5**). This molecule is blocked in the cone conformation and functionalized at the upper rim with four units of N-acetyl-lactosamine. LacNAc is well known to be a biologically important disaccharide which has been found in glycoproteins and glycolipids and it is widely distributed in different human and animal tissues[8,9]. It was also found that LacNAc can efficiently inhibit galectins in a more potent way compared to galactosides[3] and for this reason we focused on the synthesis of calixarenes functionalized with this sugar moiety. To obtain this result two pathways were explored: the first one, described in the next paragraph **2.2.1.1**, follows a classical chemical synthesis; the second one, described in paragraph **2.2.1.2**, combines chemical reactions with enzymatic methodologies (figure **2.5**).

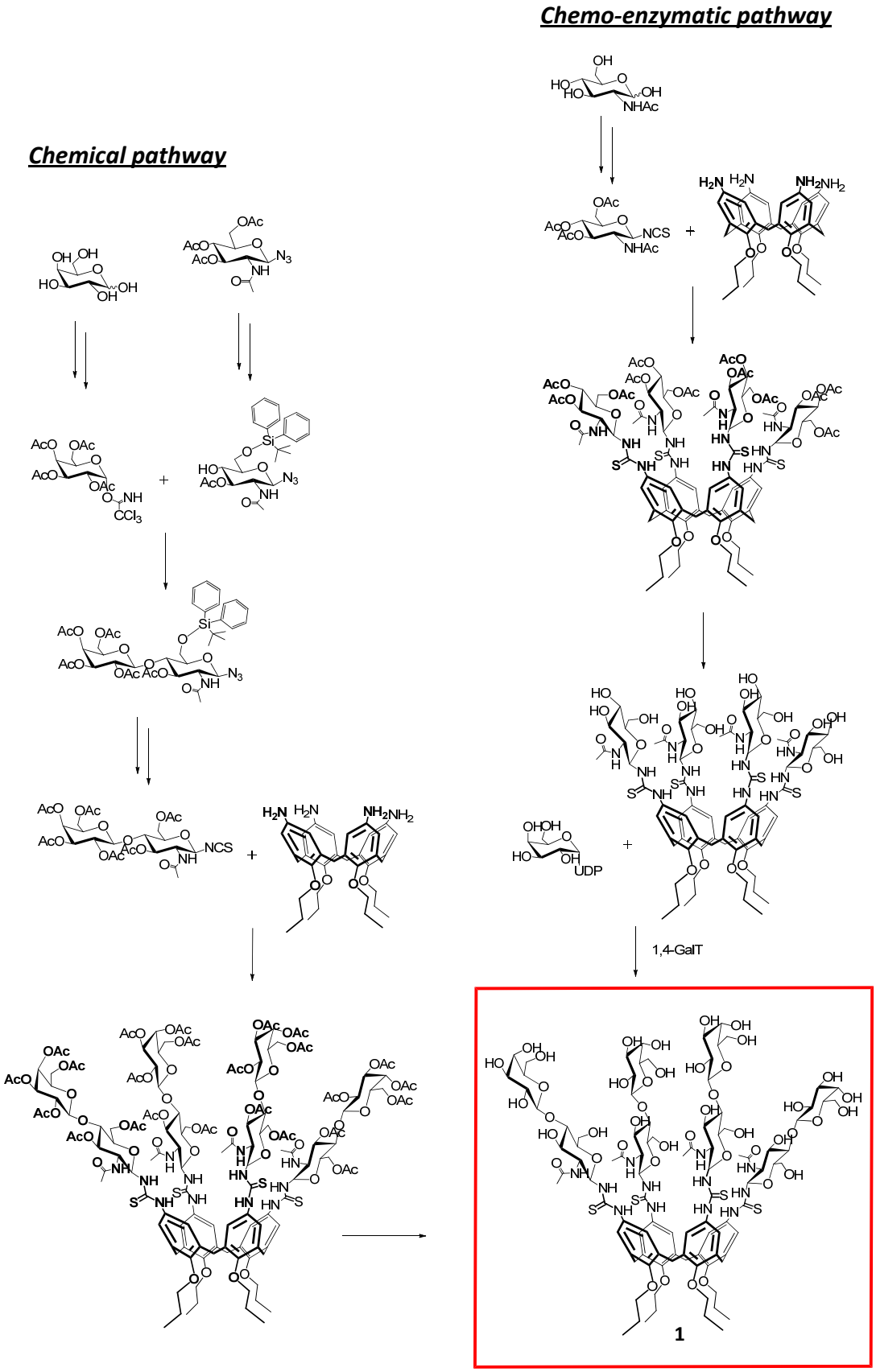


Figure 2.5 Synthetic strategies to obtain the target molecule *cone*-4LacNAc[4]Prop calixarene **1** via the chemical (left) and the chemoenzymatic (right) approach.

2.2.1.1 Chemical synthesis

To synthesize compound **1** using a fully chemical route, the used approach was to separately prepare the β-anomer of LacNAc functionalized at the anomeric position with an isothiocyanate group and the tetraamine calixarene scaffold. Subsequently, they were conjugated one to the other through a "click" reaction originating a thiourea unit. A suitable LacNAc derivative was prepared through a glycosylation reaction (figure **2.10**) between the galactosyl donor **4** and the N-acetyl-glucosamine acceptor **11**. The latter compound bears, in anomeric position, an azido group that can be converted into an isothiocyanate for the further conjugation to the aminocalixarene **20**. Both monosaccharides were properly synthesized in order to have the glycosyl donor **4** with the anomeric position activated and the glycosyl acceptor **11** only with the hydroxyl at 4 position deprotected. This allows the exclusive formation of a 1-4 bond between the two monosaccharides.

To obtain galactosyl donor **4** (figure **2.6**) the first step was the peracetylation of D-galactose to give compound **2**, which was then selectively deprotected only at the anomeric position thanks to the use of hydrazine acetate in DMF[10]. In this way, compound **3** was obtained as a mixture of α and β anomers that was used without purification in the following step. Reaction of compound **3** with trichloro acetonitrile led to galactosyl donor **4**, activated in anomeric position thanks to the introduction of the trichloro acetimidate group. Trichloro acetimidate can be selectively inserted in α or β position according to the reaction conditions[11] and therefore it is possible to direct the subsequent S_N2 O-glycosylation and selectively obtain β and α products, respectively. The α-orientation of this activating group in compound **4** was assessed via ^1H NMR spectroscopy (doublet at 6.60 ppm with coupling constant of 3.3 Hz for the anomeric proton H_1).

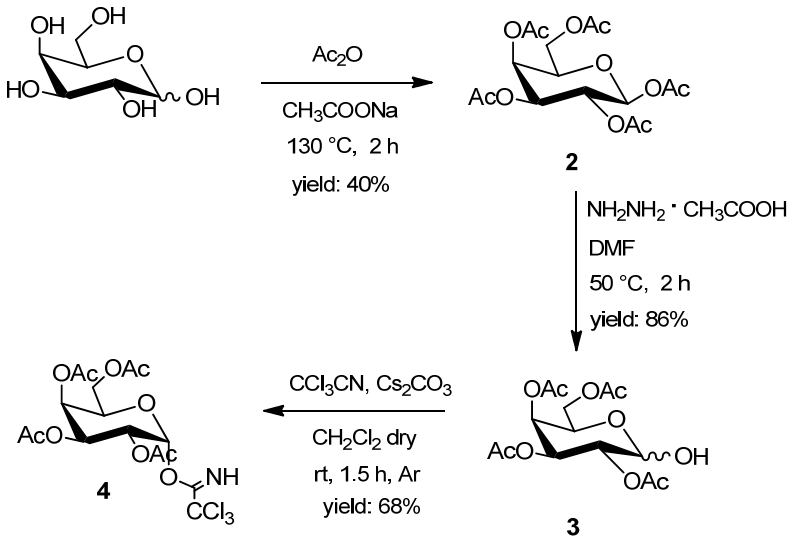


Figure 2.6 Synthesis of the galactosyl donor **4**.

In a first attempt, we planned to use the glucosylamine acceptor **6** (figure **2.7**) bearing two free hydroxyl groups for the glycosylation reaction. Gan and coworkers[12], in fact, reported that it is possible to get a regiospecific 1→4 glycosylation when 6-*O*-*tert*-butyldiphenylsilyl-*N*-acetylglucosamine derivative **6** is used, despites it bears two free hydroxy groups as potential acceptor centers. The regiospecificity was attributed to the presence of the bulky *tert*-butyldiphenylsilyl protecting group that hinders the top side of saccharide **6** disfavoring the approach of the glycosyl donor to the 3-OH position. Starting from the commercially available 2-acetamido-2-deoxy-3,4,6-tri-*O*-acetyl-2-β-D-glucopyranosyl azide, the hydroxyl groups could be removed by using the Zemplén method. The following reaction of compound **5** with *tert*-butyl(chloro)diphenyl silane (TBDPSiCl) in pyridine allowed the regioselective protection of the highly nucleophilic primary hydroxy group at 6 position of glucosamine[12] to get compound **6**.

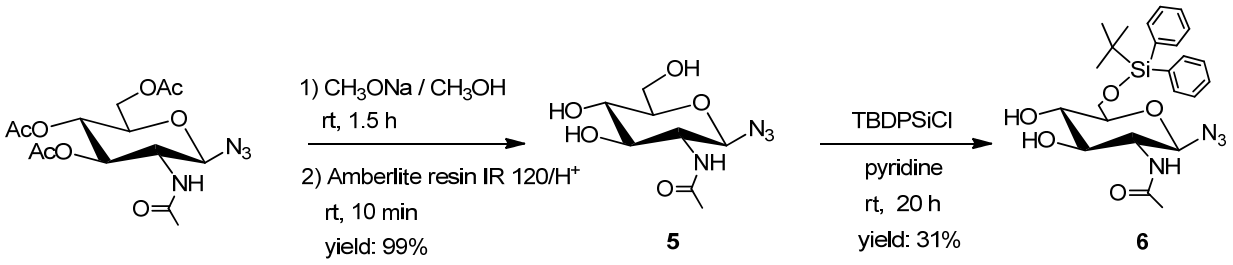


Figure 2.7 Synthesis of the glycosyl acceptor **6**.

The condensation reaction between glycosyl donor **4** and glycosyl acceptor **6** to get disaccharide**7** (figure **2.8**) was carried out at -45 °C in dry CH_2Cl_2 in presence of boron trifluoride etherate as promoter. In these conditions, it was possible to obtain the condensation between the two carbohydrates with the requested β(1→4) bond in 24% yield. The β configuration of the disaccharide **7** was unambiguously deduced from ^{1}H NMR spectra which showed the H_1' as a doublet at 4.72 ppm with a coupling constant of 8.0 Hz, typical of β anomers. However, differently from what reported by Gan *et al.*[12], a relevant presence of byproducts such as the 1→3 glycosylation product or the α anomers were also found.

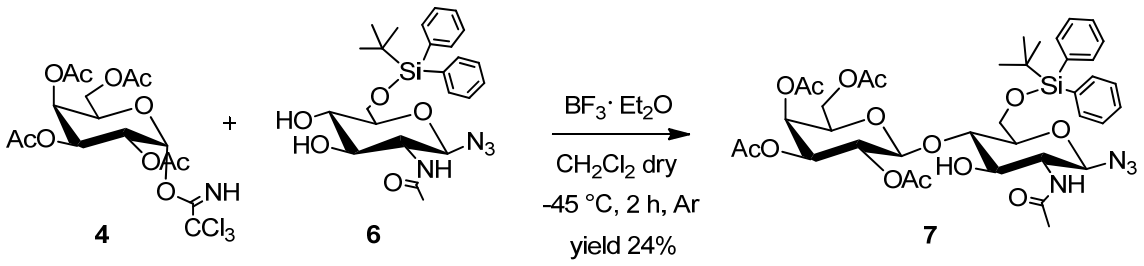


Figure 2.8 Glycosylation reaction between **4** and **6** to get compound **7**.

In a longer but safer pathway, the hydroxyl group at the 3-position of the acceptor was also protected to definitely avoid the formation of 1→3 glycosylation products. This strategy requires five protecting-deprotecting steps, instead of the two necessary to obtain product **6**, but it presents the advantage of obtaining the glucoside **11** as acceptor, which has only the hydroxyl group in position 4 available for the conjugation reaction (figure **2.9**). As for the synthetic route to obtain **6**, also here the first step consisted in the deprotection of the acetyl groups of 2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-β-D-glucopyranosyl azide with the Zemplén method to get compound **5** in quantitative yield. This saccharide was then reacted with the dimethylacetal of benzaldehyde in presence of p-toluene sulfonic acid (PTSA)[13] to obtain compound **8**, which bears the hydroxyl group in 4 and 6 position protected thanks to the formation of a benzylidene acetal. Derivative **8** was then acetylated in pyridine in presence of acetic anhydride to obtain compound **9**, that was subsequently deprotected from the benzylidene acetal with TFA[13] to get compound **10**, which is soluble in polar organic solvents or water. The use of pyridine to quench the reaction led to the formation of a large

amount of pyridinum salt that made the purification step not so straightforward. The following reaction consisted in the regioselective protection of the 6-OH group to get compound **11** having, as planned, only the 4-OH group free for the glycosylation reaction.

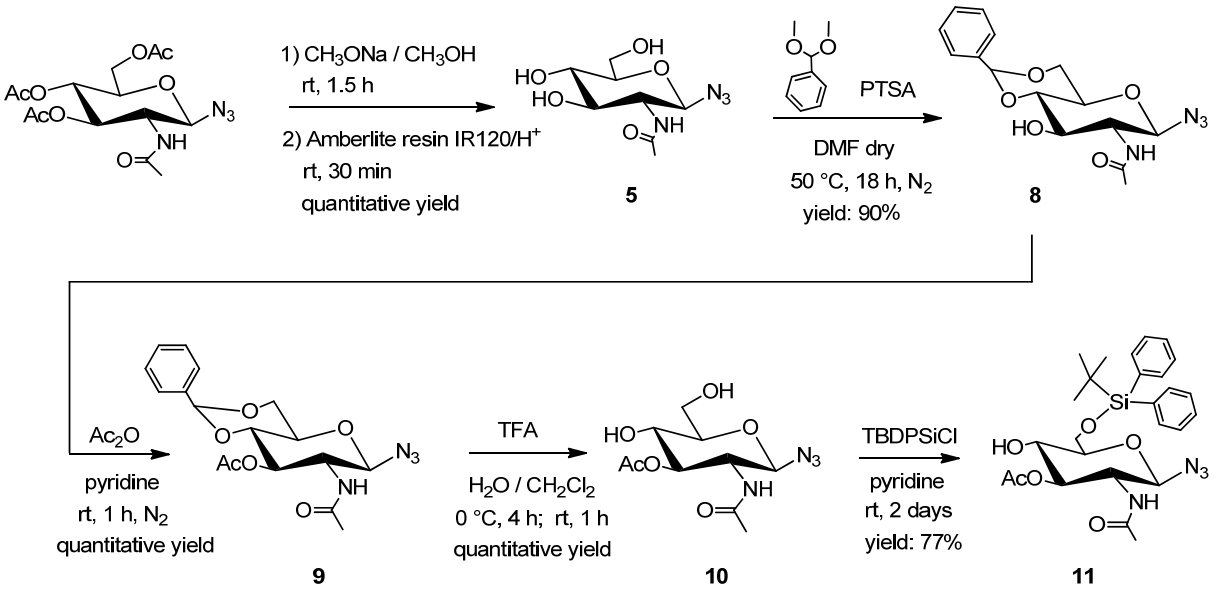


Figure 2.9 Synthesis of the glycosyl acceptor **11**.

All these synthetic steps to get compound **11** present high yields (at least 77%) with three out of five compounds obtained in quantitative yield.

The glycosylation reaction between compound **4** and **11** to get disaccharide **12** (figure **2.10**) was performed in analogous conditions as those to get compound **7**.

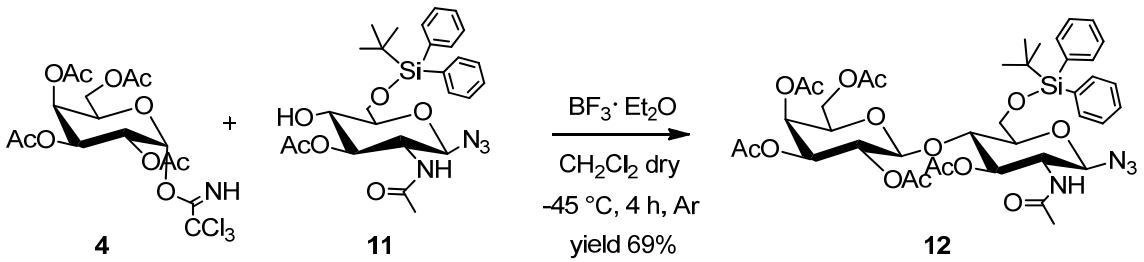


Figure 2.10 Glycosylation reaction between **4** and **11** to get compound **12**.

However, in this case it was possible to obtain the desired compound **12** with satisfactory yield (69%). A small amount of α linkage product was also detected, that made the purification of the β(1→4) anomer via column chromatography particularly tricky. Even if the reaction was not complete when it was quenched after 4 hours (presence of both reagents observed on TLC), further attempts to get better yield by adding more equivalents of boron trifluoride etherate as promoter (up to 2 equivalents) or by prolonging the reaction time (up to 30 hours) only led to the formation of larger amounts of byproducts without any improvements of the final yield.

At this point, desilylation reactions were carried out on the two disaccharides **7** and **12** with tetrabutyl ammonium fluoride and acetic acid in tetrahydrofuran[14]. The obtained compounds **13a** and **13** were subsequently peracetylated to give **14**. Subsequently, the azide group was reduced through catalytic hydrogenation in presence of Pd/C in order to obtain the amine compound **15** that was then immediately reacted with thiophosgene to give LacNAc-β-isothiocyanate **16** (figure **2.11**).

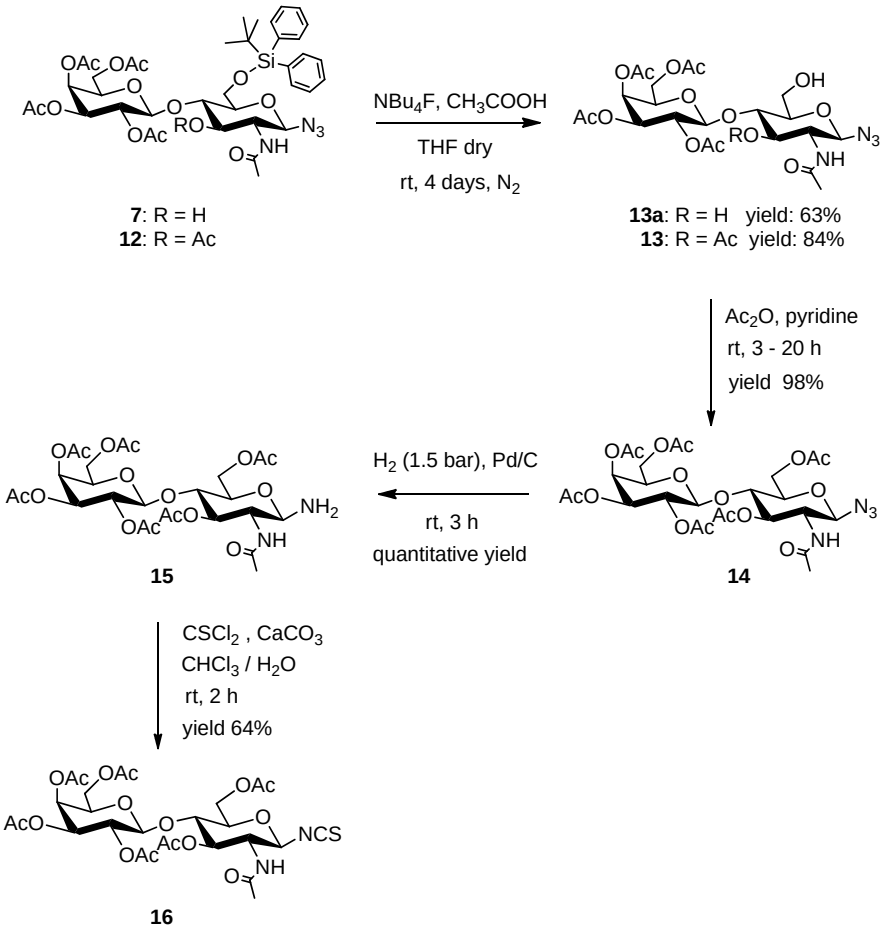


Figure 2.11 Synthesis of the LacNAc derivative **16**.

Large amounts of by-products were observed during the reduction of azido compound **14** to amino derivative **15** when the reaction time or the temperature were not strictly controlled. Mass spectra of these compounds proved the formation of a tetrasaccharide due to the reaction of the amine group of a LacNAc molecule with the anomeric position of another. Transglycosylation reactions between two unprotected glycosylamines are mentioned in several papers that describe the propensity of this type of molecules to rearrange in solution via an immonium-ion intermediate[15,16]. To reduce the amount of the transglycosylated product formed, the reaction was carried out for a maximum of 3 hours at a maximum of 20 °C. Moreover, since the presence of acid can promote the condensation between two molecules of glycosylamine to form the dimer, it was very important during the work-up procedures of precursor **14** to completely remove all the acetic acid formed during the exhaustive acetylation reaction. As soon as it was formed, compound **15** was reacted immediately with thiophosgene to afford compound **16**.

To obtain tetra-amine calixarene **20**, necessary to anchor the carbohydrate functionalities, standard procedures in the chemistry of these macrocycles were used (figure **2.12**). *p-tert*-Butyl-calix[4]arene **17** was alkylated at the lower rim by using propyl iodide in the presence of NaH as base to fix the macrocycle in the cone conformation and obtain compound **18**[17,18]. With the subsequent ipso-nitration reaction compound **19** was formed[19,20]. This reaction led to the desired compound **19** in high yields and with complete substitution of the *tert*-butyl groups with nitro groups, as confirmed by ^{1}H NMR that shows the total disappearance of the *tert*-butyl signals at around 1 ppm. By treating compound **19** with hydrazine and Pd/C (10%), product **20** was obtained[21]. The complete reduction of the nitro groups to amines was easily confirmed by recording a ^{1}H NMR spectrum, where a broad signal at 3.1 ppm, corresponding to the amine groups, appears and an high-field shift for the signal of the aromatic protons, compared to that of the nitro derivative **19**, is observed.

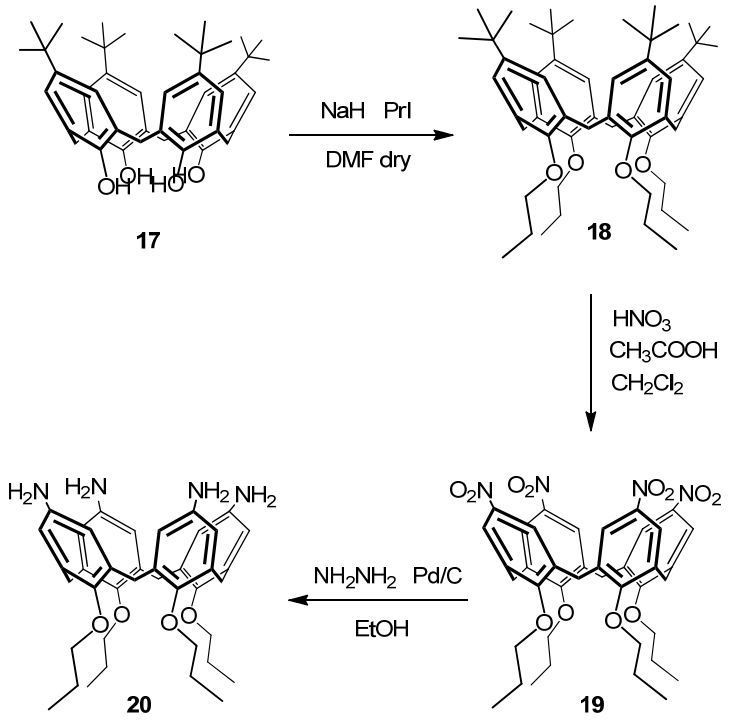


Figure 2.12 Synthesis of tetra-amino calix[4]arene **20**.

Coupling reaction between isothiocyanate LacNAc **16** and tetra-amino calix[4]arene **20** led to the formation of compound **21** where the glycosyl units are connected to the calixarene backbone through thiourea spacers (figure **2.13**). This "click" reaction gave compound **21** in high yield, and the purification was necessary only to remove the excess of carbohydrate used. No partially functionalized intermediates were detected. Deprotection of **21** using the Zemplén method[22] gave target compound **1** in 90% yield (figure **2.13**). Complete removal of the acetyl groups could be achieved in 4 hours as determined by ESI-MS analyses.

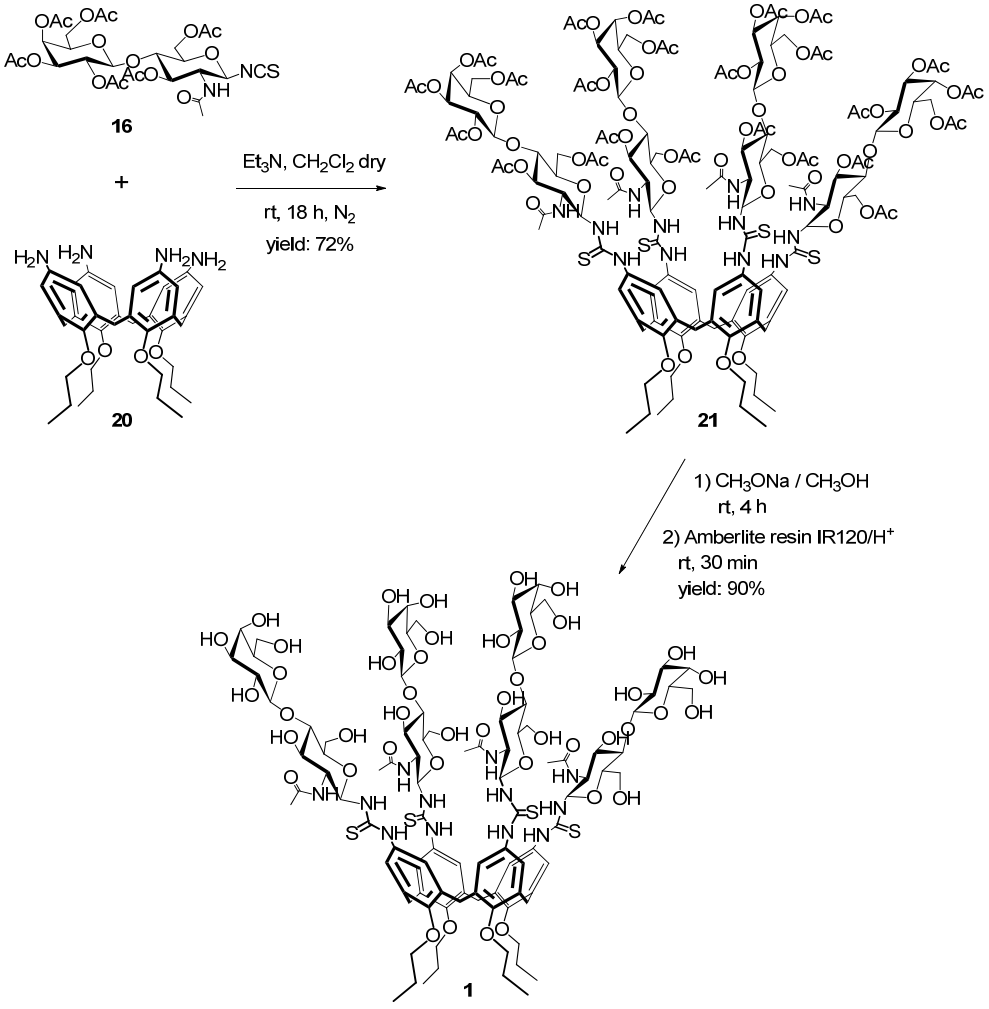


Figure 2.13 Coupling reaction between isothiocyanate derivative **16** and amino calixarene **20** to give compound **21**, and subsequent deprotection reaction to obtain *cone*-4LacNAc[4]Prop **1**.

Cone-4LacNAc[4]Prop is soluble in water up to the concentration of 0.1 M, but it tends to form aggregates as detected by NMR spectra in D_2O (figure **2.14**). By increasing the temperature from 25 °C to 70 °C it is possible to observe the NMR signals to become sharper as a proof of disaggregation of compound **1.**

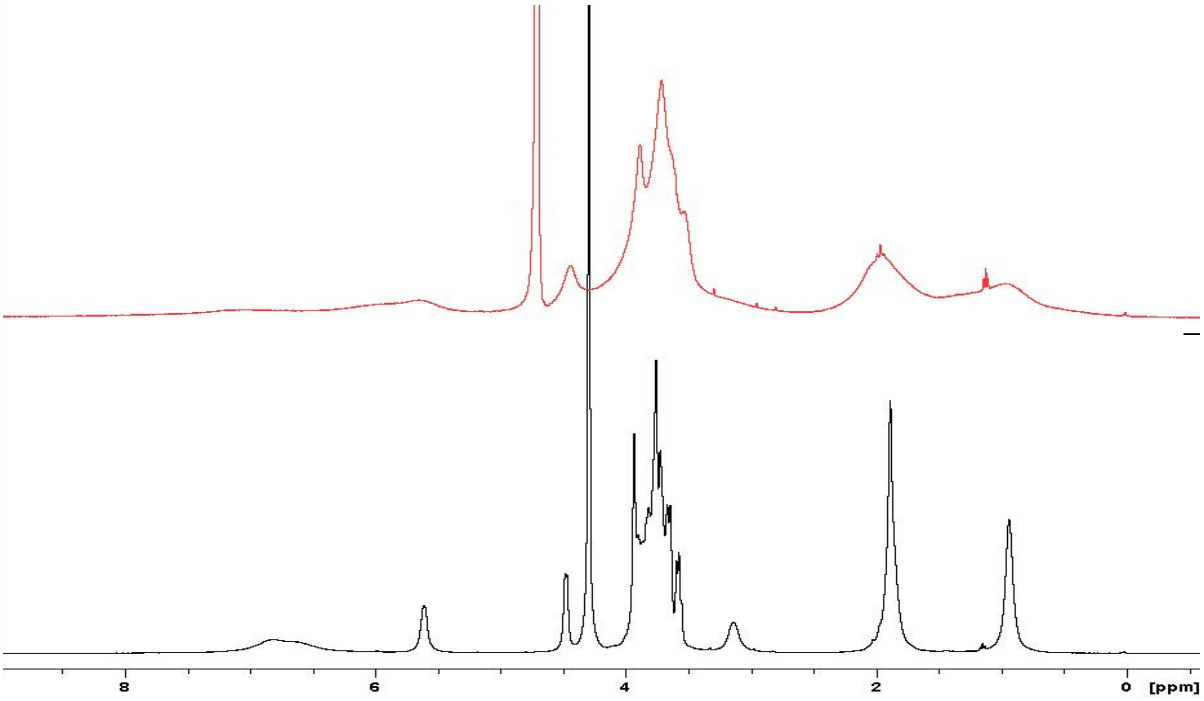


Figure 2.14 ^{1}H-NMR in D_2O (400 MHz) of *cone*-4LacNAc[4]Prop **1** at 25 °C (top) and at 70 °C (bottom).

Also the addition of an organic solvent such as methanol induces the aggregates to break (figure **2.15**) since, by solvating the lipophilic tails of the calixarene, the alcohol weakens the hydrophobic interactions responsible of the aggregation process.

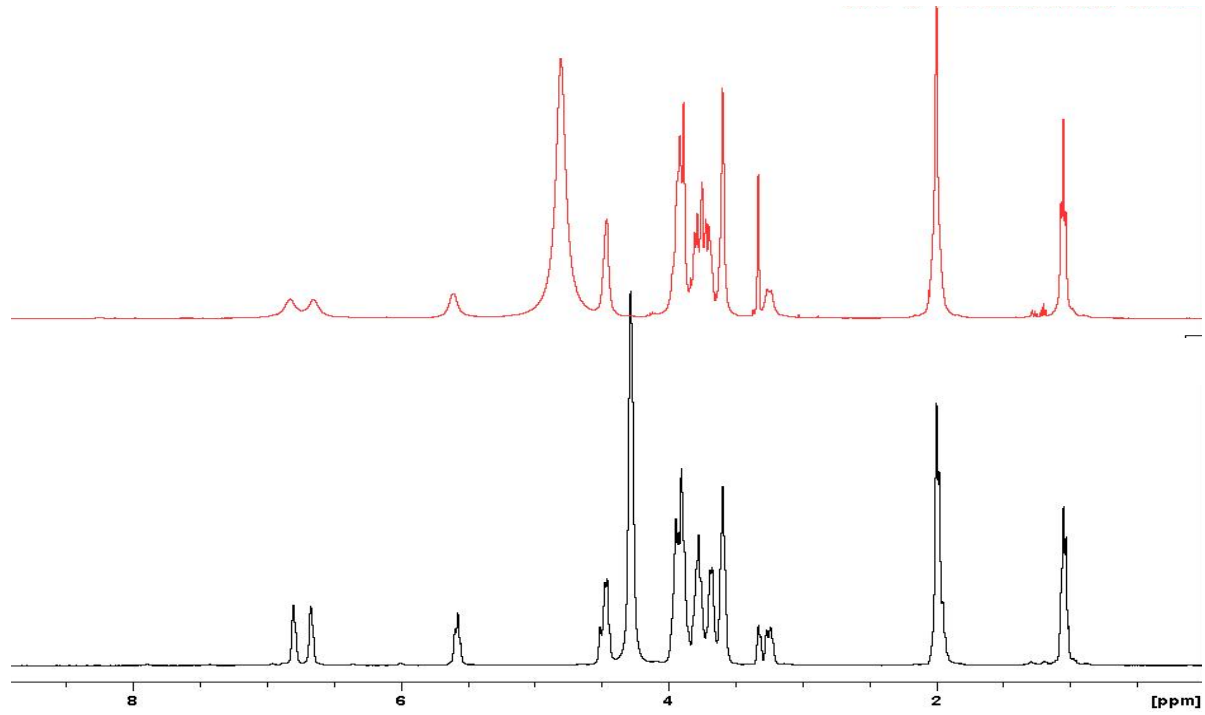


Figure 2.15 ^{1}H-NMR in CD_3OD/D_2O 5:1 (400 MHz) of *cone*-4LacNAc[4]Prop **1** at 25 °C (top) and at 70 °C (bottom).

Even at room temperature, the quality of the ^{1}H NMR spectrum registered by using a CD_3OD/D_2O 5:1 mixture increases with respect to the one recorded in pure D_2O. Especially the resonances of the aromatic protons, that are split in two signals because of the amide resonance and of the chirality of the substituents, became rather sharp when the ^{1}H NMR spectrum was recorded at 70 °C. This, together with ESI-MS analyses, gave confirmation that the target molecule **1** was effectively obtained.

2.2.1.2 Chemoenzymatic synthesis[§]

Considering that the above discussed chemical pathway to obtain calix[4]arene **1** showed to be quite long and not so efficient in some steps, an alternative route was investigated. It is well known that chemoenzymatic synthesis is a powerful tool to obtain complex carbohydrates because it combines the flexibility of chemical synthesis to the highly selectivity of enzymatic synthesis. The possibility of using a simple glycocalixarene as substrate for enzymes to build more complicated sugars directly on the macrocyclic scaffold was therefore investigated. This new strategy exploits the enzymatic galactosylation on the *cone*-4GlcNAc[4]Prop **28** functionalized at the upper rim with four units of N-acetyl-glucosamine (GlcNAc) linked to the macrocycle with thioureido groups. This method would allow to obtain oligosaccharide-functionalized macrocycles, such as N-acetyl-lactosamine calixarenes, with a minimum synthetic effort, avoiding all the time demanding steps of selective functionalization of the sugar hydroxyl groups through protection-deprotection steps necessary to direct the glycosylation reactions only at specific positions.

First of all, it was necessary to chemically synthesize compound **28**. Tetraamine-tetrapropyl-calix[4]arene **20** and peracetylated N-acetyl-GlcNAc-β-isothiocyanate **26** were prepared and conjugated thanks to the isothiocyanate-amine "click" reaction. The synthesis of compound **26** (figure **2.16**) started with the reaction of commercially available GlcNAc **22** with acetyl chloride that acts both as solvent and reagent, to give compound **23**. In one step it was possible to achieve simultaneously the acetylation of the hydroxyl groups of the carbohydrate and the introduction of a chlorine atom in anomeric α position[23]. This reaction

[§] *Part of the research presented in this paragraph was carried out at the Darmstadt University (Germany), during a two months visit in the group of Prof. Wolf-Dieter Fessner.*

leads exclusively to the α isomer, as confirmed by the ^{1}H NMR spectrum, where it is possible to observe the presence of a doublet signal at 6.10 ppm with a coupling constant of 3.0 Hz, diagnostic for the α anomer. Once obtained compound **23**, two pathways are possible to synthesize the isothiocyanate derivative **26** (figure **2.16**).

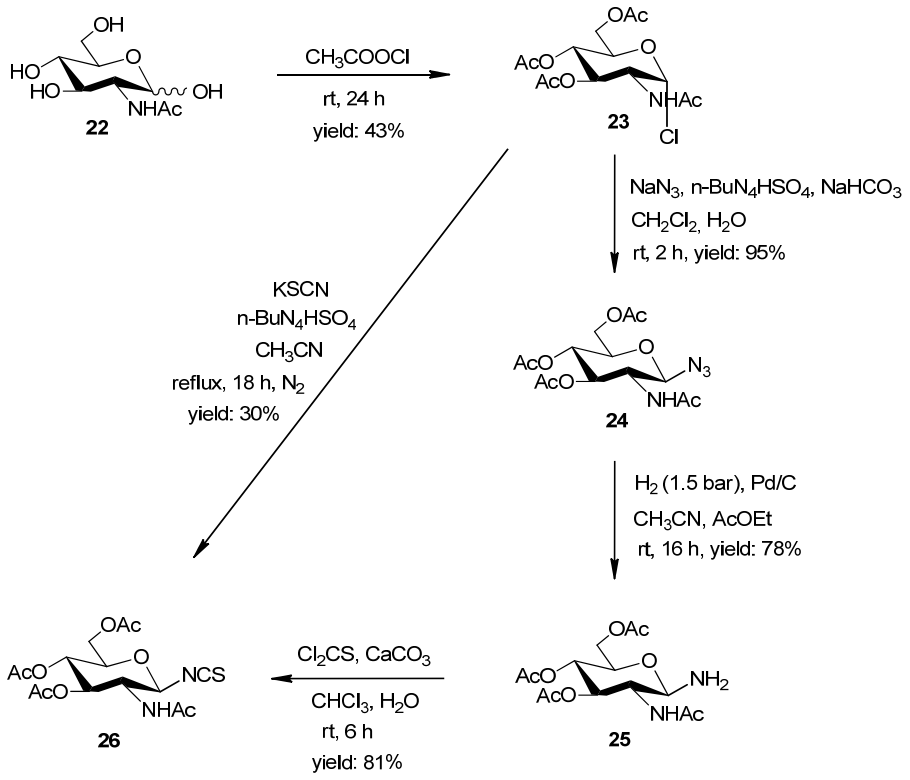


Figure 2.16 Synthesis of GlcNAc-β-isothiocyanate **26**.

The first route provided compound **26** in just one step[24] by using potassium isothiocyanate and tetrabutylammonium hydrogen sulfate. The yield of this reaction was very low, only 30% obtained, probably due to the high temperature used for the reaction (80 °C) that led to a consistent formation of byproducts that were difficult to separate even via column chromatography. Another possibility to obtain compound **26** was to proceed through a three steps pathway. First, the GlcNAc-α-chloride **23** was converted into the corresponding β-azide **24** by S_N2 reaction with phase transfert catalysis[25,26]. This reaction led to the complete configuration interconversion at the anomeric carbon, as confirmed by ^{1}H NMR experiments, that evidenced the resonance of the signal relative to the H_1 proton shifted downfield to 4.74 ppm compared to the reagent and a coupling constant of about 9 Hz, typical for the β

anomer of a 1,2-trans glycoside. Azide compound **24** was then reduced by catalytic hydrogenation in presence of Pd/C (10%), giving β anomerically pure amine derivative **25** subsequently transformed into the corresponding isothiocyanate by reaction with thiophosgene. The overall yield of the three steps is 60%, that is the double of the yield obtained with the one-step methodology directly from chloride derivative **23** to compound **26**. Besides, all the reactions led to pure compounds without the need of column chromatography purification. It is possible to conclude that the second pathway provides a better and easier method for the synthesis of isothiocyanate carbohydrates despite the higher number of synthetic steps.

GlcNAc derivative **26** was then linked to the calix[4]arene core **20** through the formation of the thiourea unit (figure **2.17**).

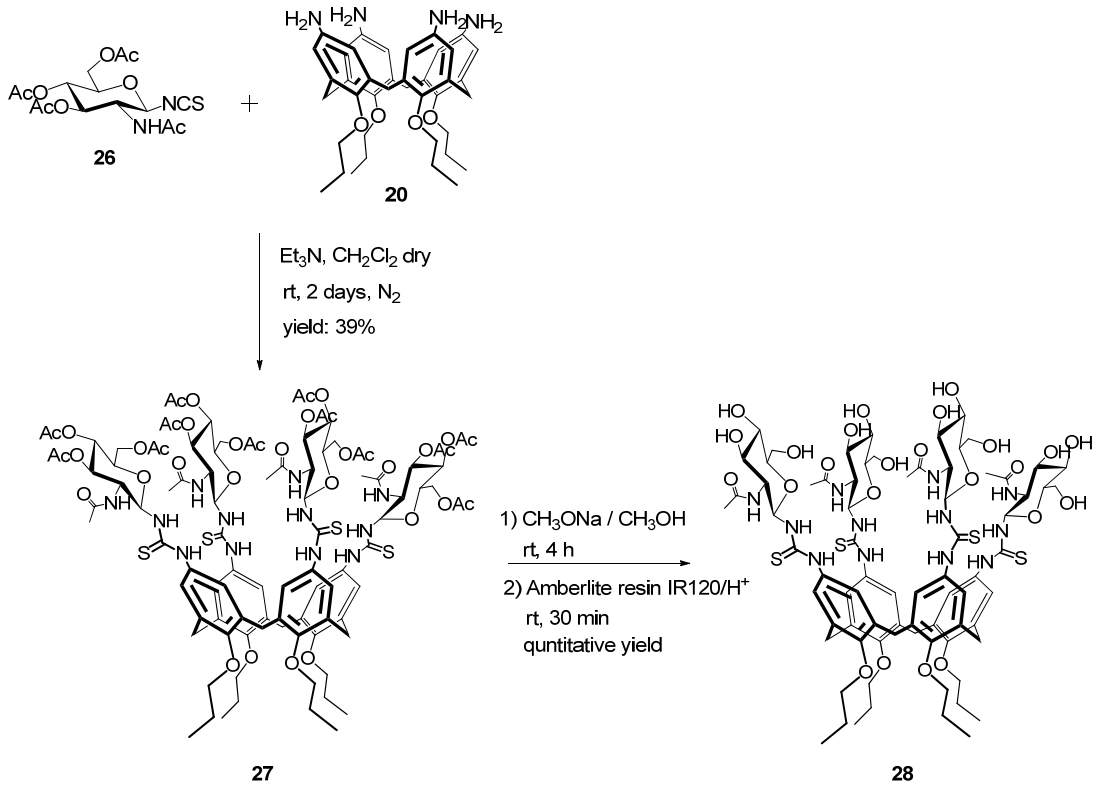


Figure 2.17 Synthesis of *cone*-4GlcNAc[4]Prop **27** via "click" reaction and subsequent deprotection to yield compound **28**.

This "click" reaction gave compound **27** in high yield, and the purification was necessary only to remove the excess of carbohydrate used. NMR analyses confirmed the success of

the reaction. The ^{1}H NMR spectrum (300 MHz, DMSO-*d6*, figure **2.18**) shows the presence of the thiourea NH protons at 9.63 and 7.49, and of two doublets at 4.32 and 3.22 ppm attributed to the axial and equatorial protons of the calixarene methylene bridge, respectively. From these data we could conclude that the calixarene, in the cone structure, had a C_4 symmetrical substitution at the upper rim with four identical saccharide units. The β anomeric protons, resonating at 5.51 ppm, confirmed that the sugar maintained, as expected, its configuration during the "click" reaction. ESI-MS analyses provided a further confirmation of the tetra-glycosylation to obtain **27**.

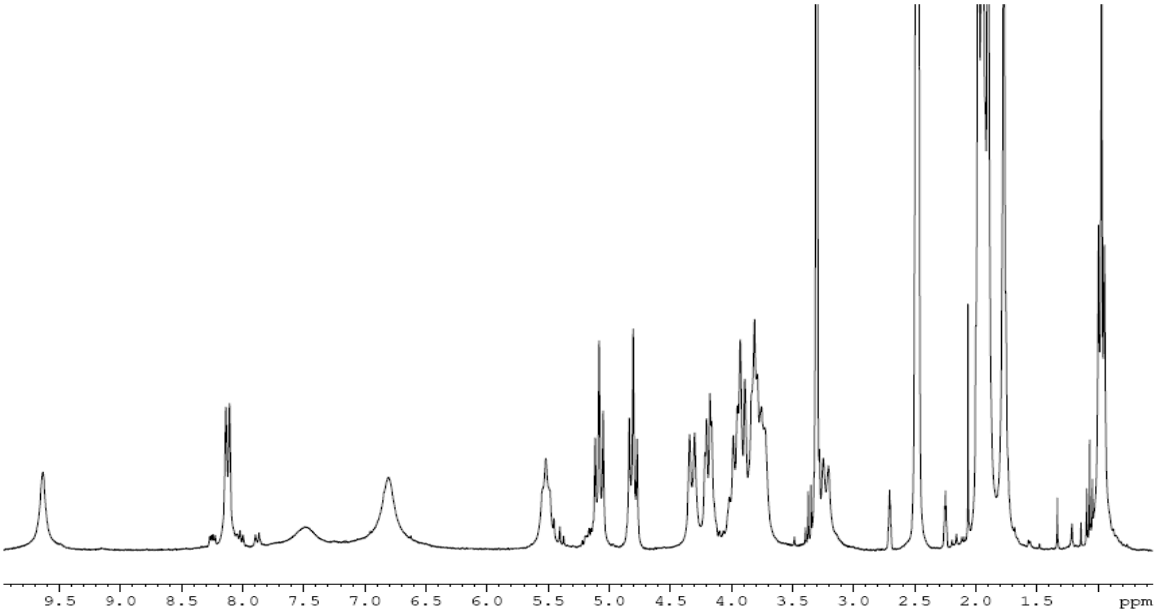


Figure 2.18 ^{1}H NMR spectrum (300 MHz, DMSO-*d6*) of peracetylated GlcNAc-calix[4]arene **27**.

The subsequent step was the deprotection of the peracetylated GlcNAc-calix[4]arene **27** to give product **28**. The hydrolysis of the acetate groups was carried out at room temperature by adding a freshly prepared sodium methoxide solution in methanol to compound **27** according to the standard Zemplén procedure[22]. Complete removal of the acetyl groups could be achieved in 4 hours as determined by ESI-MS analyses: the peak at m/z 1724.3 corresponding to the $[M+Na]^{+}$ species unambiguously confirmed that the desired compound **28** was obtained. Product **28** was completely characterized with NMR techniques, with all the peaks assigned thanks to 2D NMR experiments. ^{1}H NMR spectrum

of **28** (figure **2.19**) shows the disappearance of the methyl singlets of the acetyl groups between 2.10-1.60 ppm, with the only exception of the C2 amide signal at 2.03 ppm.

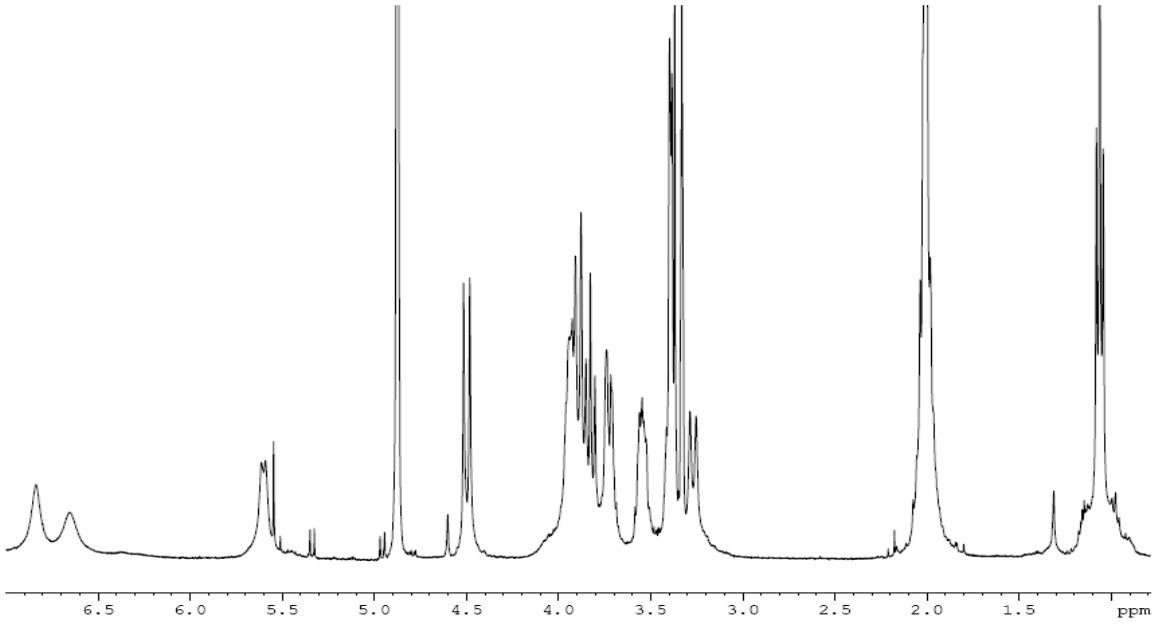


Figure 2.19 ^{1}H NMR spectrum (400 MHz, CD_3OD) of cone-4GlcNAc[4]Prop **28**.

In collaboration with Prof. Fessner and his research group at Darmstadt University (Germany), enzymatic glycosylation of GlcNAc-calix[4]arene **28** was at this point explored in order to obtain *cone*-4LacNAc[4]Prop **1** (figure **2.20**). Galactose-UDP (Gal-UDP) as glycosyl donor and 1,4-galactosyltransferase (1,4-GalT) as enzyme to drive the reaction were used (both compounds provided by Prof. Fessner). The reaction was performed by dissolving 1 mg of *cone*-4GlcNAc[4]Prop **28** in 400 µL of buffer (TRIS-HCl 25 mM, $MgCl_2$ 5 mM, NaCl 0.5 M, pH adjusted to 7.5). A sample of Triton X-100 was added to the reaction mixture in order to have a concentration of 0.3% (v/v). This allows to improve the solubility of calixarene **28** and to avoid its self-aggregation that could prevent the active site of the enzyme to reach the molecule with consequent impossibility for the glycosylation reaction to take place. The reaction was stirred at 30 °C in presence of 1,4-GalT and an excess of Gal-UDP.

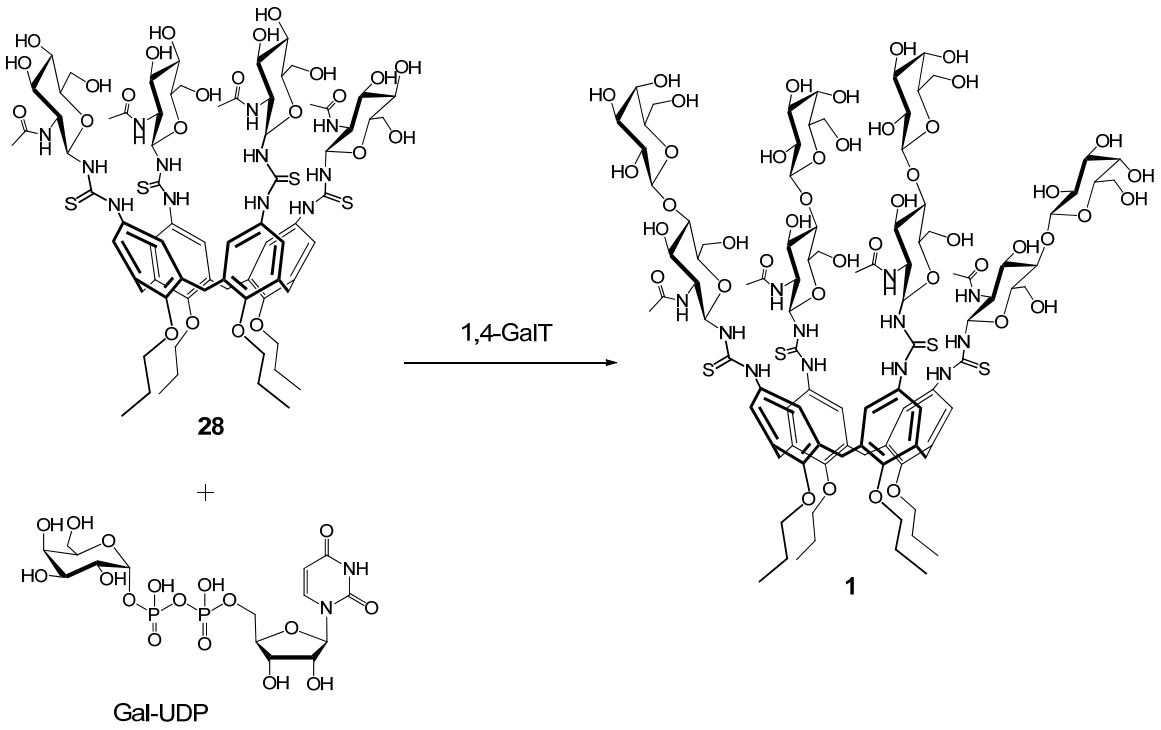


Figure 2.20 Enzymatic synthesis of *cone*-4LacNAc[4]Prop **1**.

The reaction was continuously monitored via HPLC using a PRONTOSIL 120-5-Phenyl 5.0 µm column and isopropanol/water (0.1% formic acid) as eluent (gradient from 30% to 60% of alcohol in 20 minutes). In these analyses we were helped by the fact that we could use both the calixarene reagent **28** and the product **1** already obtained via classical chemical pathway (see paragraph **2.2.1.1**) as standards. After 12 hours, the HPLC analysis showed the presence of five chromatographic peaks (figure **2.21**) in the range of elution time between 11 and 16 minutes. The peak with the highest elution time (15.7 min) corresponds to the reagent **28**, while the one with the lowest elution time (11.1 min) and intensity is given by the tetra-functionalized product **1**. This was confirmed by injecting together a sample from the enzymatic reaction mixture and the chemically obtained product **1** used as reference compound. The other three peaks between the slowest and the fastest eluted compounds have been assigned to partially glycosylated intermediates. The reaction can also be followed, even if in a more rough way, via TLC by using butanol/acetone/AcOH/water 40:40:5:15 as eluent that was found to be proper to separate the reagent and the product.

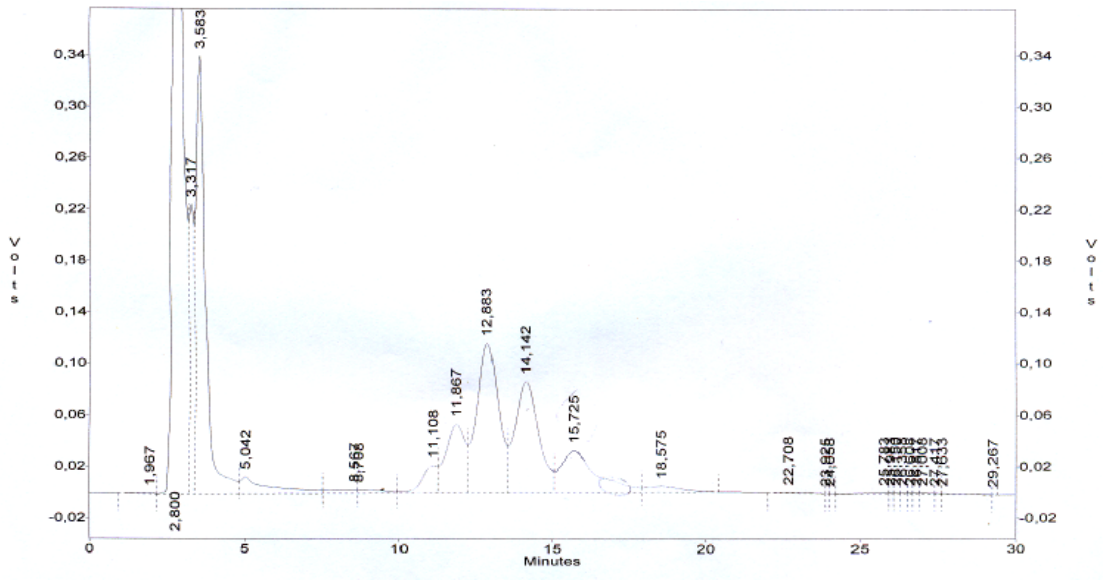


Figure 2.21 HPLC chromatogram of the galactosylation reaction after 12 hours.

The mixture was allowed to react for 48 hours and further additions of Gal-UDP and 1,4GalT were made in the meanwhile. After this portion of time, the HPLC profiles evidenced the almost complete disappearance of the peaks of the reagent and of the partially functionalized compounds, with a considerable increase of intensity of the peak having a retention time of 11.2 min and corresponding to product **1** (figure **2.22**).

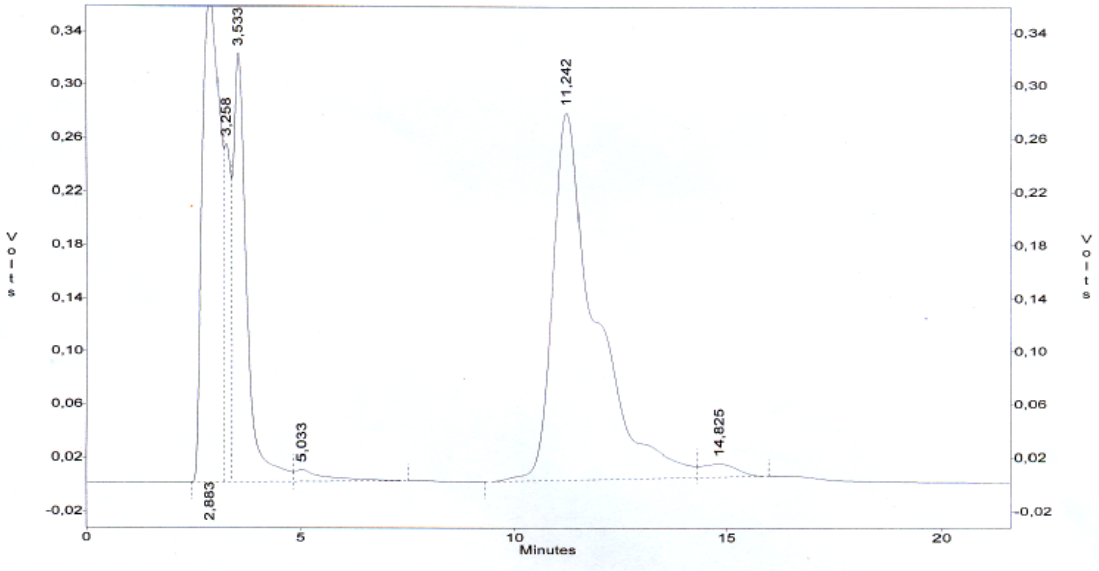


Figure 2.22 HPLC chromatogram of the galactosylation reaction after 48 hours.

The reaction was then quenched by addition of ethanol to denature 1,4-GalT and therefore block the galactosylation reaction. ESI-MS analyses confirmed the reaction was successful. It was in fact possible to observe the signal at m/z 1197.9 for the disodium molecular ion $(M+2Na)^{2+}$ of tetra-LacNAc macrocycle **1**, and only a much less intense peak at m/z 1117.8 due to the presence of traces of tri-functionalized calix[4]arene.

This result is particularly interesting and important because it demonstrates for the first time that suitable calixarene derivatives can be exhaustively and efficiently glycosylated by using glycosyltransferases. A non-natural substrate as a glycosyl-calixarene can therefore be recognized and transformed by this type of enzymes, opening the possibility to build, step by step, more complex carbohydrates directly on the calixarene scaffolds.

2.2.2 Enzymatic sialylation of *cone*-4Lac[4]Prop calixarene[**]

Although all galectins show specificity towards β-galactosides, the preference that they exhibit in binding specific carbohydrates is dependent on the variability in their Carbohydrate Recognition Domain (CRD) sequence. Multiple studies are reported on the selectivity of galectins towards sialylated glycans. It is known that the presence of sialic acid α-linked to the 3' position of Lac and LacNAc can further improve the intrafamily selectivity, for example towards Gal-8[27], while other studies suggest that most galectins exhibit diminished binding to β-galactosides decorated with 2,6 sialic acid[28]. Anyway Cummings *et al*.[29] have demonstrated a certain selectivity of galectins towards α-2,3 and α-2,6 sialylated glycans: for example Gal-2 exhibits significantly reduced binding to all sialylated carbohydrates, whereas Gal-1 can bind to α-2,3 but not α-2,6-sialylated compounds, and Gal-3 instead shows affinity either for some α-2,3- or α-2,6-sialic acid terminating glycans.

Hence, looking at a further step in the selectivity ability of our glycocalixarenes, the preparation of sialylated clusters was planned. The perspective to further increase the complexity of the saccharide portion on the calixarene backbone, skipping the tedious need of several protection-deprotection steps requested by a completely chemical approach, combined with the novelty of the use of calixarene derivatives as enzyme substrates,

[**] *The research presented in this paragraph was performed at the Darmstadt University (Germany), during a two months visit in the group of Prof. Wolf-Dieter Fessner.*

prompted us to go on in the investigation of the enzymatic approach applied to our unnatural macrocyclic derivatives.

Always in collaboration with Prof. Fessner's group, sialylation reactions with 2,3- or 2,6-sialyl transferase (2,6-SialT and 2,3-SialT, respectively) were carried out on *cone*-4Lac[4]Prop calix[4]arene **30**[1] in presence of the activated glycosyl donor CMP-sialic acid (Sial-CMP) to synthesize the corresponding tetra-sialylated clusters **31** and **32** (figure **2.23**). Derivative **30** was mainly used as more easily available model to verify the reactivity and set up the experimental conditions to be used for sialylation of the more valuable *cone*-4LacNAc[4]Prop calixarene **1**.

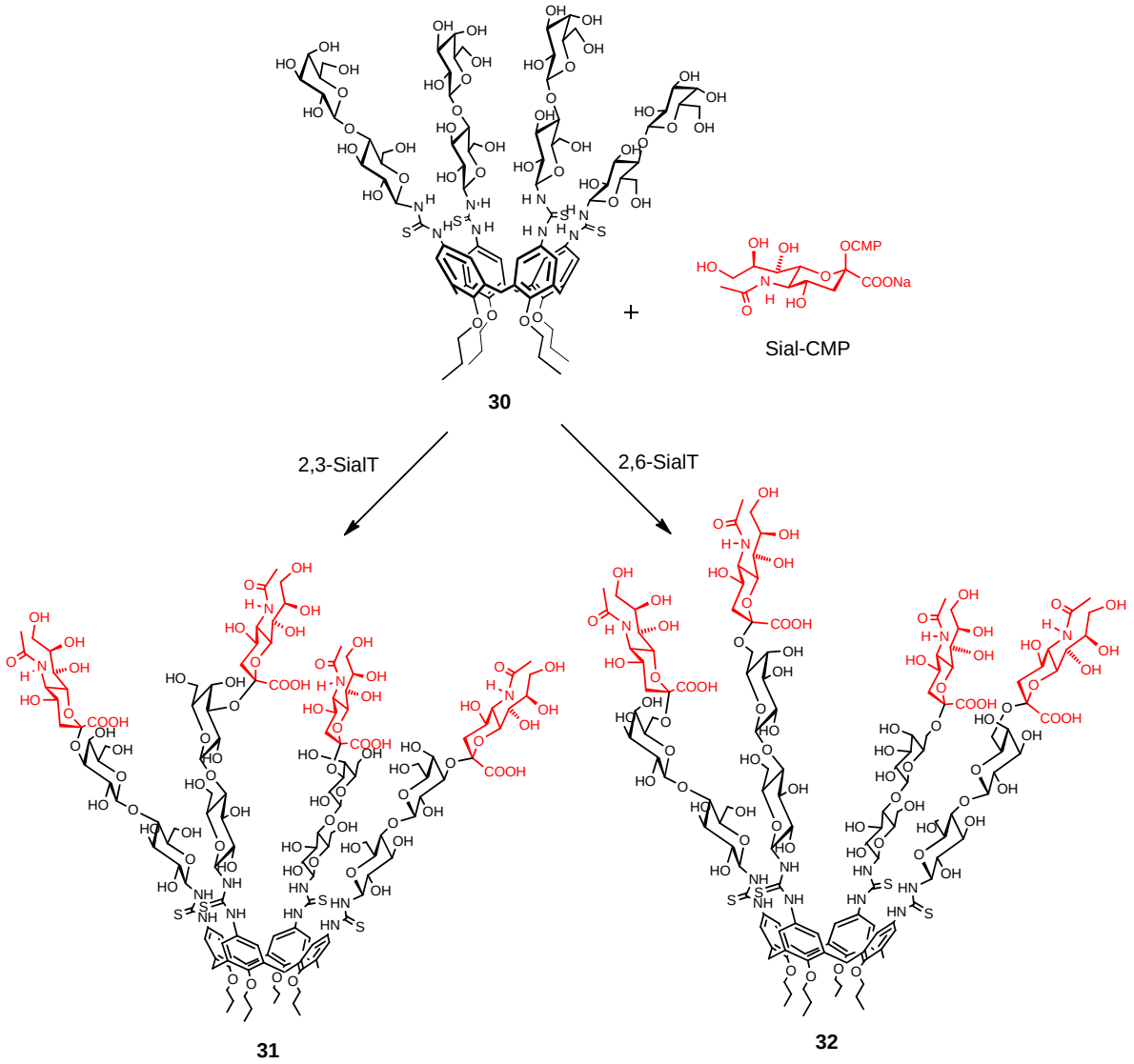


Figure 2.23 Enzymatic sialylation of *cone*-4Lac[4]Prop **30** with 2,3-SialT and 2,6-SialT in presence of Sial-CMP to obtain the tetrasialylated compounds **31** and **32**, respectively.

Chemoenzymatic strategies in the studied cases would represent a great tool to obtain the target molecules, since it would be possible to skip all the time demanding and compound consuming steps of protection-deprotection of the hydroxyl groups that chemical synthesis requires. With the final goal to obtain compounds **31** and **32,** several reactions having different conditions were carried out on *cone*-4Lac[4]Prop **30**. Solvent and buffer composition, temperature, pH, glycosyl-donor presentation (pure Sial-CMP reagent or enzymatically produced *in situ*) are all parameters that were varied in order to find the optimal experimental conditions to obtain the tetrafunctionalized glyco-clusters. The best results were obtained by dissolving calixarene **30** in buffer (TRIS-HCl 25 mM, $MgCl_2$ 5 mM, NaCl 0.5 M, pH adjusted to 8 for 2,6-SialT and to 7.5 for 2,3-SialT) at 4 mM concentration and subsequent addition of a first portion of previously purified Sial-CMP (2 equivalents). Since the glycosyl-donor is not completely stable in the reaction conditions, it was added batchwise during the total reaction time (8 equivalents on the whole). Subsequently, also 2,3-SialT or 2,6-SialT were added to the mixtures, that were gently stirred at 30 ^{o}C. The reactions were carried out in the presence of alkaline phosphatase that hydrolyses CMP preventing its inhibition activity towards glycosyltransferases. HPLC analyses revealed a very slow progression of both the reactions and therefore additional amounts of enzymes were added daily to the reaction mixtures. After ten days no more improvements could be observed and both the reactions with 2,3-SialT and 2,6-SialT were quenched by addition of ethanol.

HPLC and ESI-MS analyses of the reaction mixtures revealed that tetrasialylated compounds **31** and **32** were not obtained. Only a partial sialylation was reached, with a clear evidence only for the mono- and di-sialylated products as detected by ESI-MS, which showed the peaks at m/z 1239.2 and m/z 1385.8 for the two compounds, respectively, corresponding to the $(M+2H)^{2+}$ adducts. The peaks present in the HPLC profiles (figure **2.24** for reaction with 2,3-SialT and figure **2.25** for reaction with 2,6-SialT), other than the starting reagent having the highest retention time and the mono- and di-functionalized molecules, could be ascribed to other sialylated compounds, but they were present in too low amount to be fully characterized.

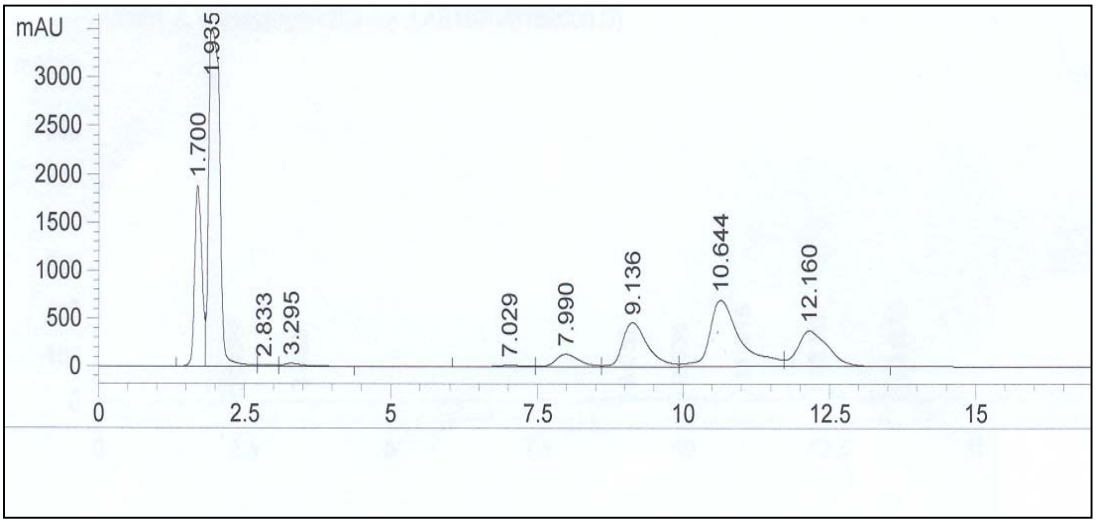


Figure 2.24 HPLC chromatogram for reaction of *cone*-4Lac[4]Prop **30** with 2,3-SialT.

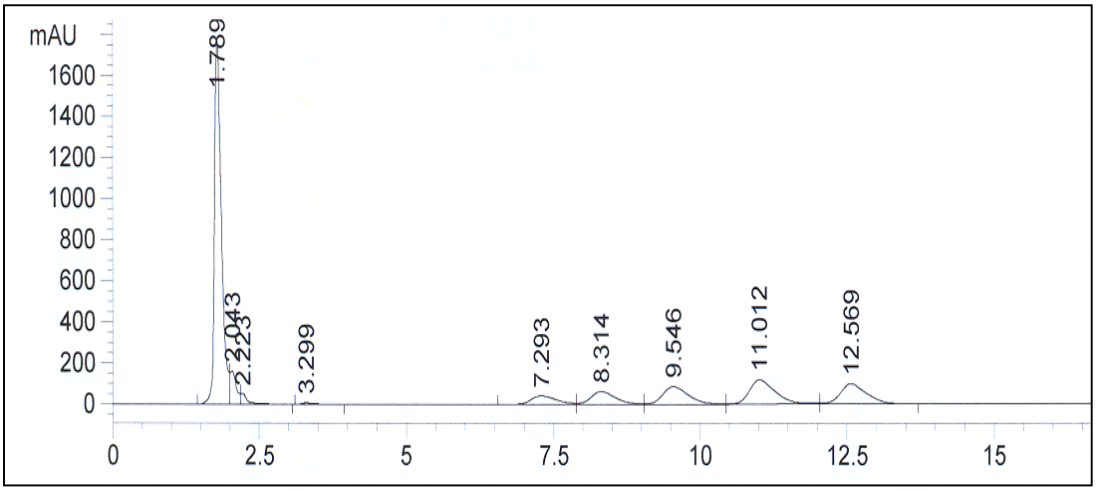


Figure 2.25 HPLC chromatogram for reaction of *cone*-4Lac[4]Prop **30** with 2,6-SialT.

These reactions show that sialyl transferase enzymes can recognize glyco-calixarenes even if they are non natural substrates. Unluckily in all the analyzed cases it was not possible to obtain the tetra-sialylated compounds. A possible explanation could be that the active binding site of the enzyme cannot reach the substrate, probably due to steric hindrance problems. To overcome this issue, and verify this hypothesis, alternative strategies described in the following paragraph **2.2.3** were explored.

2.2.3 Alternative glycocalixarene platforms for enzymatic sialylation

Enzymatic sialylation problems could arise from the difficulties of the enzyme to reach the substrate because of the steric hindrance caused by the close proximity of the four lactosyl moieties to the calixarene core. Therefore, alternative strategies were investigated. We first focused our attention on the possibility to obtain the target molecules **31** and **32** by leading the enzymatic sialylation reaction on the amino-disaccharide derivative **36** and connecting the obtained sialylated compound to tetra-isothiocyanate calix[4]arene **33**[30] via amine-isothiocyanate "click" reaction. Amino compound **36** was chosen as possible substrate for the enzymes, instead of the isothiocyanate carbohydrates used so far for the "click" reactions, to avoid undesired reactions with the amino residues present in the structure of the sialyl-transferases. Azido compounds were also not taken into consideration because we wanted to limit the number of steps to be performed on the saccharides already functionalized with sialic acid to the minimum. Tetra-isothiocyanate calix[4]arene **33** was synthesized in a rather straightforward manner from the corresponding amino-calix[4]arene **20**[21] by using thiophosgene in dry toluene (figure **2.26**).

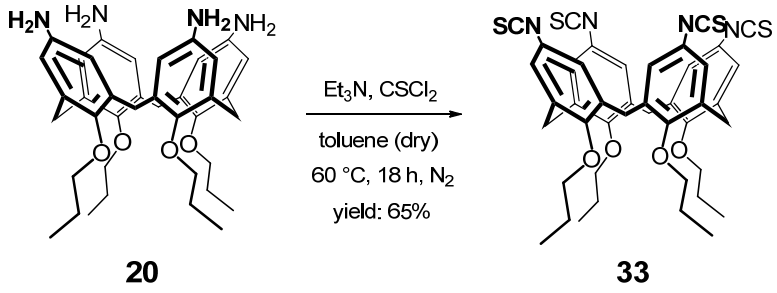


Figure 2.26 Synthesis of *cone*-tetra-isothiocyanate calix[4]arene **33**.

Compound **36** was synthesized from the corresponding hepta-acetylated-β-lactosyl azide **34**[26] via deprotection with the Zemplén method to get compound **35**, and subsequent reduction with Pd/C under hydrogen atmosphere (figure **2.27**).

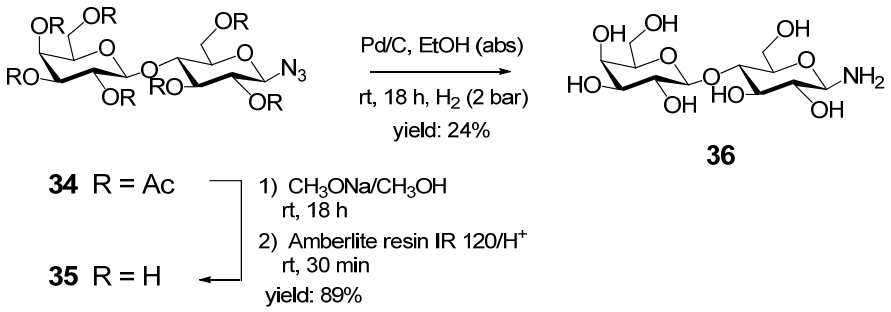


Figure 2.27 Synthesis of amino derivative **36.**

Even if it is known that *O*-unprotected glycosyl amines in solution give rise to equilibria involving their cyclic and acyclic forms[15], a recent article by López *et al.*[31] reported the synthesis of 1-β-isothiourea derivatives from these unprotected sugars, suggesting that it is possible to work in water solution with pure β-amino-glycosyl derivatives. Anyway, we observed that compound **36** undergoes fast anomerization in water. Reduction reaction under H_2 atmosphere with Pd/C in absolute ethanol led to pure β-compound, but since it is scarcely soluble in ethanol, only 24% yield was achieved. After filtering off the catalyst, the filters were washed with water to collect more product (up to 85% yield) but it was recovered as a mixture of β and α anomers in 9:1 ratio. It was immediately clear that this strategy could not be used. In fact, enzymatic sialylation reactions are performed in water medium, and an anomer mixture, that always result very difficult to be separated, would be obtained. Furthermore, a test reaction between compound **36** and **33** (figure **2.28**) failed to obtain the *cone*-4Lac[4]Prop **30** and only a mixture of mono- and di-glycosylated intermediates was isolated.

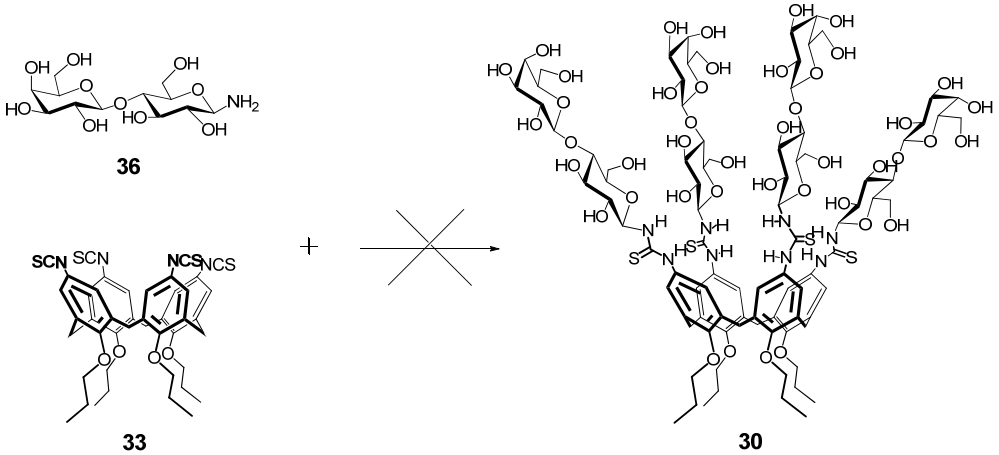


Figure 2.28 Test reaction between amino derivative **36** and *cone*-tetra-isothiocyanate calix[4]arene **33**. The formation of *cone*-4Lac[4]Prop calixarene **30** could not be achieved.

An alternative synthetic pathway was therefore designed and compound **41** became the new molecular target (figure **2.29**). Although rather similar to compound **30**, the new glycocluster **41** bears a two carbon atom spacer between the sugar moiety and the thiourea functionality. This allows an higher flexibility of the saccharide chains and hopefully a higher accessibility to the 3' or 6' positions of the sugar residues of compound **41** to the sialyl transferases.

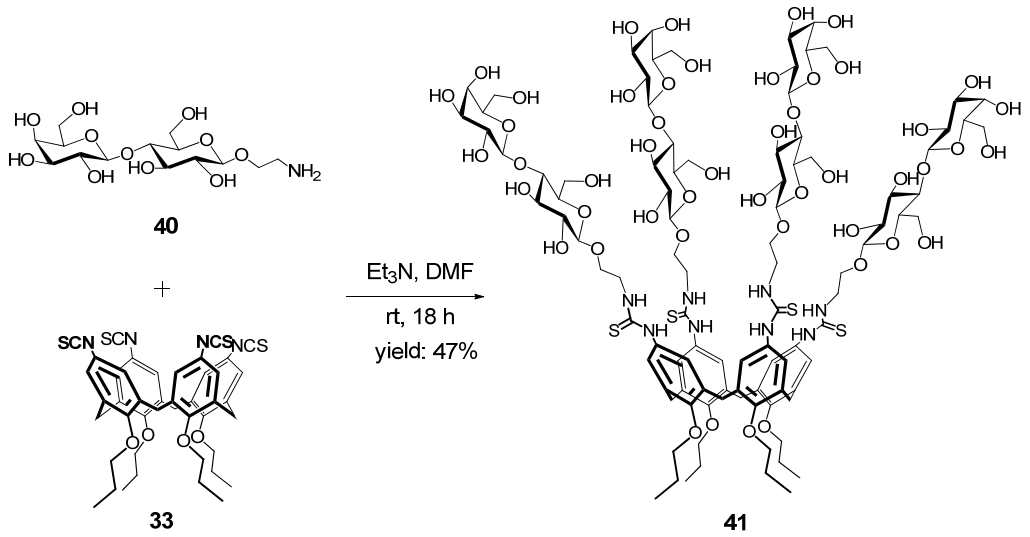


Figure 2.29 Coupling reaction between amino derivative **40** and *cone*-tetra-isothiocyanate calix[4]arene **33** to give compound **41**.

Aminoethyl β-lactoside **40** was synthesized in four steps[32] starting from peracetylated lactose (figure **2.30**). Peracetylated-β-lactose (which contained 5% of α-anomer as impurity) was coupled with 2-bromoethanol by treatment with $BF_3 \cdot OEt_2$ in dry CH_2Cl_2 to afford heptaacetyl-2-bromoethyl-β-lactoside **37** that was used without purification in the following substitution reaction with sodium azide. A column chromatography purification was then performed to obtain pure β-anomer of compound **38**. A deacetylation reaction led, in almost quantitative yield, to the sugar derivative **39** that was subsequently reduced to amino compound **40**[32] via catalytic hydrogenation using Pd/C (10%) in H_2O and ethanol.

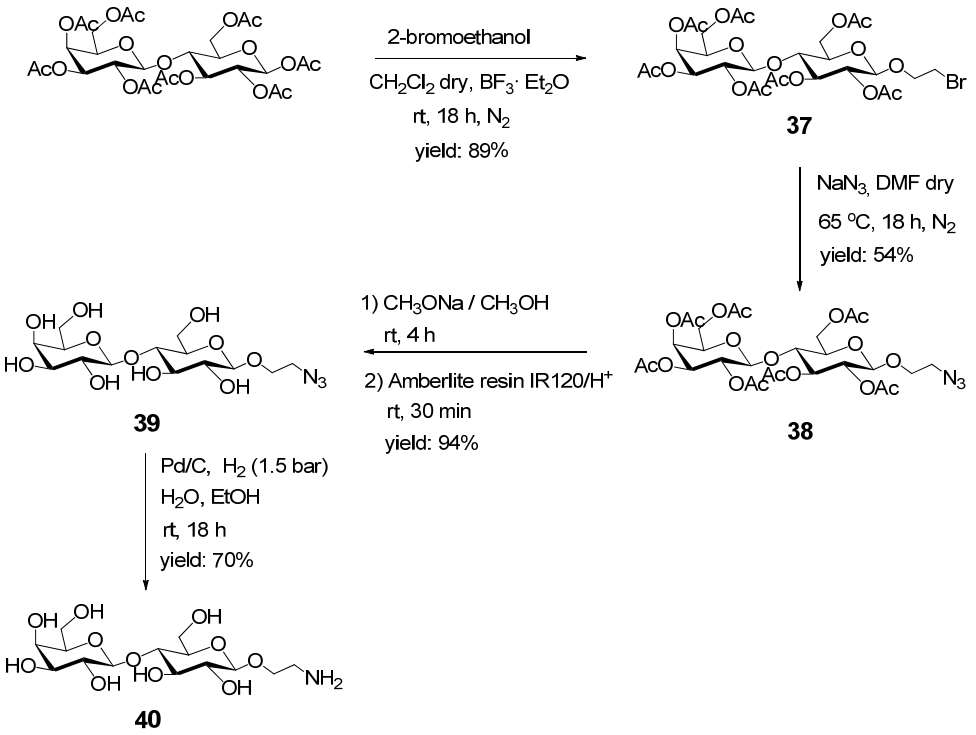


Figure 2.30 Synthesis of saccharide **40**.

Aminoethyl-β-lactoside **40** showed to be stable in water. Moreover, in the case it would not be possible to achieve the exhaustive sialylation of the four lactosides directly on glycocluster **41**, the disaccharide **40** could also represent a suitable substrate for the sialyl transferases thus disclosing a new way for the preparation of our sialylated target compounds.

Tetraisothiocyanate calixarene **33** showed satisfying reactivity towards **40**, and glycocluster **41** was obtained without particular difficulties and in reasonable (47%) yield. It is worth noting that in this case the nucleophilicity of the amino-sugar **40** is clearly superior to that of the 1-β-aminoglycoside **36** of figure **2.28**. No partially functionalized intermediates were detected and the glycocluster **41** was fully characterized by NMR spectroscopy. Curiously, despite the clearness of the NMR spectra, ESI-MS analyses evidenced the presence of peaks corresponding to partially functionalized calixarenes still bearing few isothiocyanates groups. This apparent inconsistency was clarified by MS/MS electrospray experiments, which showed that these peaks are indeed due to a fragmentation with loss of saccharide units and simultaneous rearrangement back to the isothiocyanate moiety. Compound **41** resulted to be well soluble in water, regardless of the small aliphatic chain present between the

saccharide and thiourea units, and it is now ready to be tested as substrate for sialyl transferases. As already observed for *cone*-4LacNAc[4]Prop **1**, also compound **41** tends to form aggregates in water, as suggested by the presence of rather broad peaks in the ^{1}H NMR spectrum in D_2O (figure **2.31**). Also in this case it was possible to observe the NMR signals to become sharper either by increasing the temperature (figure **2.31** bottom), or by dissolving it in an organic solvent (figure **2.32**) as a proof for disaggregation of the assembly.

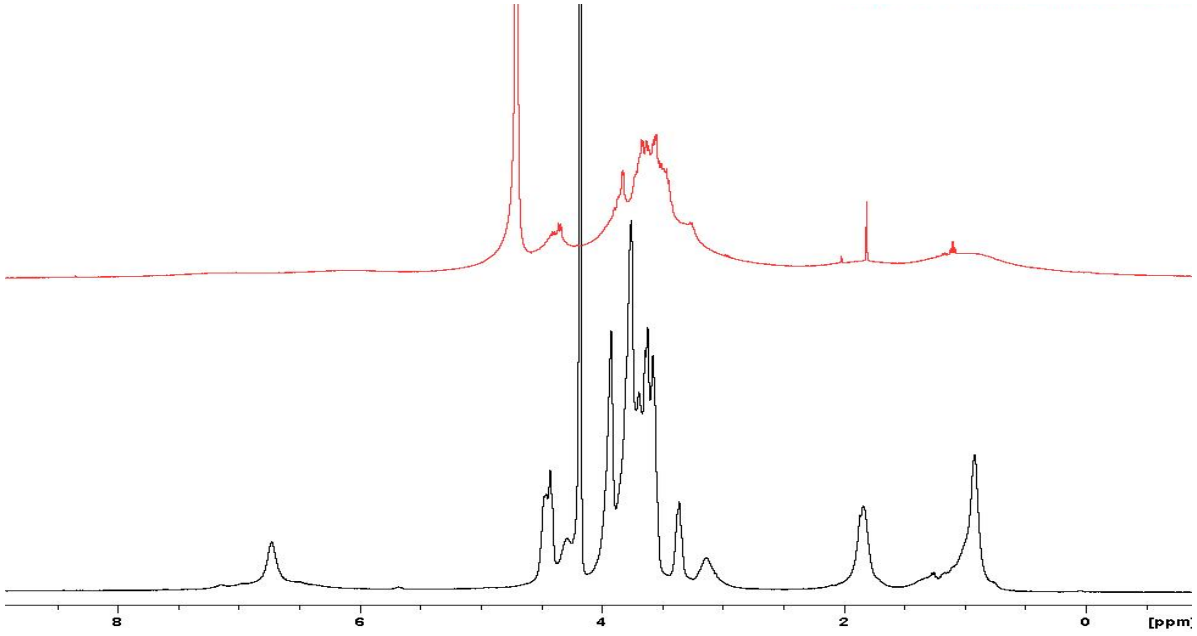


Figure 2.31 ^{1}H NMR in D_2O (400 MHz) of compound **41** at room temperature (top) and at 80 °C (bottom).

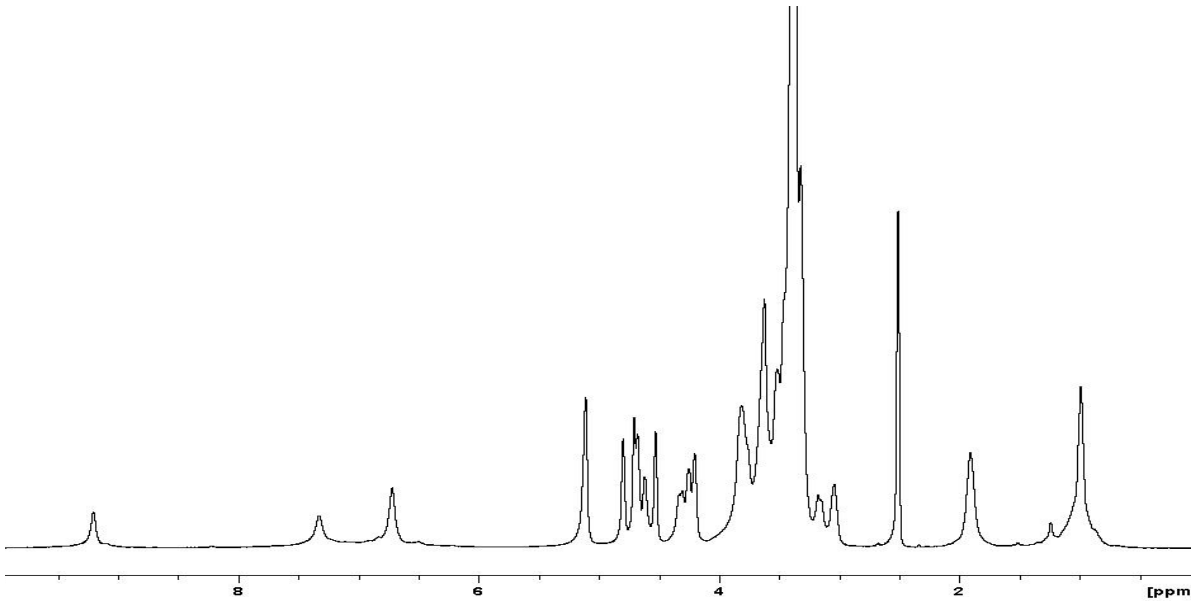


Figure 2.32 ^{1}H NMR in DMSO (400 MHz) of compound **41** at room temperature.

2.2.4 Combining carbohydrate modification at bioinspired positions with multivalent presentation††

To achieve strong and selective inhibition of galectins, it results very important to tailor the nature of the carbohydrate ligands to fully match the galectin's individual microenvironment of the binding site for the best-possible sugar-protein contact and complementarity. Besides, combining these modified carbohydrates with a suitable cluster presentation allows to exploit the benefits of multivalency. When the scaffold makes accessible different geometries for the ligand presentation, as in the case of calixarenes, the possibility to gain in selectivity increases, as previously demonstrated with the lactosylcalixarenes[1]. The decision on where to introduce structural modifications on carbohydrates, comes from the knowledge that galectins share affinity to the core disaccharide LacNAc, but exhibit different responses in terms of affinity on substitutions at the 3' position[33]. Analyses by crystallography and flexible docking provided a detailed structural rationale for the preference towards aromatic substituents in 3' position in the case of human galectin-3[34]. Nilsson *et al.*[5] have recently reported the synthesis of LacNAc derivatives modified by the addition of aromatic functionality at the 3' position of the galactose moiety. These compounds showed increased affinity for Gal-3 (K_d down to 320 nm) compared to LacNAc (K_d = 67 μm), due to the interaction of the aromatic ring with the side chain of an arginine residue Arg144 in the proximity of the binding site. On the other hand, introduction of aromatic substituents at the 2-N position of LacNAc could establish additional interactions with Arg 186 in Gal-3, which is located in close proximity to the C2 of a bound lactose or LacNAc moiety[4,5]. The presence of an arginine in that position is also conserved in other galectins[35], such as Gal-1, Gal-7 and Gal-9N (N-terminal domain), but lacks for example in Gal-8N, where the corresponding position is occupied by an isoleucine. Therefore also 2-N-position of LacNAc has attracted interest as target for the introduction of aromatic substituents[2].

In this context, within a trilateral collaboration involving Dr. Grandjean (Université de Picardie, F), Prof. Gabius (Ludwig Maximilian Universität Münich, D) and our group, it was

†† *The research presented in this paragraph was performed in collaboration with Dr. Cyrille Grandjean (Université de Picardie, F) and Prof. Hans-Joachim Gabius (Ludwig Maximilian Universität Münich, D). Results published[36].*

planned to decorate calixarene scaffolds with LacNAc synthetically modified at the 2-N and at the 3' positions, in an effort to enhance the affinity and selectivity toward different types of galectins.

Dr. Grandjean selected and prepared two differently modified LacNAc monovalent ligands, compounds **42** and **43** (figure **2.33**) that present a benzyl group at 3' position or a benzoyl group at 2-N position, respectively.

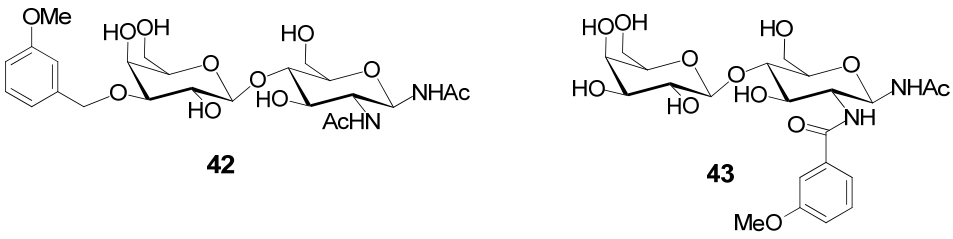


Figure 2.33 Structures of the monovalent compounds **42** and **43.**

The simultaneous disubstitution at the two positions of the disaccharide with the aromatic moieties could combine in one molecule the advantages given by the single modifications, but it did not seem the ideal choice because of solubility problems in aqueous environment.

Also the two β-isothiocyanate analogues **44** and **45** were provided by Dr. Grandjean. These two compounds were conjugated to the upper rim of alkoxycalix[4/6]arenes via a thiourea-linker followed by deprotection from acetyl groups. (figure **2.34**). For the multivalent presentation of the epitopes we chose the tetra-aminocalix[4]arene **20** blocked in the cone conformation and the conformationally mobile hexa-aminocalix[6]arene **46**. This choice was based on the binding properties towards galectins given by these architectures when decorated with lactosyl units, as previously shown[1].

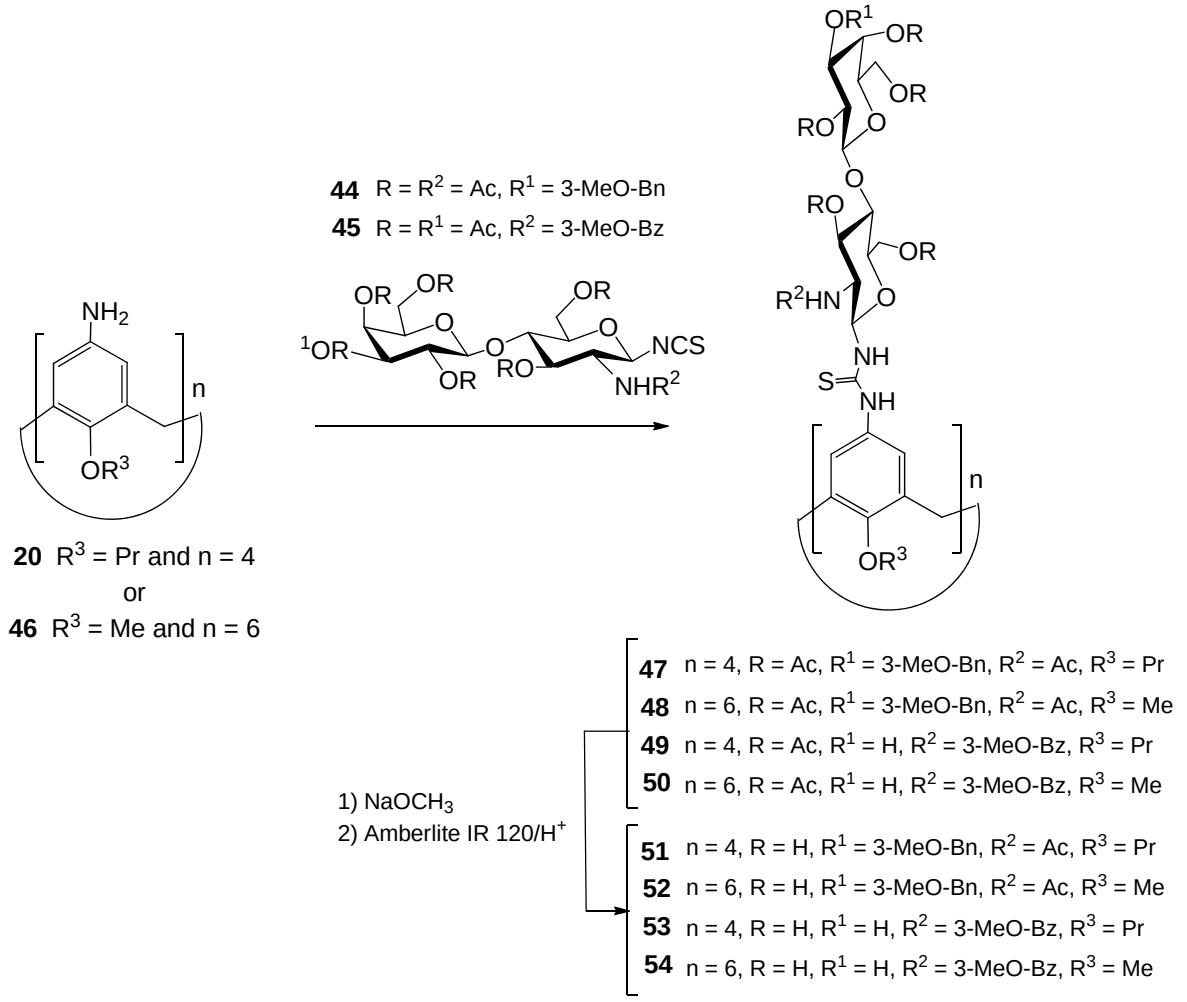
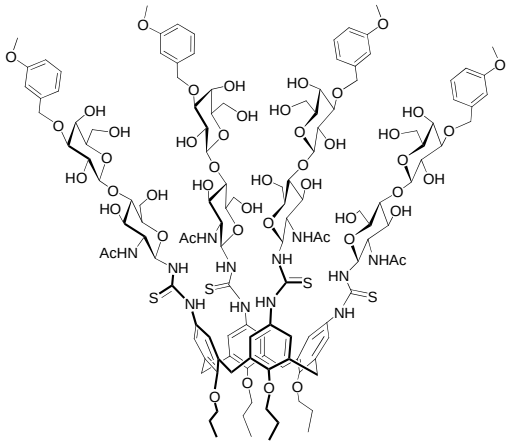
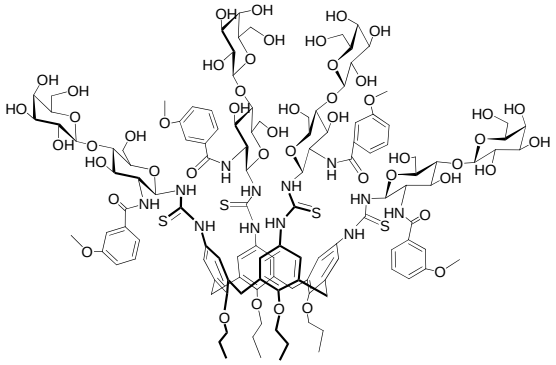


Figure 2.34 Synthetic scheme for the preparation of the calixarene based glycoclusters **51**-**54**.

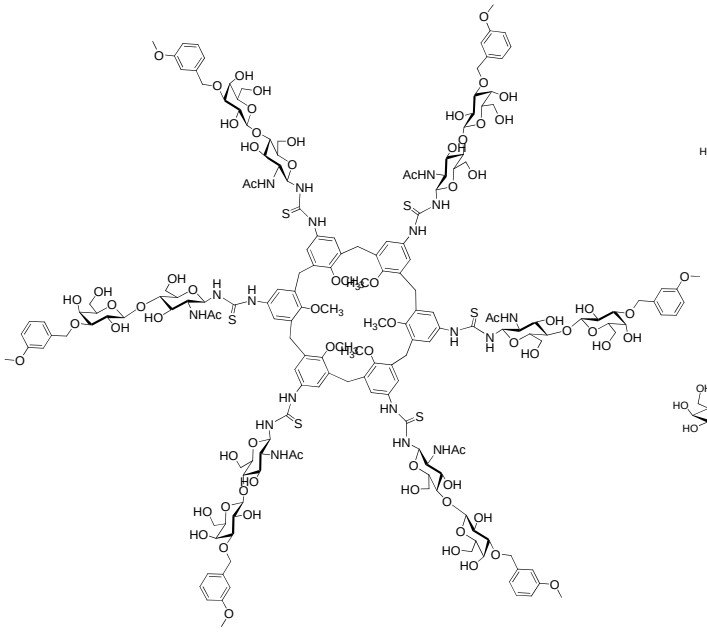
Four different clusters (**51**-**54**) were therefore synthesized: *cone*-4-(3'BnLacNAc)[4]Prop **51**, 6-(3'BnLacNAc)[6]Met **52,** *cone*-4-(2BzLacNAc)[4]Prop **53** and 6-(2BzLacNAc)[6]Met **54** (figure **2.35**). All these compounds were fully characterized by NMR and mass spectrometry. Along with the monovalent controls **42** and **43**, compounds **51**-**54** were used in the tests of inhibition of different lectins[36], which were performed by Prof. Gabius and his coworkers.



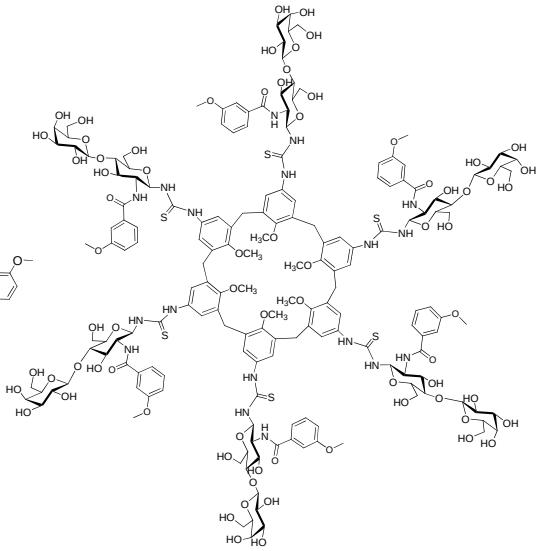
51
cone-4-(3'BnLacNAc)[4]Prop



53
cone-4-(2BzLacNAc)[4]Prop



52
6-(3'BnLacNAc)[6]Met



54
6-(2BzLacNAc)[6]Met

Figure 2.35 Structures of the four calixarene glycoclusters **51**-**54** based on 3' or 2-N substitution in the LacNAc unit.

The effects of derivatization and glycocluster formation on inhibitory potency was tested against the *Viscum Album Agglutinin* (VAA) plant toxin and five different human adhesion/growth-regulatory lectins (inter- and intra-family comparison). In contrast to galectins, the presence of substituents at the 2-N and 3' sites of LacNAc has negative effects on the sugar affinity for the galactoside-specific plant toxin VAA, that therefore provided an inherent negative control for the biological studies[37,38].

Inhibition of lectins binding to asialofetuin (ASF), a glycoprotein that possesses nine LacNAc epitopes, adsorbed on microtiter plate surface was first studied. The extent of bound lectin was determined spectrophotometrically at 490 nm using a colorimetric test which exploits the peroxidase activity on the *o*-phenylenediamine.

By comparison with lactose (table **2.1**) and with the previously determined IC_{50} values of the corresponding lactoclusters based on *cone*-calix[4]arene (*cone*-4Lac[4]Prop; 200 μM for Gal-3, 20 μM for Gal-4, see[1]) and on conformationally mobile calix[6]arene (6Lac[6]Met; 400μM for Gal-3, 5 μM for Gal-4, see[1]), titrations revealed a general increasing extent of inhibition and enabled to evaluate IC_{50} (inhibitor concentration to reach 50%-level of signal reduction)-values relative to these new monomers and clusters (table **2.1**).

Table 2.1 IC_{50} values given in μM sugar concentration for derivatives **42** and **43** and clusters **51**-**54** on the binding of lectins to immobilized ASF. Coating with constant amount of 0.5 μg ASF per well. Assays to determine the IC_{50}-value sugar were routinely performed in triplicates for up to five independent series with standard deviations not exceeding 14.2%.

Type of inhibitor	Position of aromatic substituent	Valency	VAA (0.5 μg/mL)	Gal-3 (3 μg/mL)	Gal-4 (5 μg/mL)	Gal-9 (10 μg/mL)
Lactose	-	1	300	2500	1200	3000
42	3’	1	4800	250	800	900
43	2-N	1	800	360	1250	300
51	3’	4	2800	0.15	2.2	2.8
52	3’	6	420	0.4	0.8	1.4
53	2-N	4	200	1.3	2.5	2.0
54	2-N	6	45	1.5	1.2	0.8

As expected, the tested substitutions were not favourable for the inhibition of VAA, although the hexavalent presentation of **52**, and especially of **54**, was still capable to establish some activity compared to the corresponding monovalent ligands **42** and **43**. The latter compounds, in fact, showed an IC_{50} value even higher than lactose. Compared to lactose based ligands, the inhibitory capacity was markedly increased by both types of substitution towards Gal-3 and also Gal-9, a bit less for Gal-4. The chimera-type Gal-3 maintained its characteristic sensitivity for *cone*-type tetravalent presentation, noted previously[1]. Relatively

to lactose (IC_{50} = 2500 μM), a nearly 10-fold increase in inhibition ensued from 3'-substitution, and a further 1600-fold increase due to tetravalency lowered the IC_{50}-value to 0.15 μM (in sugar), with a consequent efficiency improvement also with respect to *cone*-4Lac[4]Prop (IC_{50} = 200 μM)[1].

The hexavalent clusters **52** and **54** confirmed to be better inhibitors for the tandem-repeat-type galectins Gal-4 and Gal-9, compared to the corresponding tetravalent clusters, as already demonstrated by 6Lac[6]Met with Gal-4[1]. Again a significant decrease of IC_{50} was observed going from monovalent to multivalent ligands.

Both synthetic processes thus substantially add up to a significant gain in lectin-blocking capacity even if, for Gal-4 the aromatic substituents at 3' and 2-N positions have poor effects on the efficiency of the monovalent inhibitors **42** and **43** compared to lactose.

Parallel testing disclosed rather weak inhibitory activity of the test panel, with IC_{50}-values at least above 2 mM, by using Gal-1 and Gal-8, evidencing also an high degree of selectivity reached by these compounds.

These results encouraged us to proceed to examine these compounds in an *in vitro* system. Established human tumor lines were selected for this part of the study. Lectin binding to cells using biotinylated proteins was determined by quantitative fluorescence detection in a FACScan instrument by using streptavidin/R-phycoerythrin as indicator. In full accord with table **2.1**, the monovalent derivative **43** was not able to inhibit VAA adhesion to cells at this tested concentration, whereas lactose and the two 2-N LacNAc modified glycoclusters **53** and **54** reduced lectin binding with increasing potency. It was confirmed that tetravalent glycocluster **51** proved the most active compound against Gal-3. Therefore it constitutes a suitable scaffold to hit Gal-3, irrespective of changes in the glycomic epitope. Of note, Gal-1 binding to cells was not affected even at 10 mM sugar concentration by this calixarene (figure **2.36**), making the selection ability between Gal-1 and Gal-3 more drastic. This result is impressive considering the role played by these two lectins at biological level[39]. In fact, Gal-1 is involved in mechanisms that induce death in some tumor cells while Gal-3 interferes with this function, protecting them from the positive anoikis effect of the former galectin. It clearly appears the relevance of having an efficient inhibitor of Gal-3 that at the same time does not alter the Gal-1 activity.

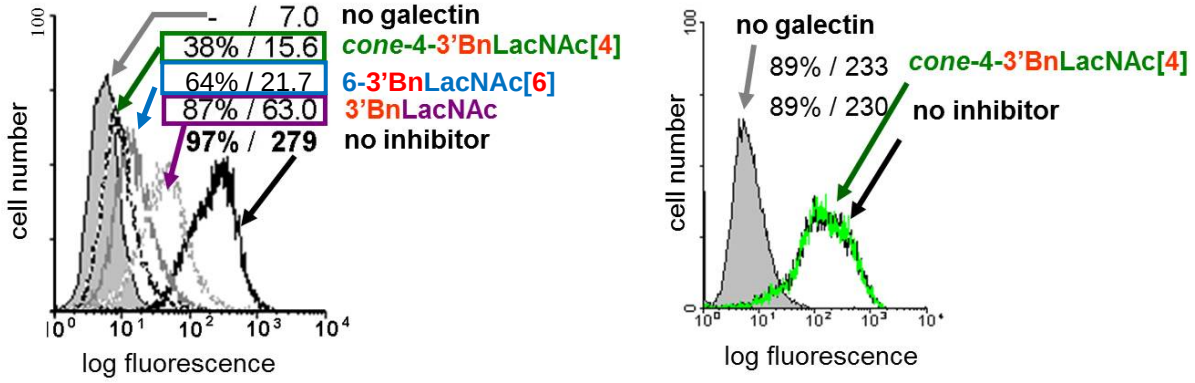


Figure 2.36 Inhibition on (left) Gal-3 (10 μg/mL) and (right) Gal-1 (10 μg/mL) binding to human colon adenocarcinoma cells. Sugar concentration was 40 and 500 μM for **51** (*cone*-4-(3'BnLacNAc)[4]Prop) and **42** (3'BnLacNAc), respectively, in test with Gal-3 and till 10 mM for **51** in test with Gal-1.

Gal-4 remained more sensitive to the hexa- than the tetravalent glycoclusters also in these experiments. The same tendency with rather similar quantitative data was observed for Gal-9, with optimal activity for the 2-N-substituted hexavalent **54** in full agreement with the data from the solid-phase assays (table **2.1**).

Overall, these results undoubtedly show the benefit of combining sugar modification with a multivalent presentation for the inhibition and selection of medical relevant lectins.

2.3 Conclusions

Since carbohydrate binding proteins, called lectins, play an important role in many biological phenomena, the research aimed at finding suitable molecules able to selectivly target them or interfere with their biological functions constitutes an important tool for analysis and possible drug development. To this aim the synthesis of several compounds has been accomplished and is described in this chapter. All these compounds presents a calixarene core functionalized at the upper rim with carbohydrate unites connected via thiourea linkers to the macrocyclic scaffold.

First of all, the synthesis of *cone*-4LacNAc[4]Prop calixarene **1** has been described. This compound was obtained through two different approaches, the first one completely chemical, the second one following a chemo-enzymatic pathway. The success of the second

route is particularly important because, for the first time, a glycosyltransferase, in particular the 1,4-galactosyltransferase, was used on a calixarene substrate leading to complete tetrafunctionalization of the starting molecule.

With the perspective in mind to further increase the complexity of the saccharide portion on calixarene backbone exploiting enzymatic reactions, sialylation reactions with 2,3- or 2,6-sialyl transferase (2,6-SialT, 2,3-SialT) were carried out on *cone*-4Lac[4]Prop **30**. With the methodologies and reagents used, complete tetrasialylation of **30** to obtain compounds **31** and **32** could not be achieved, even if evidences of partial functionalized compounds were given by HPLC and ESI-MS analyses of the studied reactions. It was hypothesized that the two enzymes have difficulties to reach the substrate because of the steric hindrance caused by the four lactosyl moieties directly connected to the calixarene core. To allow a higher flexibility of the whole molecule and hopefully a higher accessibility of sugar substrate for the sialyl transferasis enzymes, compound **41** was synthesized. This glycocluster presents a small aliphatic spacer between the lactoside moieties and the thiourea linker which hopefully could increase the flexibility of the saccharide chains which should thus results to be more accessible to the enzymes. The enzymatic sialylation of compound **41** will be attempted in the near future.

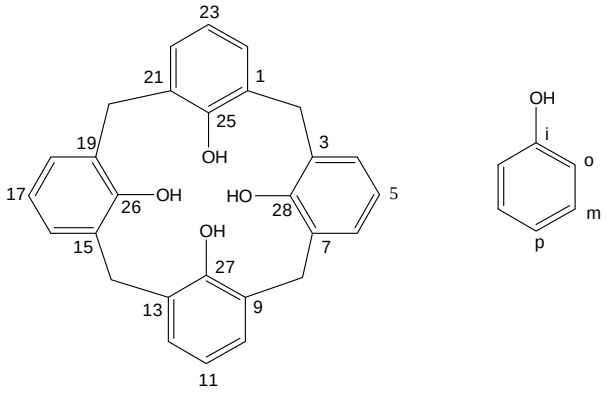
To determine the effect of structural changes in the carbohydrate units together with a clustered presentation, glycocalixarenes **51**-**54** have been synthesized.These compounds, based on *cone*-calix[4]arene and conformationally mobile calix[6]arene scaffolds, present aromatic substitutions at the 2-N or 3'-positions of N-acetyl-lactosamine. Compounds **51**-**54**, together with monovalent reference compounds **42** and **43**, were tested as lectin inhibitors on a solid-phase test and in an *in vitro* inhibition assay. In all tested cases both types of substitution showed increased inhibitory capacity compared to the analogue lactosyl clusters. Besides, multivalent glycoclusters led to much stronger and much more selective binding compared to the monovalent model ligands. The marked selectivity and efficiency shown by **51** in the inhibition of Gal-3 over Gal-1 is of particular interest because of the blocking effects of the former on a positive anoikis function of the latter towards some tumor cell lines. In general, the results obtained showed the importance of combining suitable carbohydrates to match the galectin binding sites with suitable multivalent presentation to achieve selective and potent galectin inhibition.

2.4 Experimental Part

General Information. All moisture sensitive reactions were carried out under nitrogen atmosphere, using previously oven-dried glassware. All dry solvents were prepared according to standard procedures, distilled before use and stored over 3 or 4 Å molecular sieves. Most of the solvents and reagents were obtained from commercial sources and used without further purification. Analytical TLC were performed using prepared plates of silica gel (Merck 60 F-254 on aluminum) and then, according to the functional groups present on the molecules, revealed with UV light or using staining reagents: $FeCl_3$ (1% in H_2O/CH_3OH 1:1), H_2SO_4 (5% in EtOH), ninhydrin (5% in EtOH), basic solution of $KMnO_4$ (0.75% in H_2O). Reverse phase TLC were performed by using silica gel 60 RP-18 F-254 on aluminum sheets. Merck silica gel 60 (70-230 mesh) was used for flash chromatography and for preparative TLC plates. ^1H NMR and ^{13}C NMR spectra were recorded on Bruker AV300 and Bruker AV400 spectrometers (observation of ^1H nucleus at 300 MHz and 400 MHz respectively, and of ^{13}C nucleus at 75 MHz and 100 MHz respectively). All chemical shifts are reported in part per million (ppm) using the residual peak of the deuterated solvent, which values are referred to tetramethylsilane (TMS, $\delta_{TMS} = 0$), as internal standard. All ^{13}C NMR spectra were performed with proton decoupling. For ^1H NMR spectra recorded in D_2O at 90 °C correction of chemical shifts was performed using the expression $\delta = 5.060 - 0.0122 \times T(°C) + (2.11 \times 10^{-5}) \times T^2(°C)$ [40]. Electrospray ionization (ESI) mass analyses were performed with a Waters spectrometer. HR-ESI MS spectra were recorded on a LTQ Orbitrap XL instrument in positive mode with MeOH as solvent. Melting points were determined on an Electrothermal apparatus in closed capillaries.

Nomenclature of calix[4]arene compounds.

In this thesis the simplified nomenclature proposed by Gutsche is used to name the calix[4]arene compounds. The positions on the macrocycle are numbered as indicated in the following figure.



The hydroxyl substituent defines the ipso position: subsequently the ortho, meta and para positions on the aromatic rings are identified without ambiguity.

1,2,3,4,6-penta-*O*-acetyl-β-D-galactopyranoside (2)

In a two neck round bottom flask 32.0 g of sodium acetate (0.39 mol) were suspended in 134 mL of acetic anhydride (1.38 mol) and heated to 130 °C. At this temperature D-galactose (10.0 g, 55.5 mmol) was added batchwise. The mixture was stirred at 130°C for 2 h and checked via TLC (eluent: AcOEt/hexane 1:1). When the reaction was complete it was poured into a flask containing 700 mL of ice and water (5 times the amount of acetic anhydride). The mixture was then stirred at room temperature overnight. The solid formed was recovered by filtration on a Buchner funnel and then purified via trituration in EtOH. Product **2** was obtained as white solid in 40% yield. **^{1}H NMR** (300 MHz, $CDCl_3$): δ (ppm) 5.69 (d, 1H, $J_{1\text{-}2}$ = 8.2 Hz, H_1); 5.42 (dd, 1H, $J_{4\text{-}5}$ = 0.9 Hz, $J_{4\text{-}3}$ = 3.4, H_4); 5.33 (dd, 1H, $J_{1\text{-}2}$ = 8.2 Hz, $J_{2\text{-}3}$ = 10.4 Hz, H_2); 5.07 (dd, 1H, $J_{2\text{-}3}$ = 10.4 Hz, $J_{3\text{-}4}$ = 3.4 Hz, H_3); 4.17-4.11 (m, 2H, H_{6a}, H_{6b}); 4.09-4.00 (m, 1H, H_5); 2.16 (s, 3H, Ac); 2.12 (s, 3H, Ac); 2.04 (s, 6H, 2Ac); 1.99 (s, 3H, Ac). The product shows the same spectroscopic characteristics reported in the literature[41].

2,3,4,6-tetra-*O*-acetyl-α-D-galactopyranosyl trichloroacetamidate (4)

Peracetylated galactose **2** (7.3 g, 18.7 mmol) was dissolved in DMF (50 mL) and then hydrazine acetate (2.1 g, 22.5 mmol) was added. The mixture was stirred at 50 °C for 2 h and checked via TLC (eluent: AcOEt/hexane 1:1). When completed, the reaction was quenched by addition of AcOEt (50 mL) and then washed with brine (2 x 50 mL) and water (2 x 50 mL). The organic phase was collected and concentrated under reduced pressure to give crude product **3**, a yellowish syrup, that was used without purification in the following reaction. Yield: 86% Crude product **3** (5.6 g, 16.1 mmol) was dissolved in dry CH_2Cl_2 (30 mL) under Ar

and then treated with Cs_2CO_3 (5.7 g, 17.7 mmol) and Cl_3CCN (16.1 mL, 160.8 mmol). Molecular sieves 4 Å were also added and the mixture was allowed to react at room temperature for 1.5 h under Ar. The progress of the reaction was monitored via TLC (eluent: AcOEt/hexane 1:1). The mixture was filtered over a Celite pad and the solvent removed under vacuum. The residue was purified via flash chromatography (eluent: AcOEt/cyclohexane 3:7) to give product **4** as white solid. Yield: 68%. **^{1}H NMR** (300 MHz, $CDCl_3$): δ (ppm) 8.66 (s, 1H, NH); 6.60 (d, 1H, J = 3.3 Hz, H_1); 5.56 (dd, 1H, J_{4-5} = 0.9 Hz, J_{4-3} = 3.0 Hz, H_4); 5.43 (dd, 1H, J_{2-3} = 10.8 Hz, J_{3-4} = 3.0 Hz, H_3); 5.35 (dd, 1H, J_{1-2} = 3.3 Hz, J_{2-3} = 10.8 Hz, H_2); 4.43 (td, 1H, J_{5-6}= 6.6 Hz, J_{4-5}= 0.9 Hz, H_5); 4.20-4.04 (m, 2H, H_{6a}, H_{6b}); 2.16 (s, 3H, Ac); 2.02 (s, 9H, Ac). **^{13}C NMR** (75 MHz, $CDCl_3$): δ (ppm) 170.3, 170.1, 170.0 ($\underline{C}OCH_3$); 160.9 (CN); 93.5 (C_1); 90.7 (CCl_3); 69.0 (C_5); 67.5, 67.4 (C_3, C_4); 66.9 (C_2); 61.3 (C_6); 20.7, 20.6 ($CO\underline{C}H_3$). The product presents the same physical-chemical properties reported in the literature[42].

2-Acetamido-2-deoxy-β-D-glucopyranosyl azide (5)

2-Acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-β-D-glucopyranosyl azide (2.0 g, 5.37 mmol) was dissolved in 20 mL of CH_3OH, and drops of a freshly prepared methanol solution of MeONa were added till pH 8-9. The mixture was stirred at room temperature for 1.5 hours. The reaction was monitored via TLC (CH_2Cl_2/EtOH 4:1). Amberlite resin IR 120/H^+ was subsequently added for quenching, and the mixture was gently stirred for 10 min till neutral pH. The resin was then filtered off and the solvent removed under vacuum to give product **5** as white solid in quantitative yield. **^{1}H NMR** (300 MHz, CD_3OD): δ (ppm) 4.49 (d, 1H, J = 9.2 Hz, H_1); 3.99 (dd, 1H, J_{6a-6b} = 12.1 Hz, J_{5-6a} = 1.8 Hz, H_{6a}); 3.73-3.63 (m, 2H, H_2, H_{6b}); 3.45 (t, 1H, J = 10.1 Hz, H_3); 3.36-3.29 (m, 2H, H_4, H_5); 1.98 (s, 3H, CH_3). The product presents the same physical-chemical properties reported in the literature [43,44].

2-Acetamido-6-*O*-tert-butyldiphenylsilyl-2-deoxy-β-D-glucopyranosyl azide (6)

2-Acetamido-2-deoxy-β-D-glucopyranosyl azide **5** was dissolved in 20 mL pyridine and then *tert*-butyl(chloro)diphenyl silane (2.25 mL, 8,65 mmol) was added and the mixture stirred at room temperature for 20 hours. The reaction was monitored via TLC (eluent: CH_2Cl_2/EtOH 4:1). The mixture was poured in an ice-bath and extracted with CH_2Cl_2 (3 x 30 mL). The combined organic layers were washed with HCl 0.5M (100 mL), $NaHCO_3$ saturated aqueous solution (100 mL), with $CuSO_4$ saturated aqueous solution (2 x 100 mL) and with water (3 x

100 mL). The organic extract was dried over anhydrous Na_2SO_4, filtered and concentrated under reduced pressure. The residue was purified via flash chromatography (eluent: CH_2Cl_2/ CH_3OH 94:6) and product **6** was obtained as white solid. Yield: 31%. **^{1}H NMR** (300 MHz, $CDCl_3$): δ (ppm) 7.71-7.67 (m, 4H, Ar); 7.43-7.36 (m, 6H, Ar); 5.64 (d, 1H, J = 7.1 Hz, NH); 4.69 (d, 1H, J = 9.1 Hz, H_1); 3.95 (d, 2H, J = 4.2 Hz, H_{6a}, H_{6b}); 3.77-3.66 (m, 2H, H_3, H_4); 3.55-3.45 (m, 2H, H_2, H_5); 2.08 (s, 3H, NHAc); 1.05 (s, 9H, $SiC(CH_3)_3$). **ESI-MS(+)**: m/z 507.1 [100% $(M+Na)^+$]. The product presents the same physical-chemical properties reported in the literature[12].

2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl-(1→4)-2-Acetamido-6-*O*-tert-butyldiphenyl silyl-2-deoxy-β-D-glucopyranosyl azide (7)

In a round bottom flask 30 mL of dry CH_2Cl_2 were stirred with powdered molecular sieves 4 Å for 30 min under argon atmosphere. Then galactosyl derivative **4** (0.75 g, 1.52 mmol) and glucosyl derivative **6** (0.37 g, 0.75 mmol) were added. The mixture was stirred for 2 hours at room temperature, then cooled to -45 °C. At this temperature boron trifluoride etherate (0.15 mL, 1.18 mmol) was added dropwise and the mixture stirred for 2 hours. The progress of the reaction was monitored via TLC (CH_2Cl_2/CH_3OH 10:0.3). The mixture was allowed to warm up to room temperature and then neutralized with triethylamine (0.17 mL, 1.20 mmol), diluted with CH_2Cl_2 (20 mL), filtered through a Celite pad and washed with CH_2Cl_2. The combined organic phases were concentrated under reduced pressure to give a syrup that was purified via flash chromatography (gradient from CH_2Cl_2 100% to CH_2Cl_2/CH_3OH 10:0.3). Product **7** was obtained pure as a light yellow solid. Yield: 24%. **^{1}H NMR** (300 MHz, $CDCl_3$): δ (ppm) 7.75-7.69 (m, 4H, Ar); 7.46-7.37 (m, 6H, Ar); 5.52 (d, 1H, J = 8.2 Hz, NH); 5.37 (d, 1H, $J_{3'-4'}$= 3.2 Hz, H_4'); 5.20 (dd, 1H, $J_{1'-2'}$ = 8.0 Hz, $J_{2'-3'}$ = 10.4 Hz, H_2'); 4.97 (dd, 1H, $J_{3'-4'}$ = 3.5 Hz, $J_{3'-2'}$ = 10.4 Hz, H_3'); 4.77 (d, 1H, J = 9.6 Hz, H_1); 4.72 (d, 1H, J = 8.0 Hz, H_1'); 4.14(d, 2H, J = 4.2 Hz, $H_6'_a$, $H_6'_b$); 4.00-3.78 (m, 5H, H_3, H_4, H_5', H_{6a}, H_{6b}); 3.60-3.55 (m, 1H, H_2); 3.47 (m, 1H, H_5); 2.15, 2.07, 2.05, 2.00, 1.73 (5s, 15H, OAc, NHAc); 1.08 (s, 9H, $SiC(CH_3)_3$). The product shows the same physical-chemical properties reported in the literature[12].

2-Acetamido-2-deoxy-4,6-benzyliden-β-D-glucopyranosyl azide (8)

A solution of glucosyl derivative **5** (1.30 g, 5.28 mmol) in dry DMF (15 mL) was treated with p-toluene sulfonic acid (45.5 mg, 0.26 mmol) and α,α-dimethoxytoluene (1.59 mL, 10.57 mmol). The mixture was stirred at 50 °C for 18 h under nitrogen atmosphere and monitored

with TLC (eluent: CH_2Cl_2/EtOH 4:1). The reaction mixture was then concentrated under reduced pressure and purified by flash chromatography (eluent: CH_2Cl_2/CH_3OH 10:0.5) to obtain product **8** as a pearl white solid. Yield: 90%. **^{1}H NMR** (300 MHz, CD_3OD): δ (ppm) 7.50-7.46 (m, 2H, Ar); 7.35-7.31 (m, 3H, Ar); 5.61 (s, 1H, ArCH); 4.62 (d, 1H, J = 9.1 Hz, H_1); 4.34-4.29 (dd, 1H, J = 4.5 Hz, J = 10.4 Hz, H_{6a}); 3.86-3.75 (m, 3H, H_2, H_4, H_{6b}); 3.59-3.53 (m, 2H, H_3, H_5); 1.99 (s, 3H, CH_3). **^{13}C NMR** (75 MHz, CD_3OD): δ (ppm) 174.1 ($COCH_3$); 139.3, 130.3, 129.3, 127.8 (Ar); 103.3 (ArCH); 90.8 (C_1); 82.8, 72.6, 70.2, 69.6, (C_3-C_6) 57.5 (C_2); 23.1 (CH_3). **ESI-MS(+)**: m/z 357.1 [100% $(M+Na)^+$]; 691.4 [30% $(2M+Na)^+$]. **M.p.** : 211-212.5 °C dec.

2-Acetamido-2-deoxy-3-*O*-acetyl-4,6-benzyliden-β-D-glucopyranosyl azide (9)

Compound **8** (1.58 g, 4.73 mmol) was dissolved in 10 mL of anhydrous pyridine and treated with acetic anhydride (5 mL). The mixture was stirred at room temperature for 1 h under N_2 atmosphere. The reaction was monitored via TLC (eluent: CH_2Cl_2/CH_3OH 95:5) and when finished, quenched by dilution in CH_2Cl_2 (30 mL), washed with a $CuSO_4$ saturated aqueous solution (30 mL), and water (2 x 30 mL). The combined aqueous layers were extracted with CH_2Cl_2 (2 x 50 mL) and then the combined organic phases washed with brine (50 mL), water (2 x 50 mL), dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo* to give product **9** as white solid in quantitative yield. **^{1}H NMR** (300 MHz, CD_3OD): δ (ppm) 7.43-7.39 (m, 2H, Ar); 7.35-7.31 (m, 3H, Ar); 5.59 (s, 1H, ArCH); 5.26 (t, 1H, J = 9.8 Hz, H_3); 4.82 (d, 1H, J = 9.8 Hz, H_1); 4.34 (dd, 1H, $J_{5\text{-}6a}$ = 4.8 Hz, $J_{6a\text{-}6b}$ = 10.2 Hz,H_{6a}); 3.93 (t, 1H, J = 9.8 Hz, H_2); 3.89-3.77 (m, 2H, H_4, H_{6b}); 3.70-3.60 (m, 1H, H_5); 2.01, 1.93 (2s, 6H, CH_3). **^{13}C NMR** (75 MHz,$CDCl_3$): δ (ppm) 171.5, 170.3 ($COCH_3$); 136.6, 129.2, 128.2, 125.9 (Ar); 101.3 (ArCH); 89.4 (C_1); 78.2, 71.6, 68.3, 68.1 (C_3-C_6); 53.8 (C_2); 23.2, 20.9 (CH_3). **ESI-MS(+)**: m/z 399.3 [100% $(M+Na)^+$]; 775.4 [50% $(2M+Na)^+$]. **M.p.** : 213-215 °C dec.

2-Acetamido-2-deoxy-3-*O*-acetyl-β-D-glucopyranosyl azide (10)

Compound **9** (1.6 g, 4.25 mmol) was treated with 13 mL of the system TFA/H_2O/CH_2Cl_2 (1:2:10, v/v/v) at 0 °C. The mixture was stirred at 0 °C for 4 hours and then it was let reach room temperature and stirred for another hour. The reaction was checked via TLC (eluent CH_2Cl_2/CH_3OH 95:5) and once finished it was quenched by addition of a few drops of pyridine to neutral pH. The solvent was removed under reduced pressure and the residue purified via flash chromatography (elution in gradient:CH_2Cl_2/CH_3OH from 95:5 to 9:1) to

give product **9** as a white solid in quantitative yield. **^{1}H NMR** (300 MHz, CD_3OD): δ (ppm) 8.07 (d, 1H, J = 9.3 Hz, NH); 5.04 (t, 1H, J = 9.8 Hz, H_3); 4.77 (d, 1H, J = 9.3 Hz, H_1); 3.90 (dd, 1H, $J_{5\text{-}6a}$ = 1.4 Hz, $J_{6a\text{-}6b}$ = 12.2 Hz, H_{6a}); 3.79 (t, 1H, J = 9.8 Hz, H_2); 3.74 (dd, 1H, $J_{5\text{-}6b}$ = 4.9 Hz, $J_{6a\text{-}6b}$ = 12.2 Hz, H_{6b}); 3.58 (t, 1H, J = 9.8 Hz, H_4); 3.55-3.47 (m, 1H, H_5); 2.03, 1.92 (2s, 6H, CH_3). **^{13}C NMR** (75 MHz,CD_3OD): δ (ppm) 174.0, 172.9 ($\underline{C}OCH_3$); 90.0 (C_1); 80.3, 77.1, 69.7 (C_3-C_5); 62.6 (C_6); 55.3 (C_2); 23.3, 21.4 (CH_3). **ESI-MS(+)**: m/z 311.2 [100% $(M+Na)^+$]; 599.3 [60% $(2M+Na)^+$]. **M.p.** : 105-110 °C.

2-Acetamido-2-deoxy-3-*O*-acetyl-6-*O*-tert-butyldiphenylsilyl-β-D-glucopyranosyl azide (11)

tert-Butyl(chloro)diphenyl silane (1.82 mL,7.00 mmol) was added to a solution of compound **10** (1.15 g, 4.00 mmol) in pyridine (15 mL). The mixture was stirred at room temperature for 2 days and monitored via TLC (eluent: CH_2Cl_2/CH_3OH 95:5). The mixture was then poured into icy water (50 mL) and extracted with CH_2Cl_2 (2 x 30 mL). The combined organic phases were washed with HCl 1M (2 x 50 mL), $NaHCO_3$ saturated aqueous solution (50 mL) and water to neutral pH. The solvent was then removed under reduced pressure and the residue purified via flash chromatography (CH_2Cl_2/CH_3OH 10:0.3) to give product **11** as white solid. Yield: 77%. **^{1}H NMR** (300 MHz, $CDCl_3$): δ (ppm) 7.70-7.65 (m, 4H, Ar); 7.43-7.25 (m, 6H, Ar); 6.03 (d, 1H, J = 9.2 Hz, NH); 5.07 (t, 1H, $J_{3\text{-}4}$ = 9.2 Hz, $J_{2\text{-}3}$ = 10.4 Hz, H_3); 4.54 (d, 1H, J = 9.2 Hz, H_1); 3.98-3.92 (m, 3H, H_2, H_{6a}, H_{6b}); 3.89 (t, 1H, J = 9.2 Hz, H_4); 3.55-3.50 (m, 1H, H_5); 2.12, 1.99 (2s, 6H, $COCH_3$); 1.05 (s, 9H, $SiC(CH_3)_3$). **^{13}C NMR** (75 MHz,$CDCl_3$): δ (ppm) 172.1, 170.7 ($\underline{C}OCH_3$); 135.6, 132.8, 129.9, 127.8 (Ar); 88.4 (C_1); 77.4 (C_5); 75.3 (C_3); 69.7 (C_4); 63.7 (C_6); 53.6 (C_2); 26.8 ($SiC(\underline{C}H_3)_3$); 23.2, 21.0 ($CO\underline{C}H_3$), 19.2 ($Si\underline{C}(CH_3)_3$). **ESI-MS(+)**: m/z 549.4 [100% $(M+Na)^+$]. **M.p.** : 165-166 °C.

2,3,4,6-Tetra-*O*-acetyl-β-D-galactopyranosyl-(1→4)-2-acetamido-3-*O*-acetyl-6-*O*-tert-butyl diphenylsilyl-2-deoxy-β-D-glucopyranosyl azide (12)

In a round bottom flask 30 mL of dry CH_2Cl_2 were stirred with powdered molecular sieves 4 Å for 30 min under argon atmosphere. Then galactosyl derivative **4** (1.12 g, 2.28 mmol) and glucosyl derivative **11** (0.80 g, 1.52 mmol) were added. The mixture was stirred for 2 hours at room temperature, then cooled to -45 °C. At this temperature boron trifluoride etherate (0.29 mL, 2.28 mmol) was added dropwise and the mixture stirred for 4 hours. The progress of the reaction was monitored via TLC (AcOEt/cyclohexane 1.1). The mixture was allowed to warm up to room temperature and then neutralized with triethylamine (0.32 mL, 2.28

mmol), diluted with CH_2Cl_2 (20 mL), filtered through a Celite pad and washed with CH_2Cl_2. The combined organic phases were concentrated under reduced pressure to give a syrup that was purified via flash chromatography (eluent: CH_2Cl_2/CH_3OH 98:2). Product **12** was obtained as white solid. Yield: 69%. **^{1}H NMR** (300 MHz, $CDCl_3$): δ (ppm) 7.73-7.68 (m, 4H, Ar); 7.44-7.37 (m, 6H, Ar); 6.13 (d, 1H, J = 9.4 Hz, NH); 5.30 (d, 1H, $J_{3'-4'}$ = 3.4 Hz, H_4'); 5.05-4.97 (m, 2H, H_3, H_2'); 4.88 (dd, 1H, $J_{3'-4'}$ = 3.4 Hz, $J_{3'-2'}$ = 10.2 Hz, H_3'); 4.78 (d, 1H, J = 7.9 Hz, H_1'); 4.45 (d, 1H, J = 9.2 Hz, H_1); 4.18 (t, 1H, J = 9.4 Hz, H_4); 4.17-4.05 (m, 3H, H_2, $H_6'_a$, $H_6'_b$); 3.96 (dd, 1H, J_{5-6a}= 1.2 Hz, J_{6a-6b} = 11.6 Hz, H_{6a}); 3.86 (dd, 1H, J_{5-6b} = 1.9 Hz, J_{6a-6b} = 11.6 Hz, H_{6b}); 3.74 (t, 1H, J = 6.6 Hz, H_5'); 3.37 (d, 1H, J = 9.4 Hz, H_5); 2.12, 2.06, 1.97, 1.96, 1.73 (5s, 18H, Ac); 1.07 (s, 9H, $SiC(CH_3)_3$). **^{13}C NMR** (75 MHz,$CDCl_3$): δ (ppm) 171.3, 170.4, 170.3, 170.1, 169.9, 168.7 ($\underline{C}OCH_3$); 135.8, 135.2, 133.1, 131.6, 130.0, 128.0, 127.8 (Ar); 100.1 (C_1'), 88.6 (C_1); 76.6 (C_5); 73.5 (C_4); 72.5, 69.2 (C_3, C_2'); 70.8, 70.7 (C_3', C_5'); 66.9 (C_4'); 61.1, 60.0 (C_6, C_6'); 53.2 (C_2); 26.7 ($SiC(\underline{C}H_3)_3$); 23.1, 20.8, 20.6, 20.5, 20.4, 20.3, 20.2 ($CO\underline{C}H_3$), 19.3 ($Si\underline{C}(CH_3)_3$). **ESI-MS(+)**: m/z 879.6 [100% $(M+Na)^+$]; 1737.3 [10% $(2M+Na)^+$]. **M.p.** : 58.8-61.0 °C. The spectroscopic data found are in agreement with those reported in literature[12].

General desilylation procedure of disaccharides 7 and 12 to obtain compounds 13a and 13, respectively

To a mixture of silyl disaccharide in dry THF, acetic acid and a 1M solution of n-tetrabutyl ammonium fluoride in dry THF were added. The mixture was stirred at room temperature for 4 days under nitrogen atmosphere and monitored via TLC. The solvent was then removed under vacuum, the residue dissolved in CH_2Cl_2 and washed with water (2x). The aqueous phase was extracted with CH_2Cl_2 (2x) and the combined organic phases dried over Na_2SO_4 anhydrous, filtered and concentrated under reduced pressure. Purification via flash chromatography gave the desired products as white solids.

2,3,4,6-Tetra-*O*-acetyl-β-D-galactopyranosyl-(1→4)-2-acetamido-2-deoxy-β-D-gluco pyranosyl azide (13a)

Compound **7** (0.27 g, 0.33 mmol), acetic acid (130 µL, 2.30 mmol) and n-tetrabutyl ammonium fluoride (0.52 g, 1.64 mmol) in 15 mL dry THF. TLC (eluent: CH_2Cl_2/CH_3OH 10:0.8). Flash chromatography (eluent: CH_2Cl_2/CH_3OH 10:0.2). Yield of product **13a**: 63%.**^{1}H NMR** (300 MHz, $CDCl_3$): δ (ppm) 5.80 (d, 1H, J = 8.2 Hz, NH); 5.37 (d, 1H, $J_{3'-4'}$ = 3.3 Hz, H_4');

5.20 (dd, 1H, $J_{1'-2'}$ = 8.0 Hz, $J_{2'-3'}$ = 10.4 Hz, H_2'); 5.04 (dd, 1H, $J_{3'-4'}$ = 3.3 Hz, $J_{3'-2'}$ = 10.4 Hz, H_3'); 4.87 (d, 1H, J = 9.3 Hz, H_1); 4.67 (d, 1H, J = 8.0 Hz, H_1'); 4.13-4.03 (m, 4H, OH_3, H_5', $H_6'_a$, $H_6'_b$); 3.87-3.82 (m, 2H, H_3, H_{6a},); 3.66-3.60 (m, 2H, H_4, H_{6b}); 3.59-3.45 (m, 2H, H_2, H_5); 2.40 (bs, 1H, OH_6); 2.14, 2.09, 2.04, 2.02, 1.97 (5s, 15H, OAc, NHAc). **^{13}C NMR** (75 MHz,$CDCl_3$): δ (ppm) 170.8, 170.4, 170.0, 169.9, 169.4 ($\underline{C}OCH_3$); 101.7 (C_1'); 88.2 (C_1); 80.9 (C_4); 76.3 (C_5); 71.8 (C_3); 71.3 (C_5'); 70.7 (C_3'); 68.8 (C_2'); 66.8 (C_4'); 61.4 (C_6'); 60.6 (C_6,); 55.8 (C_2); 23.3, 20.6, 20.5, 20.4 ($CO\underline{C}H_3$). **ESI-MS(+)**: m/z 599.1 [100% $(M+Na)^+$].

2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl-(1→4)-2-Acetamido-2-deoxy-3-*O*-acetyl-β-D-glucopyranosyl azide (13)

Compound **12** (1.70 g, 1.98 mmol, acetic acid (950 µL, 16.6 mmol) and n-tetrabutyl ammonium fluoride (3.75 g, 11.9 mmol)) in 60 mL dry THF. TLC (eluent: AcOEt/hexane 7:3). Flash chromatography (gradient AcOEt/hexane from 7:3 to 9:1). Yield of product **13**: 84%. **^{1}H NMR** (300 MHz, $CDCl_3$): δ (ppm) 6.31 (d, 1H, J = 9.4 Hz, NH); 5.30 (dd, 1H, $J_{3'-4'}$ = 3.2 Hz, $J_{4'-5'}$ = 0.7 Hz, H_4'); 5.09-4.95 (m, 3H, H_3, H_2', H_3'); 4.61 (d, 1H, J = 7.4 Hz, H_1'); 4.57 (d, 1H, J = 9.3 Hz, H_1); 4.09-3.86 (m, 6H,H_2, H_4, H_{6a}, H_5', $H_6'_a$, $H_6'_b$); 3.70 (dd, 1H, J_{6a-6b} = 12.4 Hz, J_{5-6b} = 3.4 Hz, H_{6b}); 3.43 (m, 1H, H_5); 2.40 (bs, 1H, OH_6); 2.11, 2.06, 2.03, 2.02, 1.95, 1.93 (6s, 21H, Ac). **^{13}C NMR** (75 MHz,$CDCl_3$): δ (ppm) 171.1, 170.5, 170.3, 170.0, 169.9, 169.3 ($\underline{C}OCH_3$); 101.0 (C_1'); 88.7 (C_1); 76.6, 74.8, 73.2, 70.7, 70.5, 69.2, 66.7 (C_3-C_5, C_2'-C_5'); 60.8, 60.3 (C_6, C_6'); 53.2 (C_2); 23.0, 20.8, 20.6, 20.5, 20.4 ($CO\underline{C}H_3$). **ESI-MS(+)**: m/z 641.3 [100% $(M+Na)^+$]. **M.p.** : 105.5-107.5 °C.

2,3,4,6-Tetra-*O*-acetyl-β-D-galactopyranosyl-(1→4)-2-acetamido-2-deoxy-3,6-di-*O*-acetyl-β-D-glucopyranosyl azide (14)

Compound **13a** (120 mg, 0.21 mmol) was dissolved in 2.0 mL pyridine and then 1.0 mL of acetic anhydride was added. The mixture was stirred at room temperature for 18 hours and monitored via TLC (eluent: CH_2Cl_2/CH_3OH 95:5). The mixture was then diluted with CH_2Cl_2 (20 mL), washed with $CuSO_4$ saturated aqueous solution (20 mL) and water (3 x 15 mL). The combined aqueous phases were extracted with CH_2Cl_2 (2 x 15 mL) and then the organic layers collected together, washed with brine (20 mL), water (2 x 20 mL) dried over anhydrous $NaSO_4$ and the solvent removed under reduced pressure to give compound **14** as a white solid. Yield: 98%.

Same procedure for compound **13** (1.10 g, 1. 78 mmol), in 6 mL pyridine and 3 mL acetic anhydride. Reaction time: 3 hours. Yield 98%.

^{1}H NMR (300 MHz, $CDCl_3$): δ (ppm) 6.01 (d, 1H, J = 9.4 Hz, NH); 5.34 (dd, 1H, $J_{3'-4'}$ = 3.4 Hz, $J_{4'-5'}$ = 0.7 Hz, H_4'); 5.10-5.03 (m, 2H, H_3, H_2'); 4.95 (dd, 1H, $J_{3'-4'}$ = 3.4 Hz, $J_{2'-3'}$ = 10.4 Hz, H_3'); 4.52-4.48 (m, 3H, H_1, H_1', H_{6a}); 4.12-4.04 (m, 4H, H_2, H_{6a}, $H_6'_a$, $H_6'_b$); 3.87 (td, 1H, $J_{4'-5'}$ = 0.7 Hz, $J_{5'-6'}$ = 7.0 Hz, H_5'); 3.80 (t, 1H, J = 9.3 Hz, H_4); 3.70-3.65 (m, 1H, H_5); 2.14, 2.12, 2.08, 2.05, 2.04, 1.96, 1.94 (7s, 21H, Ac). **^{13}C NMR** (75 MHz,$CDCl_3$): δ (ppm) 170.9, 170.3, 170.0, 169.9, 169.2 ($\underline{C}OCH_3$); 101.2 (C_1'); 83.3 (C_1); 75.8 (C_4); 74.5 (C_5); 72.8, 69.0 (C_2', C_3); 70.7, 70.6 (C_3', C_5'); 66.4 (C_4'); 61.8 (C_6); 60.6 (C_6'); 53.0 (C_2); 23.0, 20.8, 20.7, 20.5, 20.4 ($CO\underline{C}H_3$). **ESI-MS(+)**: m/z 683.1 [100% $(M+Na)^+$]. **M.p.** : 90-93 °C. The product shows the same physical-chemical properties reported in the literature[45].

2,3,4,6-Tetra-*O*-acetyl-β-D-galactopyranosyl-(1→4)-2-acetamido-2-deoxy-3,6-di-*O*-acetyl-β-D-glucopyranosyl amine (15)

Azido compound **14** (200 mg, 0.30 mmol) was dissolved in 30 mL of absolute ethanol and then Pd/C (10%) was added in catalytic amount. The mixture was reacted at the Parr apparatus at 1.5 bar of H_2, at room temperature for 3 hours. Progression of the reaction was checked via TLC (eluent: CH_2Cl_2/CH_3OH 9:1). The catalyst was then filtered off and the solvent removed under reduced pressure to give product **15** as a white solid that was used immediately without purification in the following reaction. Yield: quantitative. **^{1}H NMR** (300 MHz, $CDCl_3$): δ (ppm) 6.25 (d, 1H, J = 9.3 Hz, NHAc); 5.26 (d, 1H, $J_{3'-4'}$ = 3.3 Hz, H_4'); 5.04-4.85 (m, 3H, H_3, H_2', H_3'); 4.45-4.35 (m, 2H, H_1, H_1'); 4.04-3.80 (m, 6H, H_2, H_{6a}, H_{6b},H_5', $H_6'_a$, $H_6'_b$); 3.66 (t, 1H, J = 9.4 Hz, H_4); 3.49-3.44 (m, 1H, H_5); 2.23 (bs, 2H, NH_2); 2.07, 2.04, 2.00, 1.98, 1.97, 1.89, 1.88 (7s, 21H, Ac). **^{13}C NMR** (75 MHz,$CDCl_3$): δ (ppm) 171.1, 170.9, 170.4, 170.2, 170.0, 169.9, 169.1 ($\underline{C}OCH_3$); 101.0 (C_1'); 85.8 (C_1); 76.6, 73.7, 73.3, 70.8, 70.5, 70.0, 66.5 (C_3-C_5, C_2'-C_5'); 62.4, 60.6 (C_6, C_6'); 54.6 (C_2); 23.1, 20.8, 20.5, 20.4 ($CO\underline{C}H_3$). **ESI-MS(+)**: m/z 657.3 [100% $(M+Na)^+$]. **M.p.** : 92-96 °C.

2,3,4,6-Tetra-*O*-acetyl-β-D-galactopyranosyl-(1→4)-2-acetamido-2-deoxy-3,6-di-*O*-acetyl-β-D-glucopyranosyl isothiocyanate (16)

A sample of $CaCO_3$ (90 mg, 0.90 mmol) was added to a suspension of amino derivative **15** (192 mg, 0.30 mmol) in $CHCl_3$ (3 mL) and water (1 mL). The mixture was cooled to 0 °C and

then thiophosgene (35 µL, 0.45 mmol) was added. The mixture was stirred at room temperature for 2 hours and checked via TLC (eluent: CH_2Cl_2/CH_3OH 9:1). The solid was then filtered off, the mixture diluted with CH_2Cl_2 (30 mL) and washed with water (3 x 20 mL). The organic phase was recovered and the solvent removed under reduced pressure. The residue was purified via column chromatography (eluent: CH_2Cl_2/CH_3OH 98:2) to give product **16** as a white solid. Yield: 64%. **^{1}H NMR** (300 MHz, $CDCl_3$): δ (ppm) 6.00 (d, 1H, J = 9.4 Hz, NH); 5.35 (dd, 1H, $J_{3'-4'}$ = 3.2 Hz, $J_{4'-5'}$ = 0.6 Hz, H_4'); 5.14-5.05 (m, 2H, H_3, H_2'); 5.00-4.93 (m, 2H, H_1, H_3'); 4.51-4.46 (m, 2H, H_1', H_{6a}); 4.20 (q, 1H, J = 9.4 Hz, H_2); 4.16-4.03 (m, 3H, H_{6a}, $H_6'_a$, $H_6'_b$); 3.88 (t, 1H, J = 7.0 Hz, H_5'); 3.80 (t, 1H, J = 8.6 Hz, H_4); 3.70-3.65 (m, 1H, H_5); 2.15, 2.13, 2.10, 2.06, 2.05, 2.00, 1.96 (7s, 21H, Ac). **^{13}C NMR** (75 MHz,$CDCl_3$): δ (ppm) 171.0, 170.4, 170.3, 170.2, 170.0, 169.9, 169.3 ($\underline{C}OCH_3$); 142.6 (CS); 101.3 (C_1'); 83.9 (C_1); 75.7 (C_4); 74.3 (C_5); 72.6, 68.9 (C_2', C_3), 70.7 (C_3', C_5'); 66.4 (C_4'); 61.7 (C_6); 60.6 (C_6'); 54.5 (C_2); 23.0, 20.8, 20.7, 20.6, 20.4 ($CO\underline{C}H_3$). **ESI-MS(+)**: m/z 699.3 [100% $(M+Na)^+$]. **M.p.** : 98.5-102.0 °C.

Cone-5,11,17,23-tetrakis[(2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl-(1→4)-2-acetamido-2-deoxy-3,6-di-_O_-acetyl-β-D-glucopyranosyl)thioureido]-25,26,27,28-tetrapropoxycalix[4]arene (21)

Isothiocyanate derivative **16** (500 mg, 0.75 mmol) was added to a solution of amino-calix[4]arene **20**[21] (96 mg, 0.15 mmol) in 15 mL of dry CH_2Cl_2. Et_3N (82 µL, 0.60 mmol) was also added and then the mixture was stirred at room temperature for 18 h under N_2 atmosphere. Reaction checked via TLC (eluent: CH_2Cl_2/CH_3OH 95:5). The solvent was then removed under vacuum and the residue purified via column chromatography (eluent: AcOEt/hexane/CH_3OH 5:5:1) to give product **21** as a light yellow solid. Yield: 72%. **^{1}H NMR** (300 MHz, CD_3OD): δ (ppm) 6.74 (bs, 4H, Ar); 6.67 (bs, 4H, Ar); 5.76 (d, 4H, J = 9.6 Hz, H_1); 5.34 (d, 4H, J = 3.2 Hz, H_4'); 5.15 (t, 4H, J = 9.7 Hz, H_3); 5.09 (dd, 4H, $J_{2'-3'}$ = 10.4 Hz, $J_{3'-4'}$ = 3.2 Hz, H_3'); 5.00 (dd, 4H, $J_{1'-2'}$ = 7.6 Hz, $J_{2'-3'}$ = 10.4 Hz, H_2'); 4.70 (d, 4H, J = 7.6 Hz, H_1'); 4.59-4.53 (m, 4H, H_{6a}); 4.45 (d, 4H, J = 12.7 Hz, *ax*-Ar$C\underline{H}_2$Ar); 4.19-4.04 (m, 20H, H_2, H_{6b}, H_5', $H_6'_a$, $H_6'_b$); 3.94-3.88 (m, 12H, H_4, $OC\underline{H}_2CH_2CH_3$); 3.79-3.75 (m, 4H, H_5); 3.24 (d, 4H, J = 12.7 Hz, *eq*-Ar$C\underline{H}_2$Ar); 2.14 (s, 12H, Ac), 2.11-1.98 (m, 44H, 3Ac, $OCH_2C\underline{H}_2CH_3$); 1.94 (s, 12H, Ac); 1.92 (s. 12H, Ac); 1.04 (t, 12H, J = 7.4 Hz, $OCH_2CH_2C\underline{H}_3$). **^{13}C NMR** (75 MHz, CD_3OD): δ (ppm) 183.0 (CS); 173.8, 172.4, 172.0, 171.5, 171.2 (CO); 156.0 (Ar-ipso); 136.7 (Ar-ortho); 132.6 (Ar-para); 126.2, 125.8 (Ar-meta); 101.9 (C_1'); 84.7 (C_1); 78.3, 77.5, 75.6, 74.8, 72.5, 71.6, 70.7,

68.5 (C_3-C_5, C_2'-C_5', O<u>C</u>$H_2CH_2CH_3$); 63.8, 62.1 (C_6, C_6'); 54.1 (C_2); 31.8 (Ar<u>C</u>H_2Ar); 24.4 (OCH_2<u>C</u>H_2CH_3); 23.1, 21.4, 21.1, 20.9, 20.8, 20.6, 20.5 (CO<u>C</u>H_3); 10.8 (OCH_2CH_2<u>C</u>H_3). **ESI-MS(+)**: m/z 1702.8 [100% $(M+2Na)^{2+}$]. **M.p.** : 190-192 °C dec.

***Cone*-5,11,17,23-tetrakis[(β-D-galactopyranosyl-(1→4)-2-Acetamido-2-deoxy-β-D-glucopyrano syl)thioureido]-25,26,27,28-tetrapropoxycalix[4] arene (1)**

Peracetylated-lactosyl-calix[4]arene **21** (350 mg, 0.10 mmol) was dissolved in 7 mL of CH_3OH, and drops of a freshly prepared methanol solution of MeONa were added till pH 8-9. The mixture was stirred at room temperature for 4 hours. The progress of the reaction was checked via ESI-MS analysis. Since a precipitate was formed, drops of water were added to solubilize it. Amberlite resin IR 120/H^+ was subsequently added for quenching, and the mixture was gently stirred for 30 min till neutral pH. The resin was filtered off and the solvent removed *in vacuo* to obtain product **1** as light yellow solid. Yield: 90%. **^{1}H NMR** (400 MHz, CD_3OD/D_2O 5:1): δ (ppm) 6.82 (bs, 4H, Ar); 6.65 (bs, 4H, Ar); 5.61 (bs, 4H, H_1); 4.64 (bs, 8H, H_1', *ax*-Ar<u>C</u>H_2Ar); 3.95-3.85 (m, 24H); 3.84-3.65 (m, 20H); 3.64-3.55 (m, 12H); 3.24 (bs, 4H, *eq*-Ar<u>C</u>H_2Ar); 2.06-1.98 (m, 20H, Ac,OCH_2<u>C</u>H_2CH_3); 1.05 (t, 12H, J = 7.0 Hz, OCH_2CH_2<u>C</u>H_3). **^{13}C NMR** (100 MHz, CD_3OD/D_2O 5:1): δ (ppm) 179.9 (CS); 172.0 (<u>C</u>OCH_3); 153.2 (Ar-ipso); 133.9 (Ar-ortho); 129.9 (Ar-para); 123.1 (Ar-meta); 102.0 (C_1'); 82.3 (C_1); 75.4 (O<u>C</u>$H_2CH_2CH_3$); 77.6, 75.1, 74.1, 71.7, 71.4, 69.7, 67.3 (C_3 - C_5, C_2'- C_5'); 59.6, 58.8 (C_6, C_6'); 52.7 (C_2); 28.9 (Ar<u>C</u>H_2Ar); 21.46 (OCH_2<u>C</u>H_2CH_3); 20.5 (CO<u>C</u>H_3); 8.0 (OCH_2CH_2<u>C</u>H_3). **ESI-MS(+)**: m/z 1197.9 [100% $(M+2Na)^{2+}$]. **M.p.** > 190 °C dec.

2-Acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-α-D-glucopyranosyl chloride (23)

N-acetyl-glucosamine **22** (2 mL, 9 mmol) was slowly added in a round bottom flask containing acetyl chloride (4 mL, 56 mmol). The mixture was stirred at room temperature for 24 hours. TLC (eluent: AcOEt/hexane 6:4) were performed to control the progress of the reaction. The reaction was quenched by addition of CH_2Cl_2 (20 mL) and then poured into an ice-water mixture (30 mL). The organic phase was extracted, washed with water (20 mL), $NaHCO_3$ saturated aqueous solution (20 mL), again water till neutral pH, dried over anhydrous Na_2SO_4, filtered and the solvent removed under reduced pressure. The crude was purified by flash chromatography (elution in gradient: AcOEt/hexane 6:4 → 8:2). Product **23** was obtained pure as a white solid. Yield: 43%. **^{1}H NMR** (300 MHz, $CDCl_3$): δ

(ppm) 6.26 (d, 1H, J = 9.0 Hz, NH); 6.10 (d, 1H, J = 3.0 Hz, H_1); 5.26 (t, 1H, J = 9.0 Hz, H_3); 5.10 (t, 1H, J = 9.0 Hz, H_4); 4.50-4.42 (m, 1H, H_2); 4.22-4.16 (m, 2H, H_5, H_{6a}); 4.05-4.00 (m, 1H, H_{6b}); 2.00, 1.95, 1.90 (3s, 12H, CH_3CO). **^{13}C NMR** (100 MHz, $CDCl_3$): δ (ppm) 171.1, 170.4, 170.3, 169.0 (C=O); 93.7 (C_1); 70.7 (C_5); 69.6 (C_3); 67.1 (C_4); 61.1 (C_6); 53.2 (C_2); 22.8 ($NHCOCH_3$); 20.5, 20.4 (CH_3CO). **ESI-MS(+)**: m/z 388.3 [100% $(M+Na)^+$]; m/z 753.4[94% $(2M+Na)^{2+}$]. The product presents the same spectroscopic characteristics reported in literature[23].

2-Acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-β-D-glucopyranosyl azide (24)

GlcNAc chloride **23** (1 g, 2.7 mmol), NaN_3 (0.53 g, 8.2 mmol) and n-BuN_4HSO_4 (0.93 g, 2.7 mmol) were suspended in 10 mL CH_2Cl_2. A $NaHCO_3$ saturated aqueous solution (10 mL) was added and the mixture was vigorously stirred for 2 h at room temperature. The progression of the reaction was checked via TLC (eluent: AcOEt/CH_2Cl_2 8:2). Subsequently AcOEt (50 mL) was added to the mixture to quench the reaction. The organic phase was extracted, washed with water (30 mL), $NaHCO_3$ aqueous saturated solution (30 mL), brine (30 mL), dried over anhydrous Na_2SO_4, filtered and concentrated under reduced pressure to give product **23** as a pale yellow oil. Yield: 95%. **^{1}HNMR** (300 MHz, $CDCl_3$): δ (ppm) 5.54 (d, 1H, J = 8.7 Hz, NH); 5.23 (t, 1H, J = 9.6 Hz, H_3); 5.09 (t, 1H, J = 9.6 Hz, H_4); 4.74 (d, 1H, J = 9.3 Hz, H_1); 4.26 (dd, 1H, $J_{5\text{-}6a}$ = 4.5 Hz, $J_{6a\text{-}6b}$ = 12.3 Hz, H_{6a}); 4.15 (dd, 1H, $J_{5\text{-}6b}$ = 2.4 Hz, $J_{6a\text{-}6b}$ = 12.3 Hz, H_{6b}); 3.95-3.85 (m, 1H, H_2); 3.80-3.74 (m, 1H, H_5); 2.09, 2.04, 2.03, 1.97 (4s, 12H, $COCH_3$). The product presents the same spectroscopic characteristics reported in literature[46].

2-Acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-β-D-glucopyranosyl amine (25)

Compound **24** (0.97 g, 2.6 mmol) was dissolved in 40 mL CH_3CN. Under nitrogen flow a catalytic amount of Pd/C (10%) was added. The mixture was reacted under H_2 atmosphere (pressure: 1.5 bar) for 16 h at room temperature using the Parr apparatus. The reaction was monitored via TLC (eluent CH_2Cl_2/CH_3OH 9:1). The catalyst was then filtered off and the solvent removed under vacuum giving product **25** as a pale yellow solid. Yield: 78%. **^{1}H NMR** (300 MHz, $CDCl_3$): δ (ppm) 5.60 (d, 1H, J = 9.0 Hz, $NHCOCH_3$); 5.11-4.99 (m, 2H, H_3, H_4); 4.20 (dd, 1H, $J_{5\text{-}6a}$ = 4.8 Hz, $J_{6a\text{-}6b}$ = 12.3 Hz, H_{6a}); 4.13-4.07 (m, 2H, H_{6b}, H_1); 4.07-3.99 (m, 1H, H_2); 3.65-3.60 (m, 1H, H_5); 2.09, 2.05, 2.02,1.96 (4s, 12H, $COCH_3$). **ESI-MS(+)**: m/z 369.1 [100% $(M+Na)^+$]; m/z 715.3 [70% $(2M+Na)^+$]. The product presents the same spectroscopic characteristics reported in literature[46].

2-Acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-β-D-glucopyranosyl isothiocyanate (26)

Method A) To a solution of glucosyl amine **25** (700 mg, 2.02 mmol) in 8 mL $CHCl_3$, 6 mL of water and $CaCO_3$ (0.61 g, 6.06 mmol) were added. The mixture was cooled with an ice-water bath and then thiophosgene (0.23 mL, 3.04 mmol) was slowly added. The biphasic mixture was stirred at room temperature for 6 hours and constantly monitored with TLC (eluent: AcOEt/hexane 6:4). Subsequently the solid was filtered off, and the organic phase washed with water (3 x 10 mL), dried over anhydrous Na_2SO_4 and filtered. The solvent was removed under reduced pressure to give product **26** as a pale yellow solid. Yield: 81%.

Method B) In a round bottom flask KSCN (100 mg, 1.2 mmol) and n-BuN_4HSO_4 (185 mg, 0.55 mmol) were dissolved in 25 mL CH_3CN. Molecular sieves 4 Å were added to the solution and the mixture was stirred at room temperature for 40 min under N_2 atmosphere. Subsequently glucosyl chloride **23** (200 mg, 0.55 mmol) was added and the mixture was refluxed for 18 h under nitrogen atmosphere. The reaction was monitored via TLC (eluent: AcOEt/hexane 1:1) and once finished it was filtered on a Celite bed to remove the solid present and then concentrated to dryness under reduced pressure. The residue was purified on flash column chromatography (elution in gradient: AcOEt/hexane 1:1 → 4:1) giving product **26** as a pale yellow solid. Yield: 30%. **^{1}H NMR** (300 MHz, $CDCl_3$): δ (ppm) 5.69 (d, 1H, J = 9.0 Hz, NH); 5.29 (t, 1H, J = 9.6 Hz, H_3); 5.23 (d, 1H, J = 9.0 Hz, H_1); 5.07 (t, 1H, J = 9.6 Hz, H_4); 4.23 (dd, 1H, $J_{5\text{-}6a}$ = 4.8 Hz, $J_{6a\text{-}6b}$ = 12.3 Hz, H_{6a}); 4.13 (dd, 1H, $J_{5\text{-}6b}$ = 2.1 Hz, $J_{6a\text{-}6b}$ = 12.3 Hz, H_{6b}); 4.04-3.95 (m, 1H, H_2); 3.77–3.71 (m, 1H, H_5); 2.09, 2.04, 2.02, 1.99 (4s, 12H, $COCH_3$). **ESI-MS(+)**: m/z 410.9 [100% $(M+Na)^+$]; m/z 799.2 [10% $(2M+Na)^+$]. The product presents the same spectroscopic characteristics reported in literature[24].

***Cone*-5,11,17,23-tetrakis[(2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-β-D-glucopyranosyl) thioureido]-25,26,27,28-tetrapropoxycalix[4]arene (27)**

Calix[4]arene derivative **20**[21] (230 mg, 0.36 mmol) was suspended in 20 mL dry CH_2Cl_2, then glucosyl isothiocyanate **26** (700 mg, 1.8 mmol) and triethylamine (0.2 mL, 1.4 mmol) were added. The mixture was allowed to react for 2 days at room temperature under nitrogen atmosphere. The progress of the reaction was monitored via TLC (eluent: hexane/AcOEt/MeOH 5:5:1) and ESI-MS analyses. The solvent was then removed under reduced pressure and the residue purified via flash chromatography (eluent: hexane/AcOEt/MeOH 5:5:1) and trituration in diethyl ether. Product **27** was obtained as a

light yellow solid. Yield: 39%. **^{1}H-NMR** (300 MHz, DMSO-*d6*): δ (ppm) 9.63 (bs, 4H, ArNH); 8.12 (d, 4H, J = 9.0 Hz, NHAc); 7.49 (bs, 4H, GlcNH); 6.80 (bs, 8H, Ar); 5.51 (bs, 4H, H_1); 5.08 (t, 4H, J = 9.6 Hz, H_3); 4.80 (t, 4H, J = 9.6 Hz, H_4); 4.32 (d, 4H, J = 12.6 Hz, *ax*-ArCH_2Ar); 4.18 (m, 4H, H_{6a}); 3.95–3.70 (m, 20H, H_2, H_5, H_{6b}, OCH_2CH_2CH_3); 3.22 (d, 4H, J = 12.6 Hz, *eq*-ArCH_2Ar); 1.97, 1.93, 1.89 (3s, 44H, COCH_3, OCH_2CH_2CH_3); 1.77 (s, 12H, NHCOCH_3); 0.96 (t, 12H, J = 7.2 Hz, OCH_2CH_2CH_3). **ESI-MS(+)**: m/z 1125.6 [100% $(M+2Na)^{2+}$]. The product shows the same physical-chemical properties reported in the literature[47].

***Cone*-5,11,17,23-tetrakis[(2-acetamido-2-deoxy-β-D-glucopyranosyl)thioureido]-25,26,27,28-tetrapropoxycalix[4]arene (28)**

Cone-peracetylated-glucosylcalix[4]arene **27** (56.0 mg, 25.4 μmol) was dissolved in 5 mL of CH_3OH, and drops of a freshly prepared methanol solution of MeONa were added till pH 8-9. The mixture was stirred at room temperature for 4 hours. The proceeding of the reaction was monitored via TLC (AcOEt/i-propanol/H_2O 5:2:1) and ESI-MS analysis. Amberlite resin IR 120/H^+ was subsequently added for quenching, and the mixture was gently stirred for 30 min till neutral pH. The resin was then filtered off and the solvent removed under vacuum to give product **28** as light yellow solid in quantitative yield. **^{1}H NMR** (400 MHz, CD_3OD): δ (ppm) 6.81 (bs, 4H, Ar); 6.63 (bs, 4H, Ar); 5.58 (d, 4H, J = 8.0 Hz, H_1); 4.47 (d, 4H, J = 12.8 Hz, *ax*-ArCH_2Ar); 3.92–3.76 (m, 16H, H_2, H_{6a}, OCH_2CH_2CH_3); 3.71–3.69 (m, 4H, H_{6b}); 3.54–3.50 (m, 4H, H_3); 3.39-3.35 (m, 8H, H_4, H_5); 3.24 (d, 4H, J = 12.8 Hz, *eq*-ArCH_2Ar); 2.03–1.94 (m, 20H, OCH_2CH_2CH_3, NHCOCH_3); 1.04 (t, 12H, J = 7.2 Hz, OCH_2CH_2CH_3). **^{13}C NMR** (100 MHz, CD_3OD): δ (ppm) 181.2 (CS); 173.1 (CO); 154.6 (Ar-ipso); 135.3 (Ar-ortho); 131.6 (Ar-para); 124.4 (Ar-meta); 84.0 (C_1); 78.3, 70.5 (C_4, C_5); 76.8 (OCH_2CH_2CH_3); 74.7 (C_3); 61.3 (C_6); 69.7 (OCH_2); 55.0 (C_2); 30.5 (ArCH_2Ar); 23.0 (OCH_2CH_2CH_3); 21.9 (COCH_3); 9.4 (OCH_2CH_2CH_3). **ESI-MS(+)**: m/z 873.2 [100% $(M+2Na)^{2+}$]; m/z 1724.3 [50% $(M+Na)^+$]. The product shows the same physical-chemical properties reported in the literature[47].

***Cone*-5,11,17,23-tetrakis[isothiocyanate]-25,26,27,28-tetrapropoxycalix[4]arene (33)**

Amino-calix[4]arene **20**[21] (100 mg, 0.15 mmol) was suspended in 20 mL of dry toluene. Then Et_3N (425 μL, 3,06 mmol) and thiophosgene (93 μL, 1.23 mmol) were added and the mixture was stirred at 60 °C for 18 hours under nitrogen atmosphere. The progress of the reaction was monitored via TLC (eluent: AcOEt/Et_3N 9:1) and ESI-MS. The solvent was removed under vacuum and the residue dissolved in CH_2Cl_2 (30 mL) and washed with water (2 x 20

mL), brine (20 mL), dried over anhydrous $MgSO_4$ and filtered. The solvent was removed under reduced pressure to give product **33** as a brown solid. Yield: 65%. **^{1}H NMR** (300 MHz, $CDCl_3$): δ (ppm) 6.52 (s, 8H, Ar); 4.36 (d, 4H, J = 13.5 Hz, *ax*-ArC$\underline{H}_2$Ar); 3.81 (t, 8H, J = 7.5 Hz, OC$\underline{H}_2$CH$_2$CH$_3$); 3.09 (d, 4H, J = 13.5 Hz, *eq*-ArC$\underline{H}_2$Ar); 1.88 (sext, 8H, J = 7.5 Hz, OCH$_2$C$\underline{H}_2$CH$_3$); 0.98 (t, 12H, J = 7.5 Hz, OCH$_2$CH$_2$C$\underline{H}_3$). **ESI-MS(+)**: m/z 843.3 [100% (M+Na)$^+$]. **M.p.** > 320 °C. The product shows the same physical-chemical properties reported in the literature[30].

1-Azido-β-lactose (35)

Heptaacetylated-β-lactosyl azide **34**[26] (3.00 g, 4.84 mmol) was dissolved in 40 mL of CH_3OH, and drops of a freshly prepared methanol solution of MeONa were added till pH 8-9. The mixture was stirred at room temperature for 18 hours. The progress of the reaction was checked via ESI-MS analysis. Amberlite resin IR 120/H$^+$ was subsequently added for quenching, and the mixture was gently stirred for 30 min till neutral pH. The resin was filtered off and the solvent removed under vacuum to obtain product **35** as a white solid. Yield: 89%. **^{1}H NMR** (300 MHz, CD_3OD): δ (ppm) 4.55 (d, 1H, J = 8.7 Hz, H_1); 4.36 (d, 1H, J = 7.5 Hz, H_1'); 3.92 (dd, 1H, $J_{5\text{-}6a}$ = 2.4 Hz, $J_{6a\text{-}6b}$ = 12.0 Hz, H_{6a}); 3.85 (dd, 1H, $J_{5\text{-}6b}$ = 3.9 Hz, $J_{6a\text{-}6b}$ = 12.0 Hz, H_{6b}); 3.81 (d, 1H, J = 3.0 Hz, H_4'); 3.78 (dd, 1H, $J_{5'\text{-}6'a}$ = 7.5 Hz, $J_{6'a\text{-}6'b}$ = 11.2 Hz, $H_6'_a$); 3.69 (dd, 1H, $J_{5'\text{-}6'b}$ = 4.5 Hz, $J_{6'a\text{-}6'b}$ = 11.2 Hz, $H_6'_b$); 3.64-3.46 (m, 6H, H_3, H_4, H_5, $H_{2'}$, $H_{3'}$, $H_{5'}$); 3.21 (t, 1H, J = 8.7 Hz, H_2). **^{13}C NMR** (75 MHz, CD_3OD): δ (ppm) 103.6 (C_1'); 90.5 (C_1); 78.6, 77.2, 75.7, 75.1, 71.1 (C_3, C_4, C_5, C_2', C_5'); 73.4 (C_3'), 73.0 (C_2); 68.9 (C_4'); 61.1 (C_6'); 60.3 (C_6). **ESI-MS(+)** m/z: 390.1 [100%, (M+Na)$^+$]. The spectroscopic data found are in agreement with those reported in literature[48].

1-Amino-β-lactose (36)

1-Azido-β-lactose **35** (0.10 g, 0.17 mmol) was dissolved in 40 mL ethanol absolute with the help of ultrasounds. Under nitrogen flow, a catalytic amount of Pd/C was added. The mixture was stirred under H_2 atmosphere (pressure: 2.0 bar) for 18 h at room temperature, using the Parr apparatus. The reaction was monitored via ESI-MS analysis. The catalyst was then filtered off, the filters washed with ethanol, and the solvent removed under reduced pressure to give product **36** as white solid. Yield: 24%. **^{1}H NMR** (400 MHz, D_2O): δ (ppm) 4.36 (d, 1H, J = 8.0 Hz, H_1'); 4.03 (d, 1H, J = 8.8 Hz, H_1); 3.88-3.83 (m, 2H, H_{6a}, H_4'); 3.75-3.61 (m, 4H, H_{6b}, H_5', $H_6'_a$, $H_6'_b$); 3.60-3.51 (m, 3H, H_3, H_3'); 3.50-3.43 (m, 2H, J = 8.7 Hz, H_2', H_5); 3.12

(m, 1H, H_2). **^{13}C NMR** (100 MHz, D_2O): δ (ppm) 102.9 (C_1'); 84.9 (C_1); 78.6 (C_4); 75.6 (C_5); 75.3 (C_5'); 75.1 (C_3); 73.9(C_2); 72.5 (C_3'); 70.9 (C_2'); 68.5 (C_4'); 61.0 (C_6'); 60.2 (C_6). **ESI-MS(+)** m/z: 364.1 [100%, $(M+Na)^+$]. The product shows the same physical-chemical properties reported in the literature[49].

2-Bromoethyl-2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl-(1→4)-2,3,6-di-*O*-acetyl-β-D-glucopyranosyl isothiocyanate (37)

Peracetylated lactose (2.50 g, 3.68 mmol) was dissolved in 20 mL dry CH_2Cl_2. The mixture was cooled to 0 °C with an ice-water bath, and then boron trifluoride etherate (1.87 mL, 14.74 mmol) and 2-bromoethanol (0.36 mL, 5.16 mmol) were added. The mixture was stirred at room temperature under N_2 atmosphere for 18 hours. The progress of the reaction was monitored via TLC (eluent: AcOEt/hexane 1:1). The reaction was quenched by addition of H_2O (20 mL) and brine (15 mL). The organic layer was recovered and the product extracted with CH_2Cl_2 (3 x 20 mL). The combined organic phases were washed with H_2O till neutral pH, dried over anhydrous Na_2SO_4, filtered and the solvent removed under reduced pressure to give crude product **37** as a white solid. The product was not further purified, and used directly for the next reaction. Proof that the product was obtained was given by ESI-MS mass spectrum of the crude. Yield: 89%. **ESI-MS(+)**: m/z 764.9, 767.0 [100% $(M+Na)^+$]. The physical-chemical properties of this compound are reported in the literature[50].

2-Azidoethyl-2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl-(1→4)-2,3,6-di-*O*-acetyl-β-D-glucopyranosyl isothiocyanate (38)

NaN_3 (0.63 g, 9.75 mmol) was added to a solution of lactoside derivative **37** (2.42 g, 3.25 mmol) in dry DMF (20 mL). The mixture was stirred at 65 °C for 18 hours under nitrogen atmosphere. The reaction was monitored via TLC (eluent: AcOEt/hexane 1:1) and ESI-MS analyses. When the reaction was finished the excess of NaN_3 was filtered off and the filters washed with AcOEt. The solution was washed with water (20 mL) and the product extracted with AcOEt (4 x 30 mL). The combined organic phases were dried over anhydrous Na_2SO_4, filtered and the solvent removed *in vacuo*. The residue was purified via column chromatography (eluent: AcOEt/hexane 1:1) to give product **38** as a white solid. Yield: 54%.

^{1}H NMR (300 MHz, $CDCl_3$): 5.30 (d, 1H, $J_{3'-4'}$ = 3.3 Hz, H_4'); 5.17 (t, 1H, J = 9.0 Hz, H_3); 5.07 (dd, 1H, $J_{1'-2'}$ = 7.8 Hz, $J_{2'-3'}$ = 10.5 Hz, H_2'); 4.91 (dd, 1H, $J_{3'-4'}$ = 3.3 Hz, $J_{2'-3'}$ = 10.5 Hz, H_3'); 4.88 (t, 1H, J = 9.0 Hz, H_2); 4.52 (d, 1H, J = 7.8 Hz, H_1); 4.48-4.45 (m, 2H, H_1', H_{6a}); 4.13-4.01 (m, 3H,

H_{6a}, $H_{6'a}$, $H_{6'b}$); 3.99-3.93 (m, 1H, β-OCH_a); 3.85 (t, 1H, J = 7.4 Hz, $H_{5'}$); 3.79 (t, 1H, J = 9.0 Hz, H_4); 3.68-3.57 (m, 2H, H_5, β-OCH_b); 3.48-3.39 (m, 1H, CH_aN_3); 3.27-3.19 (m, 1H, CH_bN_3); 2.12, 2.09, 2.03, 2.01, 1.93, 1.84 (6s, 18H, Ac). **ESI-MS(+)**: m/z 728.1 [100% $(M+Na)^+$]. The product shows the same physical-chemical properties reported in the literature[51].

2-Azidoethyl-(β-D-galactopyranosyl)-(1→4)-β-D-glucopyranoside (39)

Compound **38** (450 mg, 0.64 mmol) was dissolved in 35 mL of CH_3OH, and drops of a freshly prepared methanol solution of MeONa were added till pH 8-9. The mixture was stirred at room temperature for 4 hours. The progress of the reaction was checked via ESI-MS analyses. Amberlite resin IR 120/H^+ was subsequently added for quenching, and the mixture was gently stirred for 30 min till neutral pH. The resin was filtered off and the solvent removed under vacuum to obtain product **39** as a white solid. Yield: 94%. **^{1}H NMR** (300 MHz, D_2O): δ (ppm) 4.40 (d, 1H, J = 7.8 Hz, H_1); 4.36 (d, 1H, J = 7.5 Hz, $H_{1'}$); 3.96-3.89 (m, 1H, β-OCH_a); 3.86 (dd, 1H, $J_{5\text{-}6a}$ = 1.5 Hz, $J_{6a\text{-}6b}$ = 12.6 Hz, H_{6a}); 3.80 (d, 1H, J = 3.0 Hz, $H_{4'}$); 3.76-3.62 (m, 4H, β-OCH_b, H_{6b}, $H_{6'a}$, $H_{6'b}$); 3.61-3.49 (m, 5H, H_3, H_4, H_5, $H_{3'}$, $H_{5'}$); 3.48-3.41 (m, 3H, $H_{2'}$, CH_2N_3); 3.22 (t, 1H, J = 8.4 Hz, H_2). **^{13}C NMR** (75 MHz, D_2O): δ (ppm) 102.9 ($C_{1'}$); 102.1 (C_1); 78.3, 75.3, 74.8, 74.3, 72.5 (C_3, C_4, C_5, $C_{3'}$, $C_{5'}$); 72.7 (C_2), 70.9 ($C_{2'}$); 68.5 ($C_{4'}$, β-OCH_2); 61.0 ($C_{6'}$); 60.0 (C_6), 50.5 (CH_2N_3). **ESI-MS(+)** m/z: 434.0 [100%, $(M+Na)^+$]. The product shows the same physical-chemical properties reported in the literature[52]**.**

2-Aminoethyl-(β-D-galactopyranosyl)-(1→4)-β-D-glucopyranoside (40)

Azido derivative **39** (240 mg, 0.58 mmol) was dissolved in 3 mL H_2O and diluted with 20 mL EtOH. Under a nitrogen flow, a catalytic amount of Pd/C (10%) was added. The mixture was stirred under H_2 atmosphere (pressure: 1.5 bar) for 18 hours at room temperature using the Parr apparatus. The reaction was monitored via ESI-MS mass spectra. The catalyst was then filtered off and the solvent removed under reduced pressure giving product **40** as a white solid. Yield: 70%. **^{1}H NMR** (400 MHz, D_2O): δ (ppm) 4.39 (d, 1H, J = 8.0 Hz, H_1); 4.30 (d, 1H, J = 8.0 Hz, $H_{1'}$); 4.01-3.94 (m, 1H, β-OCH_a); 3.83 (dd, 1H, $J_{6a\text{-}6b}$ = 12.8 Hz, $J_{5\text{-}6a}$ = 1.5 Hz, H_{6a}); 3.83-3-80 (m, 1H, β-OCH_b); 3.78 (d, 1H, J = 3.2 Hz, $H_{4'}$); 3.70-3.58 (m, 3H, H_{6b}, $H_{6'a}$, $H_{6'b}$); 3.57-3.44 (m, 5H, H_3, H_4, H_5, $H_{3'}$, $H_{5'}$); 3.39 (dd, 1H, $J_{2'\text{-}3'}$ = 9.6 Hz, $J_{1'\text{-}2'}$ = 8.0 Hz, $H_{2'}$); 3.22 (t, 1H, J = 8.0 Hz, H_2); 3.12 (t, 2H, J = 4.8 Hz, $C\underline{H}_2NH_2$). **^{13}C NMR** (75 MHz, D_2O): δ (ppm) 102.9 ($C_{1'}$); 101.9 (C_1); 78.2, 75.3, 74.7, 74.2, 72.5 (C_3, C_4, C_5, $C_{3'}$, $C_{5'}$); 72.7 (C_2); 70.9 ($C_{2'}$); 68.5 ($C_{4'}$); 65.8 (β-

OCH_2); 61.0 (C_6'); 59.9 (C_6); 57.4 (CH_2NH_2). **ESI-MS(+)** m/z: 386.0 [100%, $(M+H)^+$]; 408.0 [100%, $(M+Na)^+$]. The product shows the same physical-chemical properties reported in the literature[51].

5,11,17,23-Tetrakis[(2-aminoethyl-(β-D-galactopyranosyl)-(1→4)-β-D-glucopyranosyl)thio ureido]-25,26,27,28-tetrapropoxycalix[4]arene (41)

Sugar compound **40** (35 mg, 90 µmol) and Et_3N (75 µL, 0.54 mmol) were added to a solution of calix[4]arene derivative **33** (15 mg, 18 µmol) in 2.5 mL DMF. The mixture was stirred at room temperature for 18 hours and monitored via TLC (reverse phase, silica gel 60 RP-18, eluent: H_2O/i-propanol 1:1) and ESI-MS analysis. The solvent was then removed under reduced pressure and the crude product purified via size exlusion chromatography (Sephadex G-25, eluent: H_2O) to give pure product **41** as a light brown solid. Yield: 47%. **^{1}H NMR** (300 MHz, DMSO-*d_6*): δ (ppm) 9.23 (bs, 4H, NH); 7.34 (bs, 4H, NH); 6.69 (bs, 8H, Ar); 5.10-5.09 (m, 8H, OH); 4.79-4.50 (m, 20H, OH); 4.32-4.16 (m, 12H, H_1, H_1', *ax*-ArC$\underline{H}_2$Ar); 3.85-3.65 (m, 16H, H_{6a}, β-OCH_a, OC$\underline{H}_2CH_2CH_3$); 3.64-3-16 (m, 52H, H_2-H_5, H_{6b}, H_3'-H_6', β-OCH_b, CH_2N); 3.15-3.01 (m, 8H, *eq*-ArC$\underline{H}_2$Ar, H_2); 1.90-1.83 (m, 8H, $OCH_2C\underline{H}_2CH_3$); 0.97 (t, 12H, J = 7.0 Hz, $OCH_2CH_2C\underline{H}_3$). **^{13}C NMR** (100 MHz, DMSO): δ(ppm) 180.5 (CS); 153.6 (Ar-ipso); 134.8 (Ar-ortho); 133.0 (Ar-para); 123.9 (Ar-meta); 104.3, 103.1 (C_1', C_1); 81.2 (C_4); 76.9 ($O\underline{C}H_2CH_2CH_3$); 76.0, 75.3, 73.7, 73.6, 71.0 (C_2', C_3', C_5', C_3, C_4, C_5); 68.6 (C_4'), 67.9 (β-OCH_2); 60.8 (C_6, C_6'); 44.2 (CH_2N); 30.8 (Ar$\underline{C}H_2$Ar); 23.2 ($OCH_2\underline{C}H_2CH_3$); 10.6 ($OCH_2CH_2\underline{C}H_3$). **HR-ESI-MS(+)**: m/z calculated for $C_{100}H_{152}N_8O_{48}S_4Na$ $(M+Na)^+$ 2384.85076, found 2384.85620. **M.p.** : 195-200 °C dec.

General procedure for conjugation between aminocalix[4/6]arenes 20 and 46 and glycosylisothiocyanates 44 and 45

In a two-neck round bottom flask 1 equivalent (0.1 mmol) of the amino-calix[4]arene **20**[21] or amino-calix[6]arene **46**[19] was dissolved in 3 mL of dry CH_2Cl_2 under N_2 atmosphere. Then for each amino group of the calixarene 1.25 equivalents of glycosyl isothiocyanate **44** or **45** and 1 equivalent of triethylamine were added. The mixture was allowed to react at room temperature under N_2 overnight and then the solvent evaporated under reduced pressure.

Each compound was obtained as follows: **47**, via preparative layer on silica gel (hexane/AcOEt/CH_3OH 5:5:1); **48**, via preparative layer on silica gel (CH_2Cl_2/CH_3OH 95:5 and then CH_2Cl_2/AcOEt/CH_3OH 5:5:0.7); **49**, via flash column cromatography (CH_2Cl_2/CH_3OH

97.5:2.5) followed by preparative layer on silica gel (hexane/AcOEt/CH_3OH 5:5:1); **50**, via flash column chromatography (hexane/AcOEt/CH_3OH from 5:5:1 to 6:3:1) followed by preparative layer on silica gel (hexane/AcOEt/CH_3OH 6:4:1). Purified acetylated products **47-50** were then dissolved in CH_3OH and drops of a freshly prepared methanol solution of CH_3ONa were added till pH 8-9. The mixture was stirred at room temperature overnight. When a precipitate was observed, H_2O was added to help complete solubilisation. Amberlite resin IR 120/H^+ was subsequently added for quenching. After neutralization, the resin was filtered off and the solvent removed from the filtrate under vacuum to give pure products (**51-54**).

***Cone*-5,11,17,23-tetrakis[2-acetamido-1,2-dideoxy-4-*O*-[3-*O*-(3-methoxybenzyl)-β-D-galacto pyranosyl)]-β-D-glucopyranosyl]thioureido-25,26,27,28-tetrapropoxycalix[4]arene (51)**

Compound **51** was obtained as a white solid in 44% yield. **^{1}H NMR** (300 MHz, CD_3OD, 333K): δ (ppm) 7.21 (t, 4H, J = 7.8 Hz); 7.06 (s, 4H); 6.99 (d, 4H, J = 6.9 Hz); 6.88-6.82 (m, 8H); 6.63 (sb, 4H); 5.61 (d, 4H, J = 9.3Hz, H_1); 4.77-4.64 (m, 8H); 4.50 (m, 8H); 4.05-3.50 (m, 64H); 3.41 (d, 4H, J = 6.0 Hz); 3.24 (d, 4H, J = 13.2 Hz); 1.98 (m, 20H); 1.05 (t, 12H, J = 6.9 Hz). **^{13}C NMR** (400MHz, CD_3OD): δ (ppm) 180.9, 173.1, 159.8, 154.4, 140.0, 135.5, 131.0, 128.9, 124.1, 119.7, 112.9, 103.8, 83.9, 80.8, 79.5, 76.7, 75.5, 72.9, 70.9, 70.5, 65.6, 61.1, 60.4, 54.3, 30.6, 23.0, 22.0, 9.4**. HR-ESI-MS(+)**: m/z calculated for $C_{132}H_{180}N_{12}O_{48}S_4Na_2$ $(M+2Na)^{2+}$ 1438.03569, found 1438.03564.

5,11,17,23,29,35-Hexakis[2-acetamido-1,2-dideoxy-4-*O*-[3-*O*-(3-methoxybenzyl)-β-D-galactopyranosyl)]-β-D-glucopyranosyl]thioureido-25,26,27,28,29,30-hexamethoxy-calix[6]arene (52)

Compound **52** was obtained as a white solid in 22% yield. **^{1}H NMR** (300 MHz, CD_3OD, 333K): δ(ppm) 7.20 (t, 6H, J = 7.5 Hz); 7.10-6.85 (m, 24H); 6.81 (d, 6H, J = 9 Hz); 5.54 (d, 6H, J = 9 Hz, H_1); 4.68 (q, 12H,J = 12 Hz); 4.42 (bs, 6H); 4.08-3.45 (m, 114H); 3.39 (dd, 6H,J = 6Hz, J = 3 Hz); 2.01 (bs, 18H). **^{13}C NMR** (400 MHz, CD_3OD): δ (ppm) 182.0, 173.2, 172.9, 159.8, 154.0, 140.0, 134.6, 128.9, 125.8, 119.7, 112.9, 103.9, 83.3, 80.8, 79.8, 76.6, 75.5, 72.8, 71.0, 70.4, 66.7, 65.6, 61.1, 60.6, 60.1, 54.3, 29.2, 21.7. **HR-ESI-MS(-)**: m/z calculated for $C_{186}H_{245}N_{18}O_{72}S_6Cl$ $(M-H+Cl)^{2-}$ 2055.70770, found 2055.71074.

***Cone*-5,11,17,23-tetrakis[1,2-dideoxy-4-*O*-(β-D-galactopyranosyl)-2-(3-methoxy) benzamido-β-D-glucopyranosyl]thioureido-25,26,27,28-tetrapropoxycalix[4]arene (53)**

After filtering off the Amberlite resin IR 120/H^+, the crude was purified via trituration in ether to give compound **53** as a white solid in 55% yield. **^{1}H NMR** (300MHz, D_2O, 348K): δ(ppm) 7.21-6.86 (m, 12H); 6.61 (bs, 4H); 6.21 (bs, 4H); 6.00 (bs, 4H); 5.68 (bs, 4H, H_1); 4.37 (bs, 4H, H_1'); 4.01–3.21 (m, 72H); 2.57 (bs, 4H); 1.58 (bs, 8H); 0.69 (bs, 12H). **^{13}C NMR** (400MHz, CD_3OD/D_2O 9:1): δ(ppm) 181.8, 169.7, 159.6, 155.2, 135.3, 129.7, 125.6, 125.1, 119.5, 117.5, 113.0, 103.6 , 83.9, 79.4, 76.7, 75.6, 73.2, 72.8, 71.2, 68.8, 62.8, 61.1, 60.5, 55.0, 54.3, 30.1, 23.0, 9.5. **HR-ESI-MS(+)**: m/z calculated for $C_{124}H_{164}N_{12}O_{48}S_4Na_2$ $(M+2Na)^{2+}$ 1381.97309, found 1381.97449.

5,11,17,23,29,35-Hexakis[1,2-dideoxy-4-*O*-(β-D-galactopyranosyl)-2-(3-methoxy) benzamido-β-D-glucopyranosyl]thioureido-25,26,27,28,29,30-hexamethoxycalix[6]arene (54)

Compound **54** was obtained as a white solid in 42% yield. **^{1}H NMR** (400MHz, CD_3OD/D_2O 3:1, 333K): δ (ppm) 7.40-7.20 (m, 18H, Ar); 7.04 (bs, 6H, Ar); 6.67 (bs, 12H); 5.77(d, 6H, J = 7.2 Hz, H_1); 4.49 (bs, 6H, H_1'); 4.16 (t, 6H, J = 9.6 Hz); 4.05–3.45 (m, 102H); 2.67 (bs, 12H). **^{13}CNMR** (400MHz, CD_3OD/D_2O 3:1): δ (ppm) 182.0, 170.1, 159.4, 155.1, 134.9, 129.8, 126.3, 119.6, 118.0, 112.8, 103.4, 84.0, 79.2, 76.7, 75.6, 73.0, 71.2, 68.8, 61.1, 60.2, 55.2, 30.1. **HR-ESI-MS(-)**: m/z calculated for $C_{174}H_{221}N_{18}O_{72}S_6Cl_2$ $(M-H+2Cl)^{3-}$ 1326.39781, found 1326.40006.

2.5 References

1 Andrè, S.; Sansone, F.; Kaltner, H.; Casnati, A.; Kopitz, J.; Gabius, H. J.; Ungaro, R. *ChemBioChem*, **2008**, *9*, 1649-1661.

2 Krzeminski, M.; Singh, T.; André, S.; Lensch, M.; Wu, A. M.; Bonvin A. M. J. J.; Gabius. H. J. *Biochim. Biophys. Acta*, **2011**, *1810*, 150-161.

3 Ahmed, H.; Vasta, G. R. *Glycobiology*, **1994**, *4*, 545-548.

4 Fort, S.; Kim H. S.; Hindsgaul, O. *J. Org. Chem.*, **2006**, *71*, 7146-7154.

5 Cumpstey, I.; Salomonsson, E.; Sundin, A.; Leffler, H.; Nilsson, U. J. *ChemBioChem*, **2007**, *8*, 1389-1398.

6 Garcia Fernández J. M.; Ortiz Mellet C. *Adv. Carbohydr. Chem. Biochem.*, **2000**, *55*, 35-135.

7 Sansone, F.; Baldini, L.; Casnati, A.; Ungaro, R. *New J. Chem.,***2010**, *34*, 2715-2728.

8 Fukuda, M.; *Biochim. Biophys. Acta,* **1985***, 119,* 780.

9 Feizi, T.; Childs, R. A. *J. Biochem.,* **1987**, *1*, 245.

10 Fiandor, J.; Garcia-Lopez, M. T.; De Las Heras, F. G.; Mendez-Castrillon, P. P. *Synthesis*, **1985**, 1121-1123.

11 Schmidt, R. R.; Michel, J. *Tetrahedron Letters.*, **1984**, *25 (8)*, 821-824.

12 Gan, Z.; Cao, S.; Wu, Q.; Roy, R. *J. Carbohydr. Chem.*, **1999**, *18* (7), 755-773.

13 D'Onofrio, J.; de Champdoré, M.; De Napoli, L.; Montesarchio, D.; Di Fabio, G. *Bioconjugate Chem.*, **2005**, *16* (5), 1299-1309.

14 Nakahara, Y.; Ogawa, T. *Carbohydrate Research*, **1989**, *194*, 95-114.

15 Isbell, H. S.; Frush, H. L. *J. Org. Chem.*, **1958**, *23*, 1309-1319.

16 Linek, K.; Alföldi, J. *Carb. Res.*, **1987**, *164*, 195-205.

17 Groenen, L. C.; Ruël, B. H. M.; Casnati, A.; Timmerman, P.; Verboom, W.; Harkema, S.; Pochini, A.; Ungaro, R.; Reinhoudt, D. N. *Tetrahedron Lett.*, **1991**, *32*, 2675-2678.

18 Kelderman, E.; Der Haeg, L.; Heesink, G. J. T.; Verboom, W.; Engbersen, J. F. J. *Angew. Chem.,* **1992**, *104*, 1107-1110. *Angew Chem. Int. Ed. Engl.*, **1992**, *39*, 1075-1077.

19 Dudič, M.; Colombo, A.; Sansone, F.; Casnati, A.; Donofrio, G.; Ungano, R. *Tetrahedron*, **2004**, *60*, 11613-11618.

20 Verboom, W.; Durie, A.; Egberink, R. J. M.; Asfari, Z.; Reinhoudt, D. N. *J. Org. Chem.,* **1992**, *57*, 1313-1316.

21 Sansone, F.; Chierici, E.; Casnati, A.; Ungaro, R. *Org. Biomol. Chem.,* **2003**, *1*, 1802-1809.

22 Zemplén, G.; Pascu, E. *Ber. Dtsch. Chem. Ges.*, **1929**, *62*, 1613-1618.

23 Robinson, A.; Fang, J. M.; Chou, P. T.; Liao, K. M.; Chu, R. M.; Lee, S. J. *ChemBioChem*, **2005**, *6*, 1899-1905.

24 Camarasa, J.; Fernandez-Resa, P.; Garcia-Lopez M. T.; De Las Heras, F. G.; Mendez-Castrillon, P. P.; San Felix, A. *Synthesis*, **1984**, 509-510.

25 Lindhorst, T. K.; Kieburg, C. *Synthesis*, **1995**, 1228-1230.

26 Tropper, F. D.; Andersson, F. O.; Braun, S.; Roy, R. *Synthesis*, **1992**, 618–620.

27 Ideo, H.; Matsuzaka, T.; Nonaka, T.; Seko, A.; Yamashita, K. *Journal of Biological Chemistry*, **2011**, *286* (13), 11346-11355.

28 Zhuo, Ya.; Bellis, S. L. *Journal of Biological Chemistry*, **2011**, *286* (8), 5935-5941.

29 Stowell, S. R.; Arthur, C. M.; Mehta, P.; Slanina, K. A.; Blixt, O.; Leffler, H.; Smith, D. F.; Cummings, R. D. *J. Biol. Chem.*, **2008**, *283*(15), 10109-10123.

30 van Wageningen, A. M.; Snip, E.; Verboom, W.; Reinhoudt, D. N.; Boerrigter, H. *Liebigs Ann./Recueil*, **1997**, 2235-2245.

31 López, O.; Maya, I.; Fuentes, J.; Fernández-Bolaños, J. G. *Tetrahedron*, **2004**, *60*, 61-72.

32 Koshi, Y.; Nakata, E.; Miyagawa, M.; Tsukiji, S.; Ogawa, T.; Hamachi, I. *J. Am. Chem. Soc.*, **2008**, *130*, 245-251.

33 (a) Sparrow, C. P.; Leffler, H.; Barondes, S. H. *J. Biol. Chem.*, **1987**, *262*, 7383-7390; (b) Oda, Y.; Herrmann, J.; Gitt, M. A.; Turck, C. W.; Burlingame, A. L.; Barondes, S. H.; Leffler, H. *J. Biol. Chem.*, **1993**, *268*, 5929-5939; (c) Hirabayashi, J.; Hashidate, T.; Arata, Y.; Nishi, N.; Nakamura, T.; Hirashima, M.; Urashima, T.; Oka, T.; Futai, M.; Müller, W. E. G.; Yagi F.; Kasai, K. I. *Biochim. Biophys. Acta* **2002**, *1572*, 232-254; (d) Wu, A. M.; Wu, J. H.; Tsai, M. S.; Liu, J. H.; André, S.; Wasano, K.; Kaltner, H.; Gabius, H. J. *Biochem. J.*, **2002**, *367*, 653-664; (e) Wu, A. M.; Wu, J. H.; Liu, J. H.; Singh, T.; André, S.; Kaltner, H.; Gabius, H. J. *Biochimie*, **2004**, *86*, 317-326.

34 (a) Sörme, P.; Arnoux, P.; Kahl-Knutsson, B.; Leffler, H.; Rini, J. M.; Nilsson. U. J. *J. Am. Chem. Soc.*, **2005**, *127*, 1737-1743. (b) Aplander, K.; Tejler, J.; Toftered, J.; Carlsson, S.; Kahl-Knutsson, B.; Sundin, A.; Lefflerb, H.; Nilsson, U. J. *Carbohydrate Research,* **2006**, *341*, 1363-1369.

35 Houzelstein, D.; Goncalves, I. R.; Fadden, A. J.; Sidhu, S. S.; Cooper, D. N.; Drickamer, K.; Leffler, H.; Poirier, F. *Mol. Biol. Evol.* **2004**, *21*, 1177-1187.

36 André, S.; Grandjean, C.; Gautier, F. M.; Bernardi, S.; Sansone, F.; Gabius, H. J.; Ungaro, R. *Chem. Commun.*, **2011**, *47*, 6126-6128.

37 André, S.; Giguère, D.; Dam T. K.; Brewer, C. F.; Gabius, H. J.; Roy, R.; *New J. Chem.*, **2010**, *34*, 2229-2240.

38 (a) Galanina, O.E.; Kaltner, H.; Khraltsova, L. S.; Bovin N. V.; Gabius, H. J. *J. Mol. Recognit*., **1997**, *10*, 139-147; (b) Jiménez, M.; André, S.; Barillari, C.; Romero, A.; Rognan, D.; Gabius, H. J.; Solís, D. *FEBS Lett.*, **2008**, *582*, 2309-2312.

39 Sancehz-Ruderisch, H.; Fischer, C.; Detjen, K. M.; Welzel, M.; Wimmel, A.; Manning, J. C.; André, S.; Gabius, H. J. *FEBS Journal* **2010**, *277*, 3552-3563.

40 Gottlieb, H. E.; Kotlyar, V.; Nudelman, A. *J. Org. Chem*. **1997**, *62*, 7512-7515.

41 Liu, Z.; Byun H. S.; Bittman, R. *Org. Lett.*, **2010**, *12* (13), 2974–2977.

42 Cheng et al. *J. Med. Chem.,***2005,** *48,* 645-652.

43 Micheel, F.; Helmut, W. *Chemische Berichte*, **1995**, *89* (6), 1521-1530.

44 Gudmundsdottir, A. V.; Nitz. M. *Org. Lett*., **2008**, *10* (16), 3461-3463.

45 Pratt, M. R.; Bertozzi, C. R. *J. Am. Chem. Soc.*, **2003**, *125*, 6149-6159.

46 Neumann, J.; Thiem, J. *Eur. J. Org. Chem.*, **2010**, 900-908.

47 Krenek, K. et al., *Carbohydrate Research*, **2007**, *342*, 1781-1792.

48 Zhang, W.; Wang, J.; Li, J.; Yu, L.; Wang P. G. *J. Carbohydrate Chemistry.*, **1999**, *18* (8), 1009-1017.

49 Ogata, M.; Hidari, K. I. P. J.; Kozaki, W.; Murata, T.; Hiratake, J.; Park, E. Y.; Suzuki, T.; Usui, T.; *Biomacromolecules*, **2009,** *10*, 1894-1903.

50 Muñoz, F. J.; Rumbero, A.; Sinisterra, J. V.; Santos, J. I.; André, S.; Gabius, H. J.; Jiménez-Barbero, J.; Hernáiz, M. J. *Glycoconj. J*., **2008**, *25*, 633-646.

51 Cheng, H. et al, *J. Med. Chem.* **2005,** *48,* 645-652.

52 Gao, Y.; Eguchi, A.; Kakehib, K.; Lee, Y. C. *Bioorg. Med. Chem.*, **2005**, *13*, 6151-6157.

Chapter 3

Studies of the lectin-glycocalixarene interactions by NMR techniques and computational methods

Abstract[§]

Recent studies have found that calix[n]arene based lacto-clusters result very efficient in the inhibition of VAA plant toxin and tumor related galactins, showing a remarkable selectivity towards given lectins[1,2]. To gain detailed insights into the way of interaction between these glycoclusters and the proteins, few representative glycocalixarenes were chosen to be studied in solution by NMR techniques. In particular, the binding of three different calix[4]arenes (presenting a *cone* or 1,3-alternate structure and functionalized either with lactose or LacNAc units) to VAA, truncated Gal-3 or Gal-9N was unraveled thanks to the use of STD and DOSY NMR experiments. Besides, a monovalent molecule, similar in structure to a monomeric unit of the calixarene ligands, was synthesized to be used as reference in the different NMR experiments. Also modeling methods were employed to provide a complete picture of the ligand-lectin interaction.

3.1 Introduction

It is well known that many biological processes start with the specific recognition of a carbohydrate determinant by a suitable receptor [3, 4]. The glycan parts of cell glycoconjugates bind to specific lectins and these processes mediate diverse aspects of cell regulation[5], such as cell proliferation, migration or adhesion. The elucidation of the protein-ligand interactions and the unraveling of these processes in details, constitute a fascinating and at the same time extremely useful field to explore. NMR spectroscopic techniques have become a powerful tool to understand and investigate the carbohydrate-lectin interactions in solution. A ligand interacting with a protein in fact perturbs the structure of both involved partners, thus causing the variation of several parameters, including NMR observables such as chemical shifts, relaxation parameters T_1 and T_2, diffusion coefficients, NOEs and so on[6]. With NMR techniques we can exploit the different physical properties of small and large molecules, especially those based on relaxation and diffusion, to elucidate the structure of the carbohydrate-protein complex. Therefore, we focused our attention in particular on two

[§] *The NMR and computational studies presented in this chapter were carried out at the CIB-CSIC of Madrid (Spain) under the supervision of Prof. Jesús Jiménez-Barbero, during a three months period.*

important NMR techniques: Saturation Transfer Difference (STD) and Diffusion Ordered SpectroscopY (DOSY) NMR.

STD-NMR is a method based on the transfer of saturation from a macromolecular receptor to a bound small ligand, which in turn moves into solution where it is detected[6,7,8] (figure **3.1**).

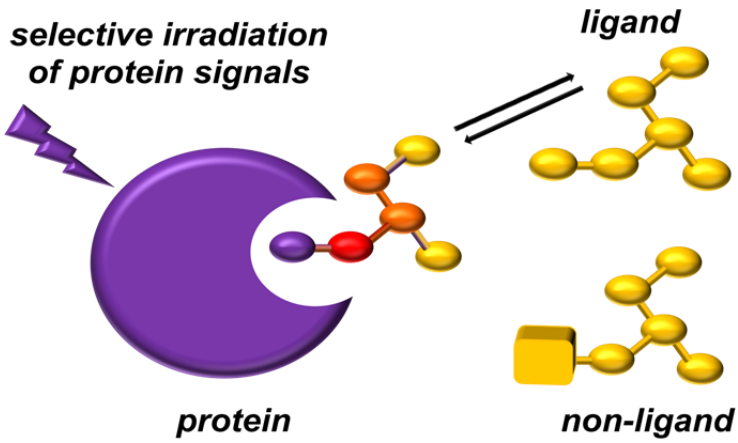


Figure 3.1 Schematic representation of the STD technique.

A STD experiment is performed by initially recording a ^{1}H NMR spectrum of the ligand-receptor solution with a selective irradiation of the protein (on-resonance spectrum). The frequency should be set at least 700 MHz far from the first ligand signal. For a 500 MHz instrument, usually -1.5 ppm is a good choice, even if the chemical nature of the ligand and of the protein has to be taken into consideration. The irradiation has usually a band width of only a few Hz, but still full saturation of the protein can usually be achieved within 50-200 ms by spin diffusion. A ligand binding to the protein will also be saturated, and the degree of its saturation is dependent on the resident time of the ligand in the protein binding pocket. Upon dissociation, the ligand transfers the saturation into solution, where it gives rise to narrow widths resonance signals. For the protons interacting through an intermolecular trNOE with the protein, a decrease of intensity is observed. However this attenuation is usually too low to be detected directly. Therefore, in a second experiment the irradiation frequency is set at a value far from any signal, both of the ligand and the protein, for example at 100 ppm, obtaining in this way a normal ^{1}H NMR spectrum (off-resonance spectrum). Subtraction of the on-resonance spectrum, in which the protein is saturated,

from the off-resonance one produces a difference spectrum where only the signals of the ligands in contact with the protein appear. All compounds without binding activity are cancelled out. In this way it is possible to identify the biologically active components directly from a mixture, when the dissociation constant of the ligands are in a range between 10^{-3} and 10^{-8} M. Moreover, in favourable cases, this technique also permits to determine the epitope map, that is to say to identify which parts of the ligand molecule are in closer contact with the macromolecule, since they receive the strongest degree of saturation. The ligand is normally used in an approximately 100-fold molar excess over the protein, allowing low µM protein concentration to be used.

To gather further information on the type of the ligand-receptor complex formed, DOSY-NMR experiments[9] can result very useful. This technique provides a method to determine the molecular size of a given compound or aggregate through the measurement of diffusion coefficients (D values). This parameter is a property of the molecule as a whole, therefore is shared by all the protons of the molecule. A change of the D values gives indication that a variation of the molecular size has occurred, and therefore binding events can be detected and the molecular weight of the resulting complex quantified. Thus, in principle, it should be possible to employ this method to unravel the stoichiometry of the complexes and prove the presence of multivalent events.

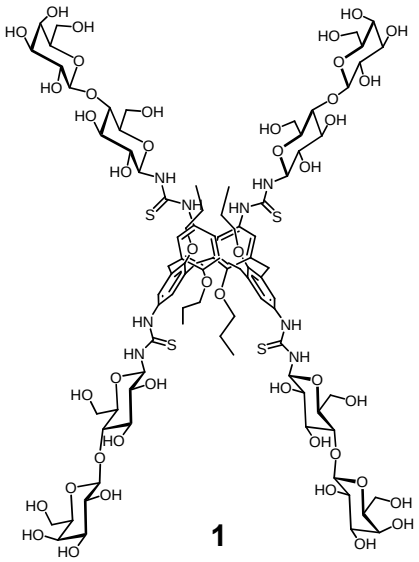
3.2 Results and discussion

3.2.1 Studied compounds

In order to obtain as much NMR information as possible about the ligand-protein interactions, some requirements are necessary. First of all, the studied compounds should be soluble in pure water or in maximum 10% DMSO water solution, in order not to damage or unfold the protein. Second, the ^{1}H NMR spectra should present distinguishable signals for the various protons of the molecule, to allow epitope mapping of the ligands by STD studies.

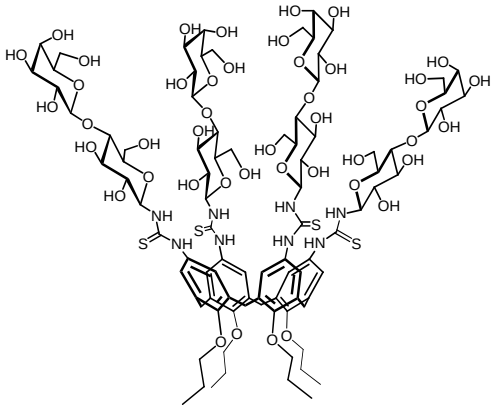
According to their structural features and solubility properties in water, three different calixarenes were selected for this study: *alt*-4Lac[4]Prop **1**, *cone*-4Lac[4]Prop **2** and *cone*-4LacNAc[4]Prop **3** (figure **3.2**). Besides, compound **4** was also synthesized to be used as

monovalent reference. Lacto-compounds **1** and **2** were synthesized according to published procedures[2], while the synthesis of tetra-LacNAc compound **3** has been exhaustively described in Chapter 2 of this thesis.



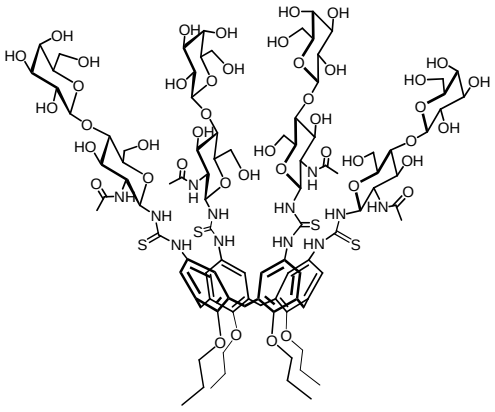
1

alt-4Lac[4]Prop



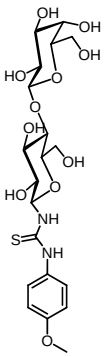
2

cone-4Lac[4]Prop



3

cone-4LacNAc[4]Prop



4

monomer

Figure 3.2 Structures and labels of the studied glycoclusters **1**-**3** and of the monovalent ligand **4**.

Monomer **4** was prepared via an isothiocyanate-amine "click" reaction between isothiocyanate-4-methoxybenzene (commercially available) and peracetylated β-lactosyl

amine **5**[10] to give compound **6**, subsequently deacetylated exploiting the Zemplén procedure[11] to obtain compound **4** (figure **3.3**).

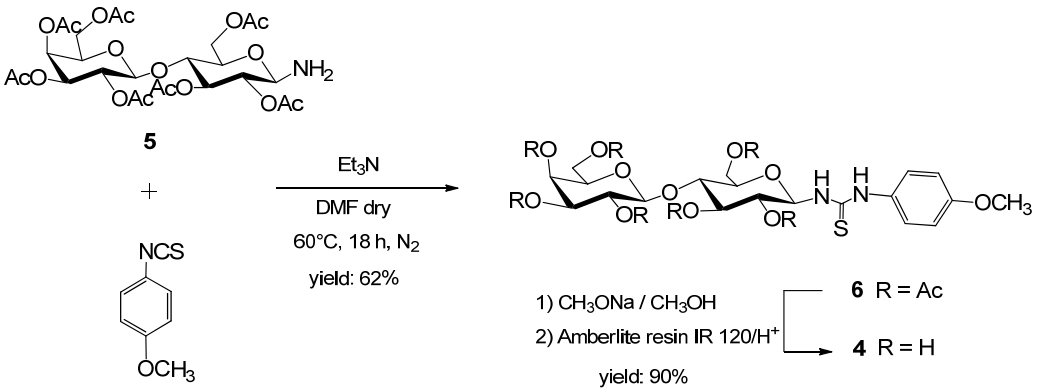


Figure 3.3 Synthesis of the monovalent ligand **4**.

The 1,3-alternate glycocluster **1** and the monomer **4** are well soluble in water, while for the *cone*-glycoclusters **2** and **3** DMSO is necessary to help their solubility in aqueous medium.
Under these conditions, all glycoclusters present ^{1}H NMR spectra with many overlapped signals (figures **3.4**-**3.7**). With the only exception of the anomeric protons H_1 and H_1', in the region between 3 and 4 ppm all the proton resonances of the lactose or LacNAc moiety are grouped together in a complex multiplet. Because of the limited dispersion of chemical shifts of non-anomeric protons, the ^{1}H NMR spectra of the carbohydrate residues are not first order. Nevertheless, standard 2D NMR techniques (COSY, TOCSY and HSQC) were successfully employed for the assignment of the ^{1}H NMR resonances of the molecules **1** and **4** (figure **3.4** and **3.7**). On the contrary, the complete assignment process was precluded for c*one*-glycoclusters **2** and **3**, since they present rather broad and overlapped signals, probably due to aggregate formation in solution.

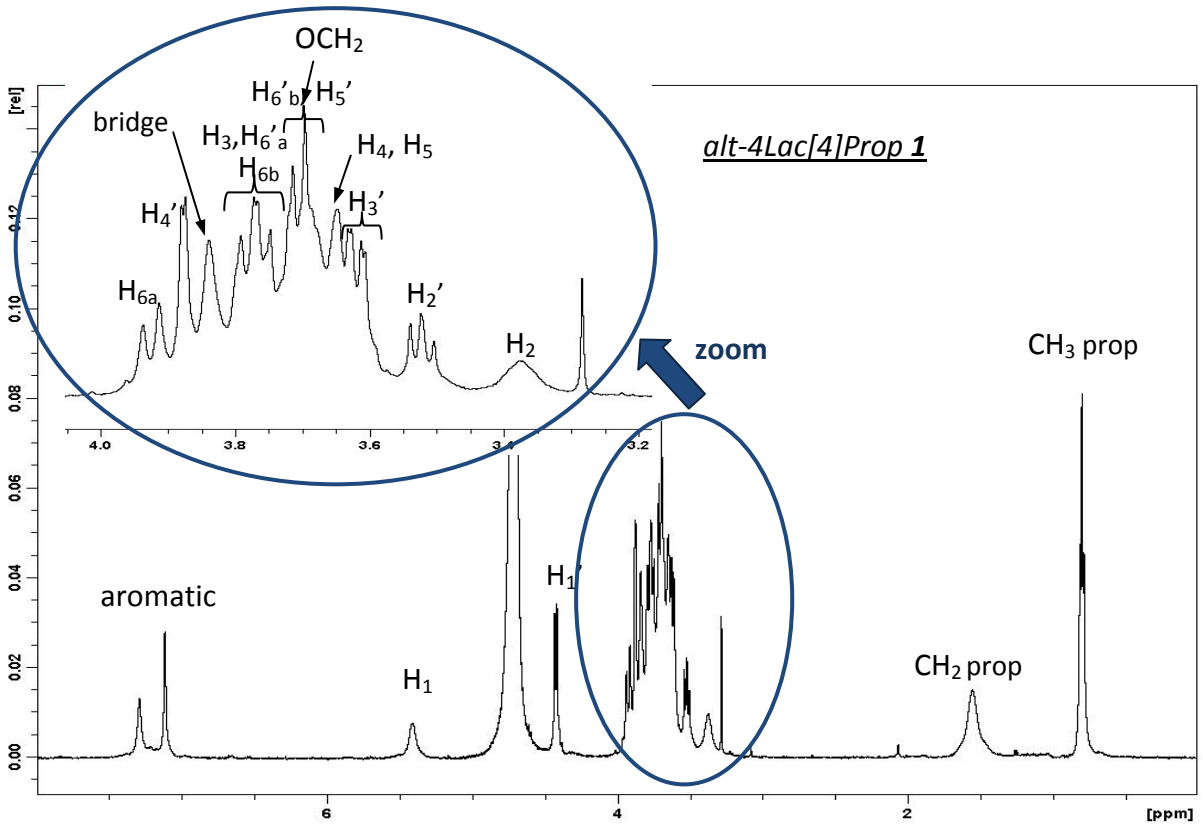


Figure 3.4 ^{1}H NMR 500 MHz of compound **1** (*alt*-4Lac[4]Prop) in D_2O at 298 K, in phosphate buffer, 100 mM NaCl, pH = 7.4.

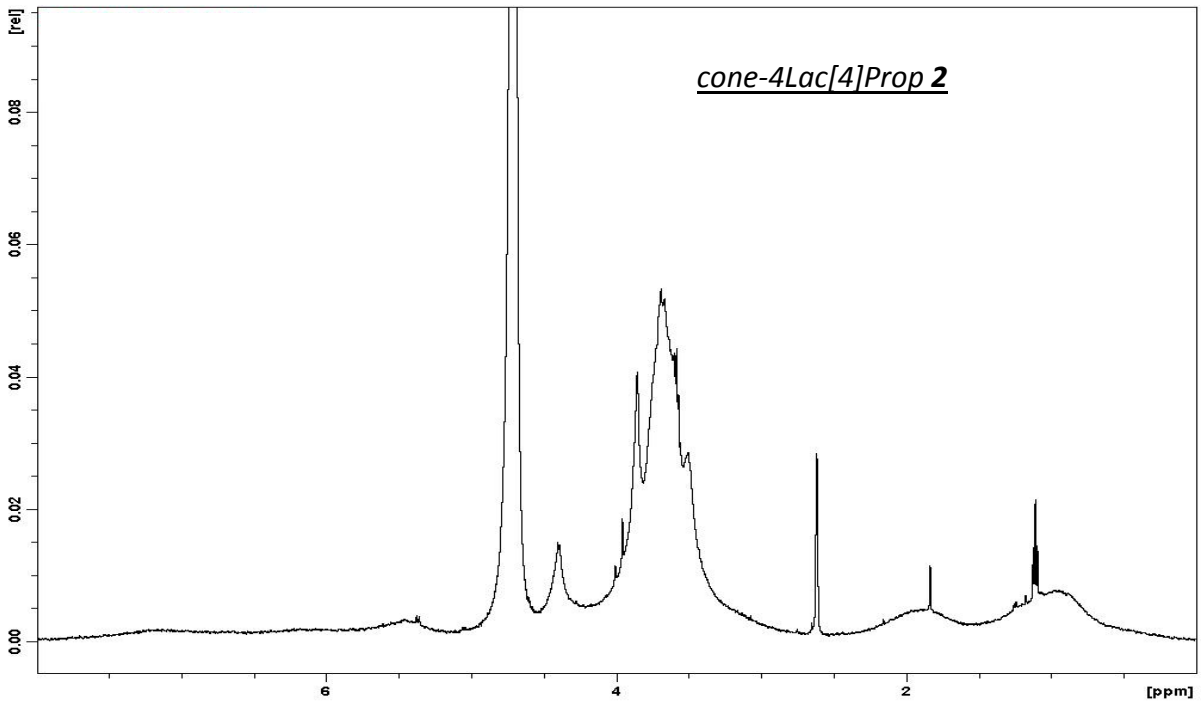


Figure 3.5 ^{1}H NMR 500 MHz of compound **2** (*cone*-4Lac[4]Prop) in D_2O:DMSO 9:1 at 298 K, in phosphate buffer, 100 mM NaCl, pH = 7.4.

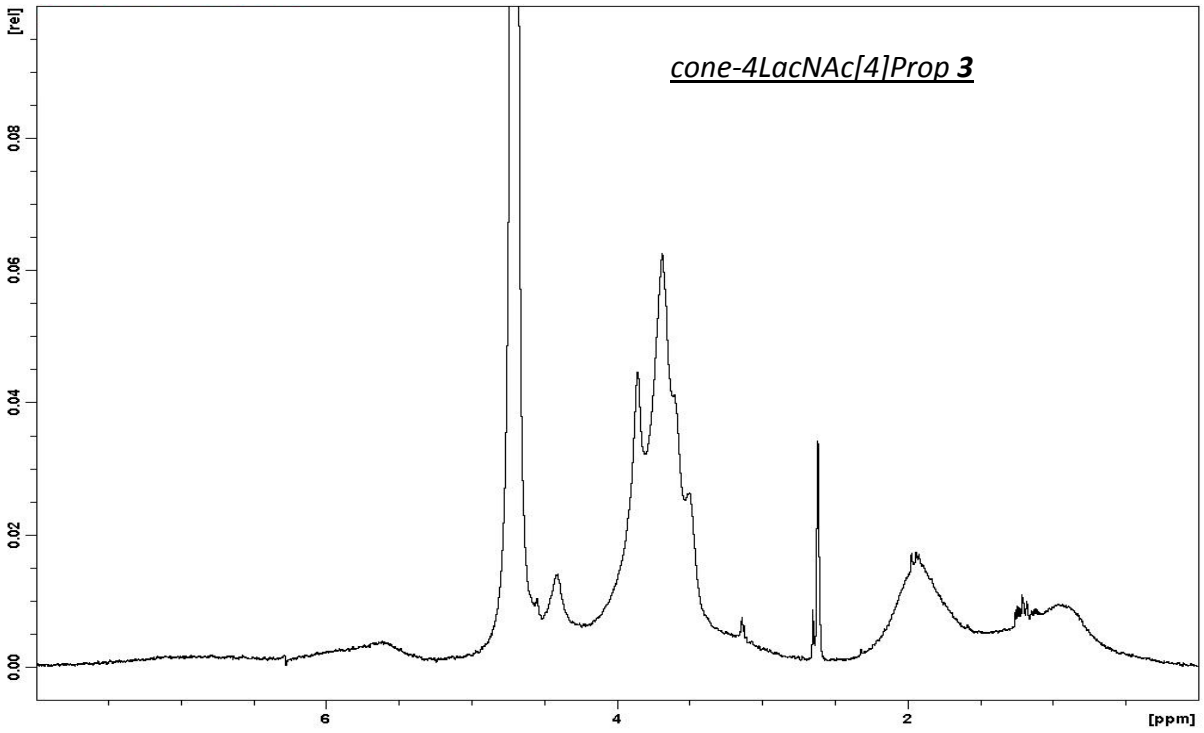


Figure 3.6 ^{1}H NMR 500 MHz of compound **3** (*cone*-4LacNAc[4]Prop)in D_2O:DMSO 9:1 at 298 K, in phosphate buffer, 100 mM NaCl, pH = 7.4.

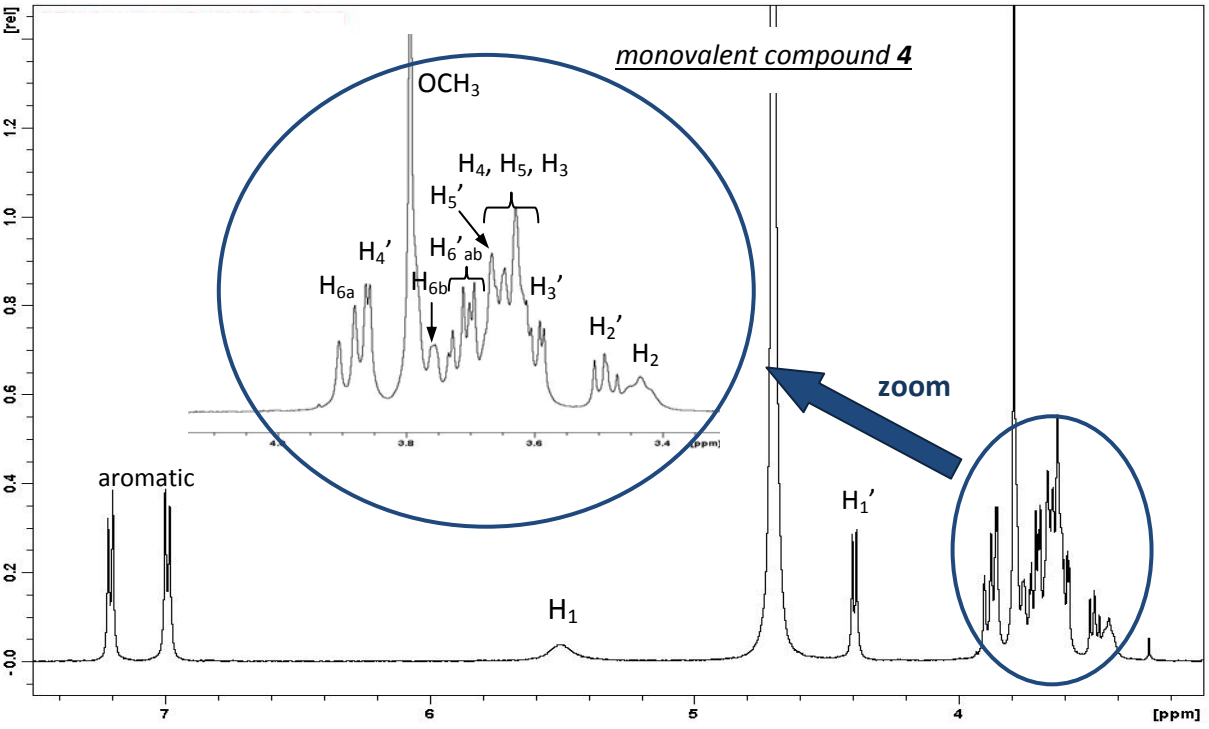


Figure 3.7 ^{1}H NMR 500 MHz of monovalent compound **4** in D_2O at 298 K, in phosphate buffer, 100 mM NaCl, pH = 7.4.

From the protein side, different lectins were selected to be studied. The plant toxin VAA, human galectin-3 in its truncated form (Gal-3), and the N-domain of human galectin-9 (Gal-9N) were chosen. Gal-3 belongs to the chimera type subgroup of galectins, while Gal-9 is member of the tandem-repeat-type subgroup. The two modified galectins were obtained by genetic engineering and furnished by Prof. Gabius (Ludwig Maximilian University of Münich, Germany). In presence of polyglycosylated ligands, galectins can agglutinate, precluding NMR studies. For this reason both galectins were used in their monovalent form, leaving only a single carbohydrate recognition domain on Gal-9N, and avoiding the formation of multivalent species of the protein for Gal-3 by removing its N-terminal non-lectin domain necessary for oligomerization.

3.2.2 Saturation Transfer Difference (STD) NMR experiments

To investigate the binding process, standard STD-NMR experiments were performed for all compounds with protein-ligand molar ratios between 1:10 and 1:200. The off-resonance frequency was set at δ = 100 ppm, while the on-resonance frequency was centred at δ = -1.5 ppm, where only the aliphatic signals of the protein are present. Only for selected experiments (as specified in the text) the on-resonance frequency was set at 0.81 ppm. Before calculating the percentage of STD, the background of the protein was subtracted to the corresponding ligand-lectin STD spectrum, either by suppressing the protein signals with a spin-lock filter in the case of VAA, or by subtracting the STD spectrum of the free protein to the one of ligand-lectin mixture in the case of the two galectins. Possible STD effects due to the self-aggregation of the ligands were also preliminary evaluated by performing STD-NMR experiments for the free compounds under the same conditions of the ones where also the protein was present. Only for *cone*-glycoclusters **2** and **3** these effects were detected, in a large extent already at 1 mM concentration in 10% DMSO aqueous solution. For the analysis of the data, the STD signal intensities were normalized with respect to that with the highest response. STD effects for individual protons of the ligand are reported in percentage and can be classified as very strong (100%-80%), strong (80%-60%), medium (60%-40%), weak (40%-20%) and very weak (< 20%). The epitope map is visualized on the molecules with spheres of different colours (purple, red, orange, yellow and white) and decreasing size corresponding

to decreasing percentage of STD effect. When the ^{1}H NMR spectrum of the studied compound showed overlapping resonances, and therefore it was not possible to distinguish between them, their STD percentage is expressed as an average.

Initially, the complexation of monovalent ligand **4** with each lectin individually (VAA, Gal-3 and Gal-9N) was analyzed.

STD spectrum of **4** (1 mM) in presence of VAA (5µM), molar ratio 200:1, is shown in figure **3.8**, while the absolute and relative STD values are listed in table **3.1**. The quantified STD effects are summarized in figure **3.9**. The most intense STD effects were observed at H_2', H_3' and H_4' protons, followed by H_5' and H_6', indicating the galactose moiety a the major epitope. Interaction of the aglycon part of the molecule with the VAA protein was also observed, although rather weak.

proton	STDabs(%)	STDrel(%)
H1'	1.4	29.6
H2'	4.8	100
H3'	4.0	84.0
H4'	4.0	84.0
H5'	3.3	68.3
H6'a	3.2	67.6*
H6'b	3.2	67.6*
H1	no STD observed	
H2	no STD observed	
H3	2.8	58.4*
H4	2.8	58.4*
H5	2.8	58.4*
H6a	1.4	28.8
H6b	1.0	20.6
H arom	1.6	33.8*
OCH_3	1.1	23.3*

Table 3.1 Absolute and relative percentage of STD for the different protons of **4** interacting with VAA. Asterisks denote overlapping signals.

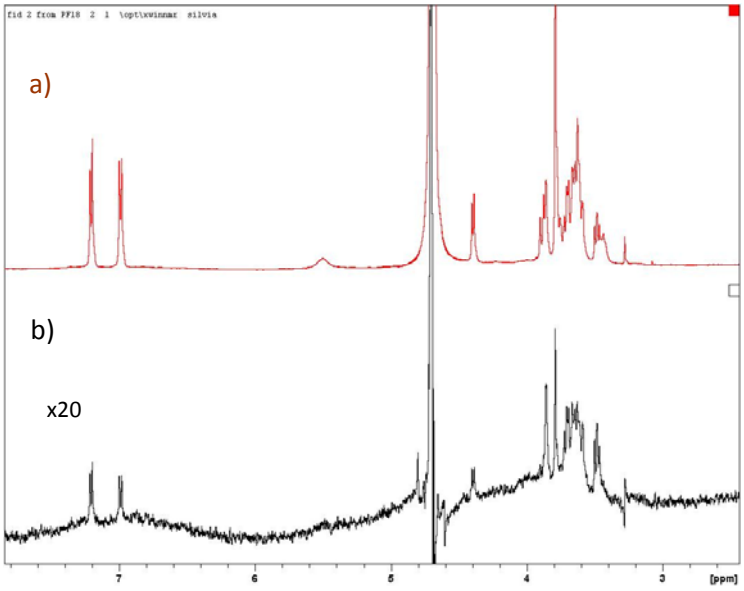


Figure 3.8 a) Reference off-resonance 1H NMR spectrum of the mixture of compound **4** (1 mM) and VAA (5 µM) at 500 MHz and 298 K. b) STD NMR spectrum of the same sample. On-resonance frequency = -1.5 ppm. The spectrum has been scaled up of a factor 20 for better comparison.

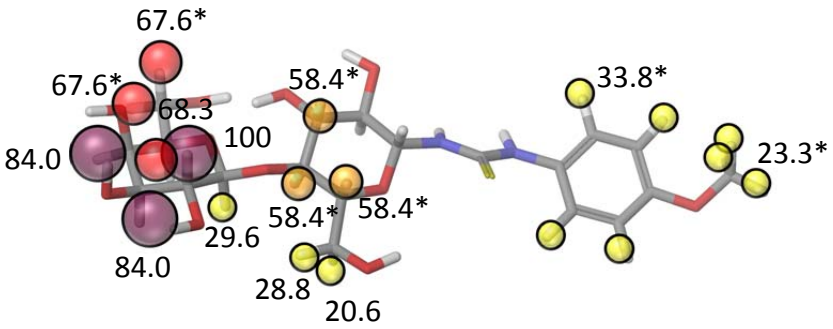


Figure 3.9. Structure of monomer **4** and relative degree of saturation in presence of VAA of the individual protons normalized to that of H_2' proton (that experiences the most intense STD effect) as determined from the STD NMR spectrum. Asterisks denote overlapping signals.

Analogous experiments were performed for compound **4** (4 mM) in presence of Gal-3 (100 µM), molar ratio 40:1. Also in this case the major STD effects were detected for the galactose moiety, but the most intense signals were here detected for H_6' protons, immediately followed by H_4' and H_5' protons (table **3.2** and figures **3.10** and **3.11**).

proton	STDabs(%)	STDrel(%)
H1'	no STD observed	
H2'	no STD observed	
H3'	0.5	51.1
H4'	0.9	92.4
H5'	0.8	82.6
H6'a	0.9	100*
H6'b	0.9	100*
H1	no STD observed	
H2	no STD observed	
H3	0.4	41.3*
H4	0.4	41.3*
H5	0.4	41.3*
H6a	0.3	32.6
H6b	0.3	32.6
H arom	0.2	16.3*
OCH_3	0.1	14.1*

Table 3.2 Absolute and relative percentage of STD for the different protons of **4** interacting with Gal-3. Asterisks denote overlapping signals.

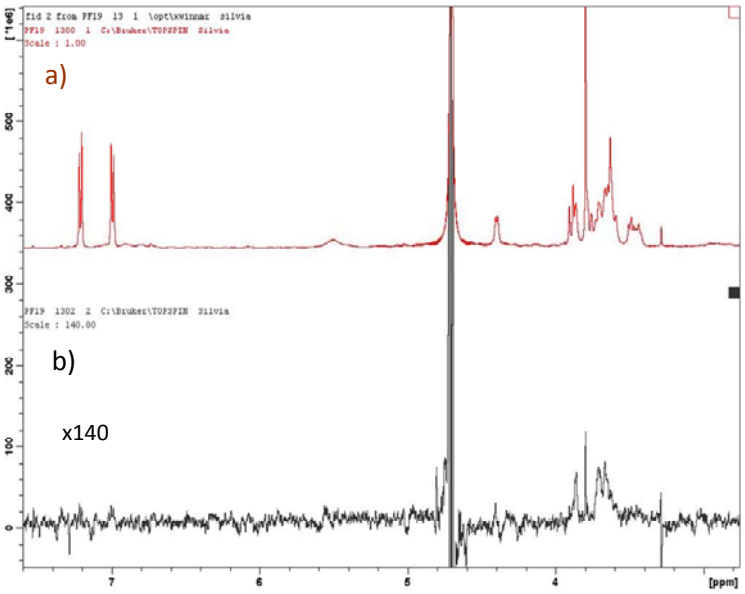


Figure 3.10 a) Reference off-resonance ^{1}H NMR spectrum of the mixture of compound **4** (4 mM) and Gal-3 (100 µM) at 500 MHz and 298 K. b) STD NMR spectrum of the same sample. On-resonance frequency = -1.5 ppm. The spectrum has been scaled up of a factor 140 for better comparison.

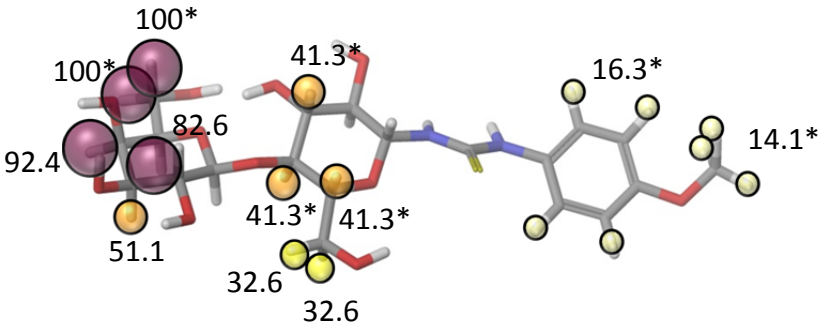


Figure 3.11 Structure of monomer **4** and relative degree of saturation in presence of Gal-3 of the individual protons normalized to that of H_6' protons (that experience the most intense STD effect) as determined from the STD NMR spectrum. Asterisks denote overlapping signals.

Thus there are subtle differences between VAA and Gal-3 in the binding of the same substrate, as already described for other galactose-containing saccharides and glycomimetics[12]. In this case, only a very weak interaction of the ligand aromatic moiety was detected (STD relative percentages below 20%), indicating its minor involvement in the binding.

For the mixture of **4** with Gal-9N, too weak STD signals were obtained setting the on-resonance at -1.5 ppm. For this reason the on-resonance frequency was shifted at 0.81 ppm to get a better signal/noise ratio for the STD spectrum. At 0.81 ppm in fact there are more protons of the protein and, as a consequence, the intensities of the STD signals increased compared to those registered at -1.5 ppm. The STD experiment with Gal-9N was performed using the same molar ratio and concentrations as for the studies with Gal-3. The results obtained in this case were also similar. As for the experiments with Gal-3, H_5' and H_6' are the protons receiving the most intense saturation effect. This shows that the C4, C5, C6 region of galactose is the portion of ligand mainly involved in the binding with galectins (table **3.3** and figures **3.12** and **3.13**), as typically observed in other cases[12].

proton	STDabs(%)	STDrel(%)
H1'	0.6	6.1
H2'	1.4	15.2
H3'	0.6	6.2
H4'	2.2	24.0
H5'	9.1	100
H6'a	4.9	54.0*
H6'b	4.9	54.0*
H1	no STD observed	
H2	2.0	22.3
H3	3.1	33.6*
H4	3.1	33.6*
H5	3.1	33.6*
H6a	0.7	7.8
H6b	0.7	7.8
H arom	1.2	13.7*
OCH_3	0.9	9.4*

Table 3.3 Absolute and relative percentage of STD for the different protons of **4** interacting with Gal-9N. Asterisks denote overlapping signals.

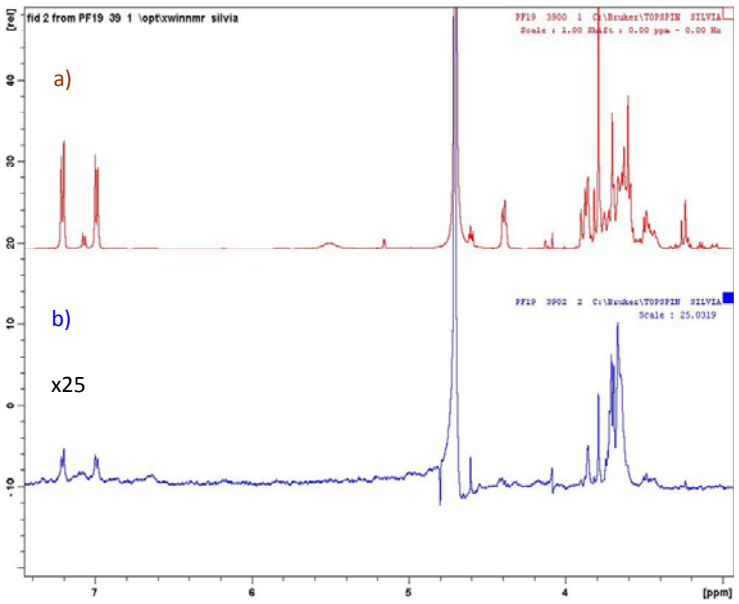


Figure 3.12 a) Reference off-resonance ^{1}H NMR spectrum of the mixture of compound **4** (4 mM) and Gal-9N (100 µM) at 500 MHz and 298 K. b) STD NMR spectrum of the same sample. On-resonance frequency = 0.81 ppm. The spectrum has been scaled up of a factor 25 for better comparison.

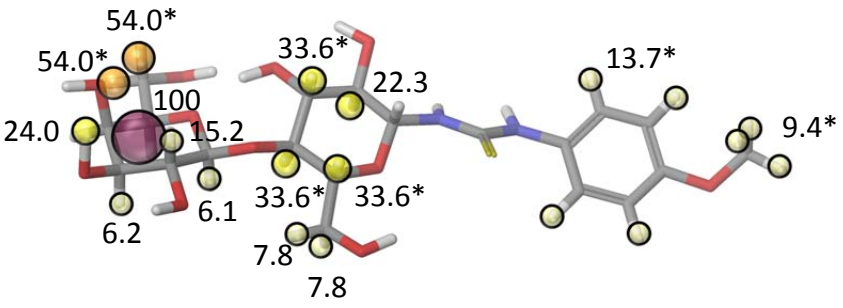


Figure 3.13. Structure of monomer **4** and relative degree of saturation in presence of Gal-9N of the individual protons normalized to that of H_5' proton (that experiences the most intense STD effect) as determined from the STD NMR spectrum. Asterisks denote overlapping signals.

STD-NMR experiments were then used to study the mode of binding of *alt*-4Lac[4]Prop **1** to the studied lectins. Glycocluster **1** is fairly soluble in water up to the tested 1 mM concentration, and therefore it was not necessary to add any DMSO as cosolvent. For this compound important overlapping of different protons was observed (see figure **3.4**). For them the percentage of STD is expressed as an average of the different signals. Besides, protons H_1 and H_2 of glucose are rather broad already in the free ligand state, due to the

presence of the thiourea linked calixarene moiety, which gives rise to equilibrium between the different possible rotamers (*vide infra* paragraph **3.2.3.1**). Therefore it resulted difficult to establish the exact percentage of STD that these two protons experience, and the values reported for them cannot be considered accurate.

STD experiments were recorded for a sample containing compound **1** (1 mM) in presence of VAA 5 µM (molar ratio 1:200). Clear STD signals were observed for most of the protons, including the aromatic and alkyl moieties, as resulted from the STD spectrum (figure **3.14**) and depicted in figure **3.15**. Absolute and relative percentage of STD for the different protons of **1** interacting with VAA are summarized in table **3.4**.

proton	STDabs(%)	STDrel(%)
H1'	0.8	58.2
H2'	1.4	100
H3'	1.0	73.0
H4'	1.2	82.3
H5'	1.0	68.1*
H6'a	1.0	70.9*
H6'b	1.0	68.1*
H1	(1.2)	(85.1)
H2	(1.1)	(78.7)
H3	1.0	70.9*
H4	1.0	69.5*
H5	1.0	69.5*
H6a	1.1	74.5
H6b	1.0	70.9*
H arom	1.2	85.1*
CH_2 bridge	1.1	79.4*
OCH_2 prop	1.0	68.1*
CH_2 prop	not determined	
CH_3 prop	not determined	

Table 3.4 Absolute and relative percentage of STD for the different protons of **1** interacting with VAA. Asterisks denote overlapping signals. Values in brackets are referred to STD values that could not be accurately determined.

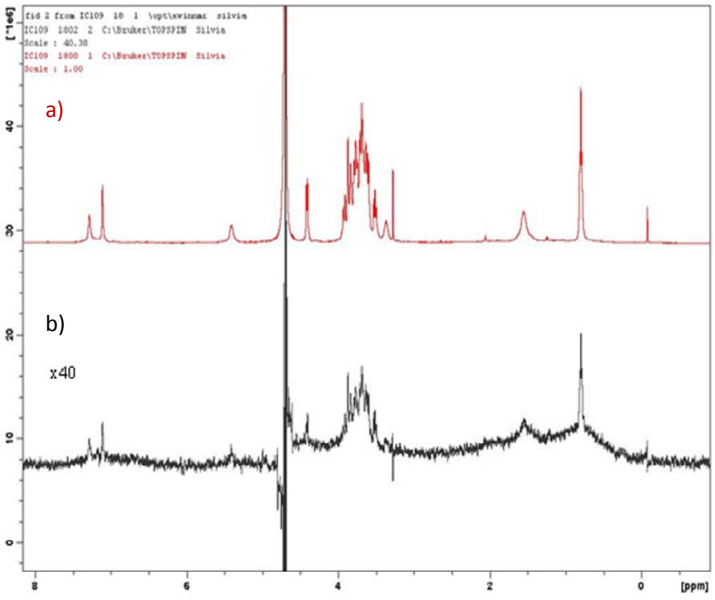


Figure 3.14 a) Reference off-resonance ^{1}H NMR spectrum of the mixture of compound **1** (1 mM) and VAA (5 µM) at 500 MHz and 298 K. b) STD NMR spectrum of the same sample. On-resonance frequency = -1.5 ppm. The spectrum has been scaled up of a factor 40 for better comparison.

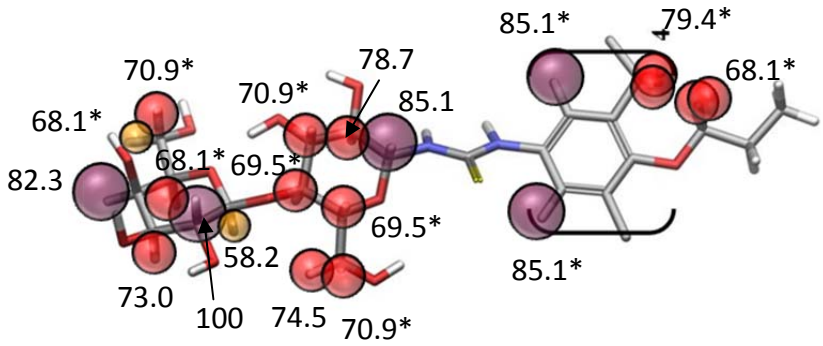


Figure 3.15. Structure of the monomeric repeating unit of calixarene **1** and relative degree of saturation in presence of VAA of the individual protons normalized to that of H_2' proton (that experiences the most intense STD effect) as determined from the STD NMR spectrum. Asterisks denote overlapping signals.

Regarding the lactose region, the interaction pattern typical for VAA was observed relatively to the galactose unit. The highest percentage of STD are in fact for H_2' and H_4' protons of this saccharide. On the other hand, all the protons of the ligand seem to receive in general a quite intense and similar STD effect. In the aliphatic region the signals of VAA are very intense and overlapped with the propyl protons of the calixarene, so that for these protons it was not possible to determine the STD percentage in an accurate way, even if it is clear from the STD spectrum that they interact with the protein.

In contrast, the STD spectra for **1** (1 mM) in the presence of 100 µM Gal-3 (molar ratio 1:10) indicated that, in this case, the interaction was fairly selective for the galactose moiety, with the major saturation transfer detected to its H_6' and H_4' protons (figure **3.16** and **3.17** and table **3.5**). The STD effects at the propyl chain of ligand **1**, and strictly confined to this part of the calixarene scaffold, could be explained considering the 1,3-alternate structure of this glycocalixarene. Because of the similar direction in space of the lactoside of one calixarene aromatic unit and of the alkyl chain of the proximal aromatic unit imposed by this conformation, the binding between the saccharide and the recognition site of the lectin could hold the propyl substituent in rather close proximity of the protein surface.

proton	STDabs(%)	STDrel(%)
H1'	no STD observed	
H2'	1.3	93.0
H3'	1.3	87.4
H4'	1.4	94.4
H5'	1.2	83.9*
H6'a	0.8	55.2*
H6'b	1.4	100
H1	no STD observed	
H2	no STD observed	
H3	0.8	55.2*
H4	0.7	51.0*
H5	0.7	51.0*
H6a	0.8	55.2
H6b	0.8	55.2*
H arom	no STD observed	
CH_2 bridge	no STD observed	
OCH_2 prop	1.2	83.9*
CH_2 prop	no STD observed	
CH_3 prop	0.7	49.0*

Table 3.5 Absolute and relative percentage of STD for the different protons of **1** interacting with Gal-3. Asterisks denote overlapping signals.

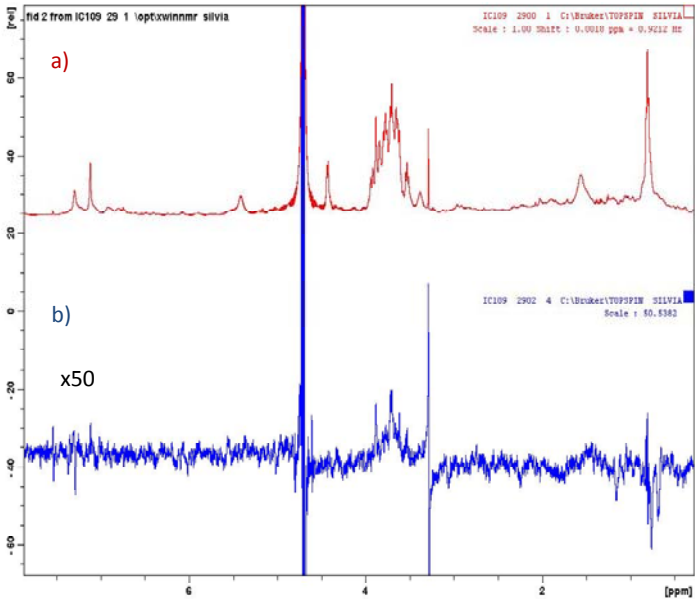


Figure 3.16 a) Reference off-resonance ^{1}H NMR spectrum of the mixture of compound **1** (1 mM) and Gal-3 (100 µM) at 500 MHz and 298 K. b) STD NMR spectrum of the same sample. On-resonance frequency = -1.5 ppm. The spectrum has been scaled up of a factor 50 for better comparison.

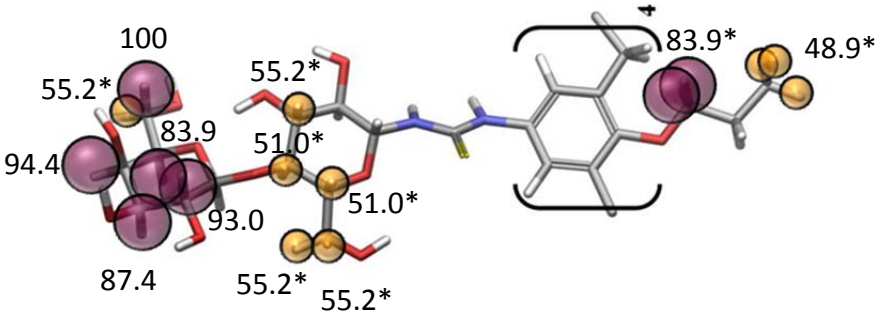


Figure 3.17. Structure of the monomeric repeating unit of calixarene **1** and relative degree of saturation in presence of Gal-3 of the individual protons normalized to that of H_6' proton (that experiences the most intense STD effect) as determined from the STD NMR spectrum. Asterisks denote overlapping signals.

Since some resonances of the calixarene ligand are in the aliphatic region due to the presence of the propyl chains at the lower rim, in the experiments with Gal-9N it was not possible to set the on-resonance frequency at 0.81 ppm. Setting the on-resonance at -1.5

ppm as for the Gal-3 experiments, no satisfying STD signal could be obtain for an epitope mapping. Nevertheless it was possible to observe the presence of an STD effect denoting that the compound **1** is however interacting with the protein Gal-9N.

In the graphs below (figures **3.18**-**3.20**) are summarized the percentage of saturation obtained for the sugar epitope for ligand **1** and **4** in the complexation with the three different lectins, allowing an easy comparison. The relative percentage of saturation is displayed for each proton of the lactose unit. Although some STD signals were found also for the aromatic and alkoxy moieties of both ligands, the aglycon part is not reported in these graphs, since we believe they are not giving rise to specific noncovalent interactions with the protein.

With VAA, for both compounds **1** and **4**, the proton that experience the highest STD effect is H_2' (figure **3.18**). Anyway, while for **4** the galactose protons are by far the ones with the greatest saturation percentage, for **1** also all the glucose protons seem to receive a quite intense and similar STD effect.

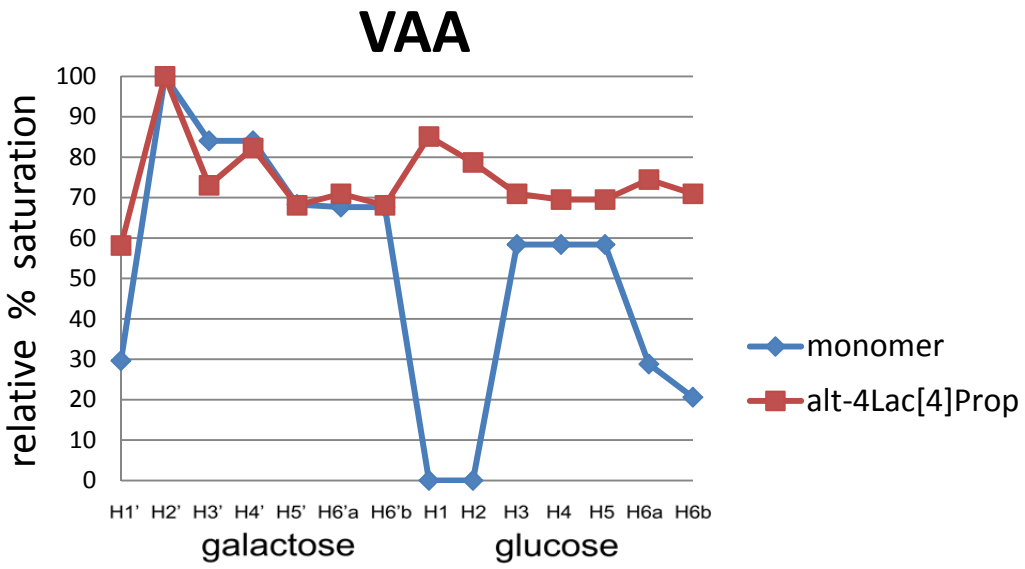


Figure 3.18 Schematic representation of the STD effect for monomer **4** and calixarene **1** with VAA.

With Gal-3, the epitope map for monomer **4** is very similar to that of lactose (used as model), with protons H_4', H_5' and H_6' of galactose the most involved in the recognition

process. Also for calixarene **1** these protons are strongly involved in the binding, even if H_2' and H_3' show comparably strong STD effects (figure **3.19**).

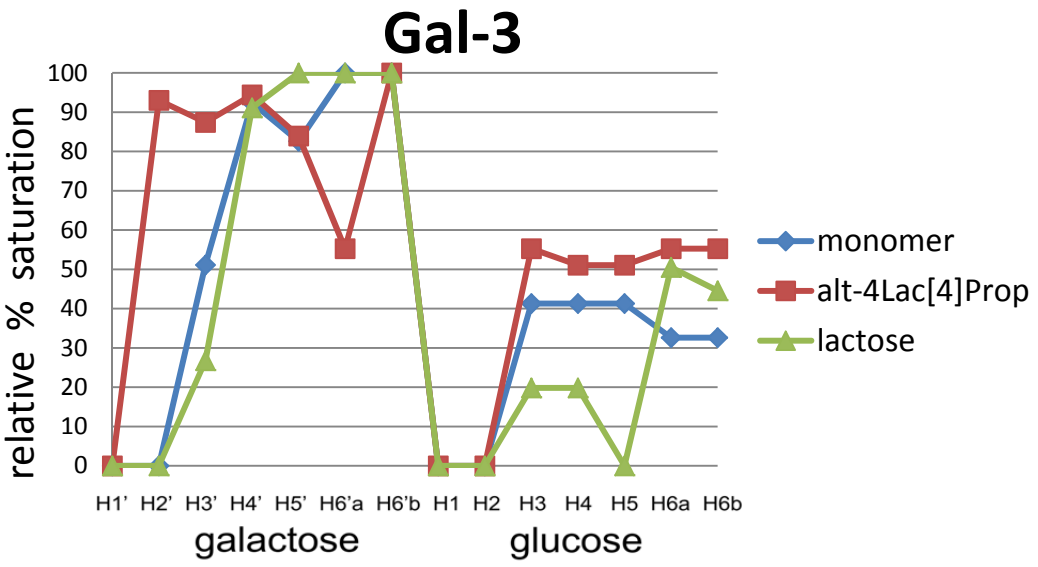


Figure 3.19 Schematic representation of the STD effect for monomer **4**, calixarene **1** and lactose (used as model) with Gal-3.

Accordingly, the major epitope for the monomer **4** and the glycocalixarene **1** is conserved in the case of VAA and of Gal-3. VAA mostly recognizes the C2-C4 region of the galactose moiety, while Gal-3 mainly interacts with the C4-C6 galactose region. VAA shows also transient interactions with the aromatic part of the monomeric ligand and even more with the calixarene scaffold, while the Gal-3 basically bind to the sugar moiety and only minor interactions with the calixarene skeleton were detected. In all cases, the recognition pattern of these molecules is consistent and similar to that observed for lactose.

With Gal-9N the epitope mapping was recorded with the on-resonance frequency set at 0.81 ppm. In this case monomer **4** presents the same epitope of lactose (used as model for comparison), with H_5' and H_6' protons of galactose experiencing the most intense STD effect (figure **3.20**). It is worth to remember that the epitope mapping for calixarene **1** could not be determined with Gal-9N because around the on-resonance frequency used in these experiments, the methyl groups of the ligand also resonate.

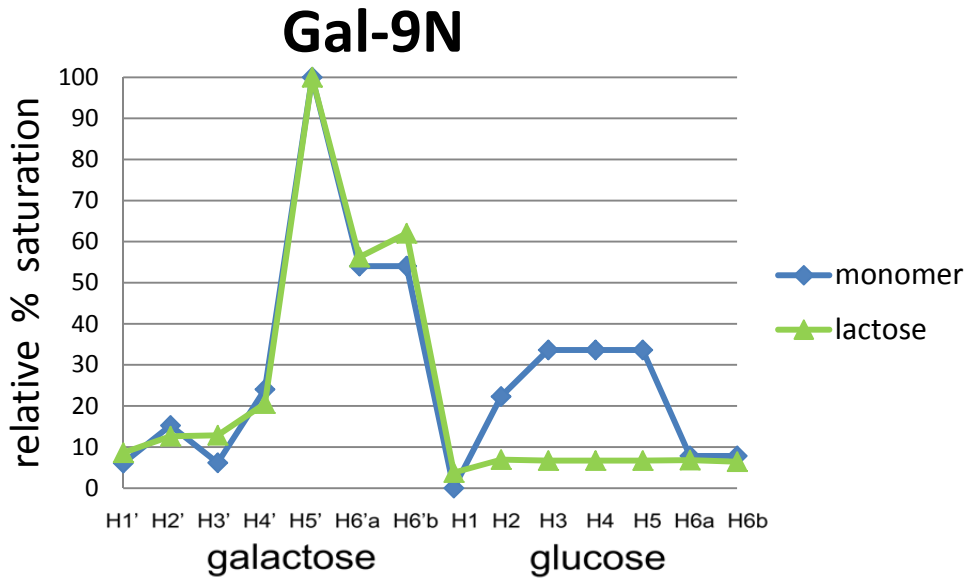


Figure 3.20 Schematic representation of the STD effect for monomer **4** and lactose (used as model) with Gal-9N.

^{1}H NMR spectra obtained for compounds **2** and **3** were of poor quality (figure **3.5** and **3.6**), with many broad signals, indicating that very important self-aggregation phenomena occur. Considering that practically all the protons for these two glycoclusters are together in a large multiplet between 3 and 4 ppm, no epitope mapping can be deduced from the analyses, nevertheless the percentage of saturation can be calculated as an overall average. Therefore STD experiments in presence of VAA and Gal-3 were anyway carried out. Since the glycoclusters **2** and **3** resulted poorly soluble in water, to all solutions DMSO (max. 10%) was added to help solubility.

Addition of any lectin (VAA 5 µM, Gal-3 100 µM) to *cone*-4Lac[4]Prop **2** (1 mM) immediately produced a white solid precipitate. This fact strongly indicates that this calixarene is able to bind to more than one lectin molecule simultaneously, causing protein agglutination. STD NMR analysis on the surnatant revealed the presence of STD signals, but too weak to be quantified.

For glycocalixarene **3** no solid precipitate was produced upon addition of VAA or Gal-3 and an overall STD effect could be detected. For both VAA (figure **3.21**) and Gal-3 (figure **3.22**), evident STD in the sugar region was observed indicating the recognition of the lactose moiety.

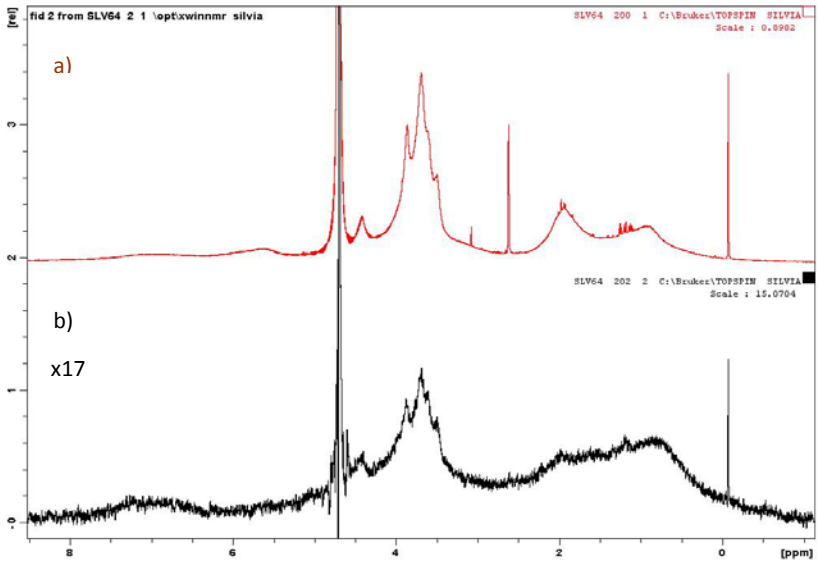


Figure 3.21 a) Reference off-resonance spectrum of a mixture of calixarene *cone*-4LacNAc[4]Prop **3** (1 mM) and VAA (5 µM) in a ratio 200:1. b) STD-NMR spectrum of the same sample. The spectrum is scaled up of a factor 17 for better comparison.

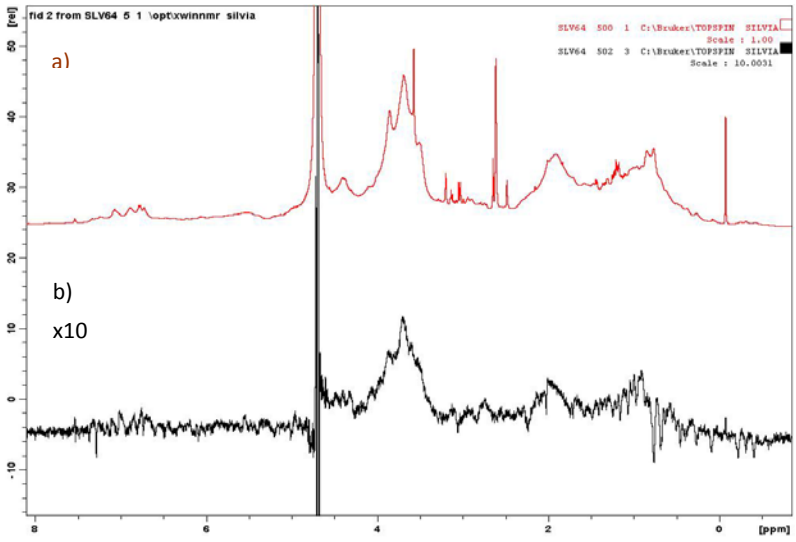


Figure 3.22 a) Reference off-resonance spectrum of a mixture of calixarene *cone*-4LacNAc[4]Prop **3** (1 mM) and Gal-3 (100 µM) in a ratio 10:1. b) STD-NMR spectrum of the same sample. The spectrum is scaled up of a factor 10 for better comparison.

3.2.3 Diffusion Ordered SpectroscopY (DOSY) NMR experiments

DOSY experiments were performed to obtain semiquantitative information about the stoichiometry (molecular size) of the protein-ligand complexes, and therefore to check if multivalent effects take place in glycocalixarene-(ga)lectin interactions.

Accurate determination of diffusion coefficients can be affected by several errors. Temperature and viscosity changes in the sample must be avoided and gradients must be correctly calibrated[13]. To overcome these challenges and obtain reliable molecular size data in a straightforward way, the use of a calibration curve can be a useful simple strategy[14]. Using five different proteins of different molecular weight, a calibration curve was therefore built. The selected proteins were: 1) aprotinin (6.5 kDa), 2) α-lactalbumin (14 kDa), 3) carbonic anhydrase (29 kDa), 4) ovalbumin (44 kDa) and 5) bovin serum albumine (66kDa), as depicted in figures **3.23** and **3.24**.

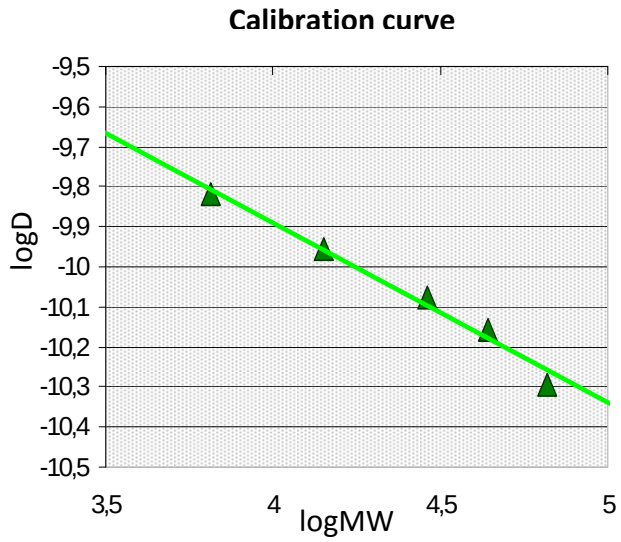


Figure 3.23 Plot of logD versus logMW for five different proteins of different molecular weight to obtain a calibration curve. A least-squares linear correlation was fitted to the data according to the equation:

logD= -0.4492logMW - 8.0937, r^2= 0.9812.

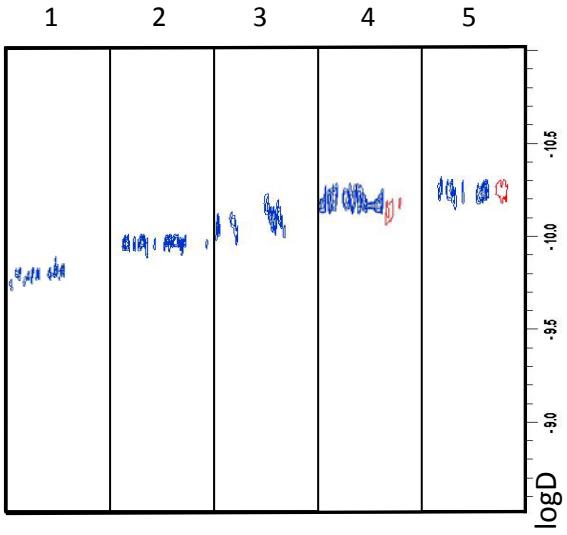


Figure 3.24 Alternative representation of the data presented in the figure **3.23**. Narrow strips of the 2D DOSY spectra, corresponding to the proton region between 0 and 2 ppm are taken and displayed side by side.

Then, the binding of the compounds **1**-**4** to Gal-3 was studied. Thanks to the calibration curve obtained, from logD values it was possible to determine the molecular weight of the protein-ligand complexes. In figure **3.25** are summarized the results obtained for calixarene **1** and for the monovalent reference **4**.

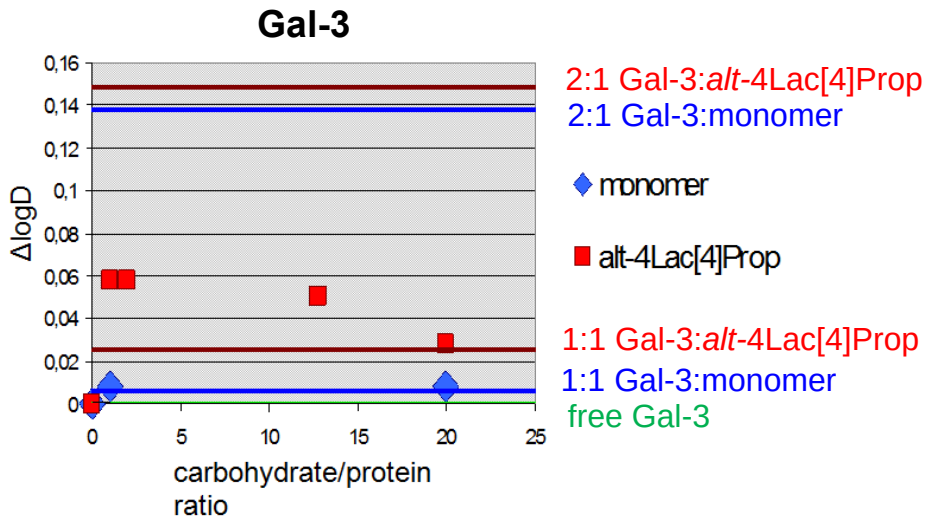


Figure 3.25 Plot of the change in logD (ΔlogD) of Gal-3 as a function of equivalents of sugar units added in case of calixarene **1** or monovalent ligand **4**. The predicted ΔlogD values for free Gal-3 as well as for the 1:1 and 2:1 protein : ligand complexes are indicated with coloured solid lines in the chart.

Sum projections of the diffusion dimension for interaction of **1** and **4** with Gal-3 at different protein-ligand ratios are displayed in figures **3.27** and **3.26**, respectively.

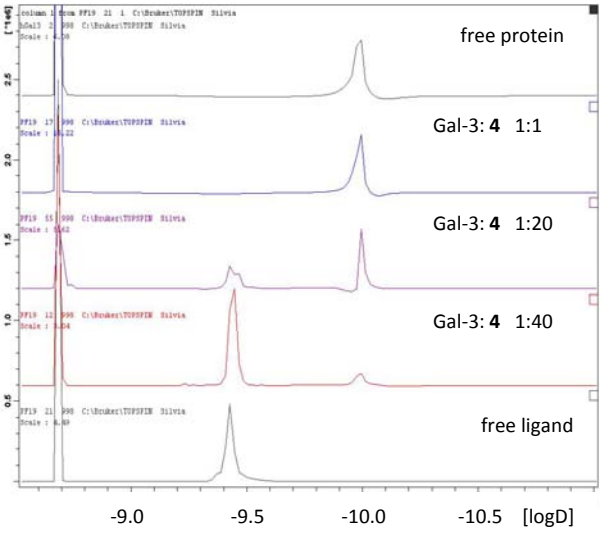


Figure 3.26 Sum projections of the diffusion dimension for monomer **4** (peak on the left) and Gal-3 (peaks on the right) at different protein:ligand ratios.

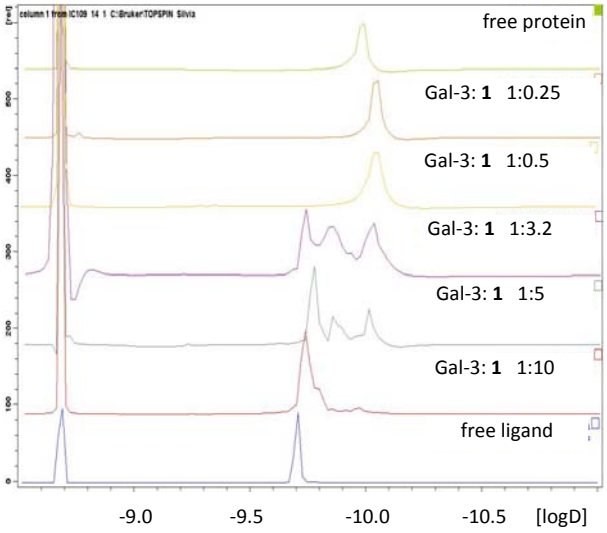


Figure 3.27 Sum projections of the diffusion dimension for 1,3-alternate calixarene **1** (peak on the left) and Gal-3 (peaks on the right) at different protein:ligand ratios.

In the case of the monomer **4**, as expected, the diffusion coefficient is always the same and independent from the equivalents of carbohydrate added. Also increasing the concentration of the ligand up to 40 equivalents (4 mM concentration), the diffusion coefficient does not change, always presenting a value compatible with a 1:1 protein:ligand complex and indicating that no multivalent effect is taking place.

Since each calixarene is functionalized with four lactose units, the equivalents of ligand used for each experiment were divided by four, in order to maintain the same lactose equivalents as for monovalent ligand **4**, to allow a direct comparison. *Alt*-4Lac[4]Prop calixarene **1** showed the stoichiometry of the complex to be between 1:1 and 2:1 protein:ligand. Titration up to 2 equivalents of sugar (corresponding to 0.5 equivalents of glycocluster, 75 µM concentration) resulted in an increase in the logD values of Gal-3 higher than the one expected for the 1:1 protein:ligand stoichiometry, approaching the 2:1 stoichiometry. The titration between 2 and 20 equivalents of sugar (ligand concentration 1 mM max.) resulted in a decrease in logD back toward the one expected for a 1:1 complex. Therefore, the DOSY experiments for this system showed the presence of multivalent effects, with **1** interacting with more than one Gal-3 at the same time.

Analogous DOSY experiments were performed with the *cone*-calixarenes **2** and **3**. The two glycoclusters showed very similar behaviour in solution with Gal-3 (figure **3.28**). In both cases the free ligand, at 1 mM concentration, presents a diffusion coefficient higher than the galectin itself (as shown in figure **3.29** for calixarene **2** and in figure **3.30** for calixarene **3**). This result strongly confirms, as anticipated by the monodimensional ^{1}H NMR spectra, that in solution *cone*-4Lac[4]Prop **2** and *cone*-4LacNAc[4]Prop **3** form aggregates. One possible explanation is that these compounds are able to form micelles. Due to the presentation of the sugar moieties at the same face of the calixarene backbone, in **2** and **3** the whole molecule distinctly shows a hydrophilic and a hydrophobic part. Very likely these compounds, in water solution, can aggregate by exposing the carbohydrate units (hydrophilic part of the molecules) to the aqueous medium and the propyl chains (hydrophobic part) facing inward. Using the calibration curve, an average number of thirteen calixarene units of either compound **2** or **3** was calculated to be aggregated together.

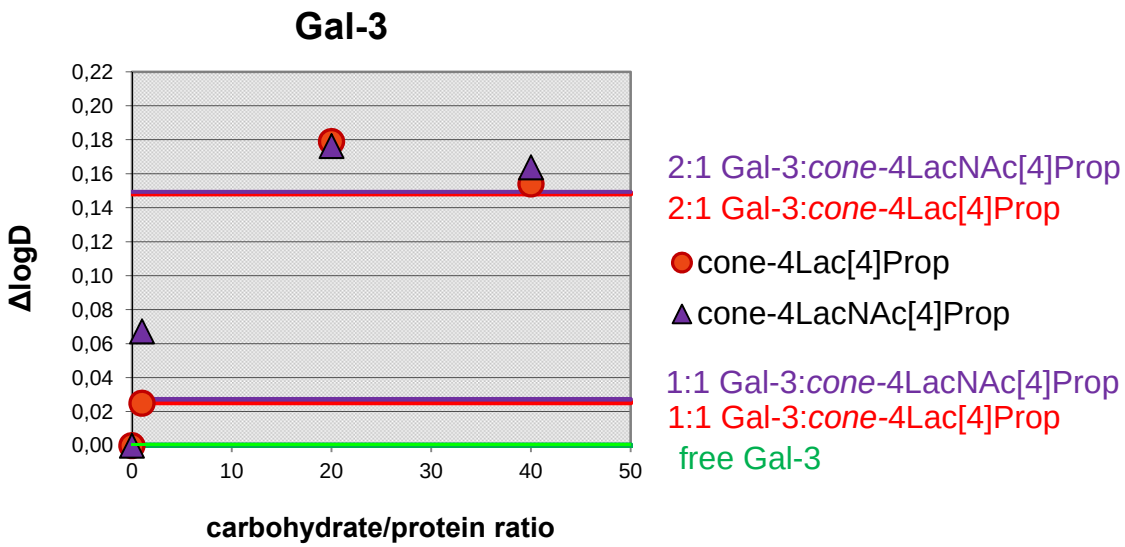


Figure 3.28 Plot of the change in logD (ΔlogD) of Gal-3 as a function of equivalents of sugar for *cone*-4Lac[4]Prop **2** or *cone*-4LacNAc[4]Prop **3** added. The predicted ΔlogD values for free Gal-3 as well as for the 1:1 and 2:1 protein : ligand complexes are indicated with coloured solid lines in the chart.

For both compounds **2** and **3**, titration up to 0.25 ligand equivalents resulted in a logD value indicating the formation of a 1:1 complex. In fact, the concentration of ligand for this first DOSY experiment was 50 µM, minimizing the aggregation process. When the equivalents of ligand were increased to 5 (1 mM concentration), a logD value higher than the one corresponding to the 2:1 protein:ligand stoichiometry was detected. Subsequent titration with 10 equivalents of ligand (always at 1 mM concentration and decreasing the protein concentration) resulted in a decrease in logD back towards the value calculated for the 2:1 protein:ligand complex. Actually, in the case of the *cone*-calixarenes, the demonstrated self-aggregation phenomena of the glycosylated ligands make very difficult the interpretation of the collected data and it is not clear the nature of the species present in solution, although indications of interactions between the glycoclusters and Gal-3 are unequivocal. The occurrence of a multiple binding between the galectins and the *cone*-derivatives (or an aggregated species of them) can be supposed on the basis of the increased inhibitory efficiency with respect to monovalent lactose units, as previously reported[2].

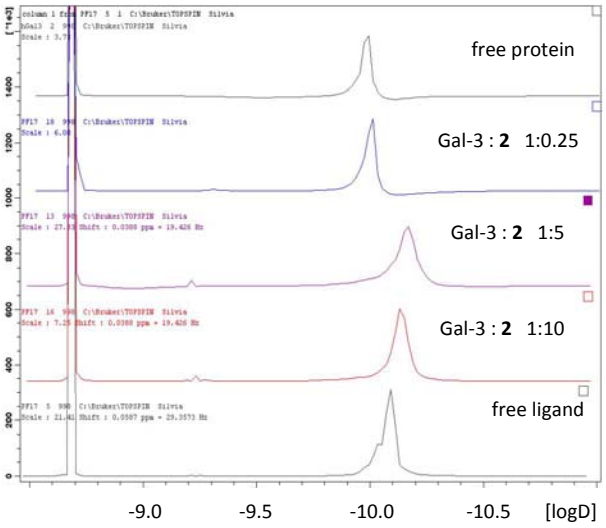


Figure 3.29 Sum projections of the diffusion dimension for *cone*-4Lac[4]Prop **2** and Gal-3 at different protein : ligand ratios.

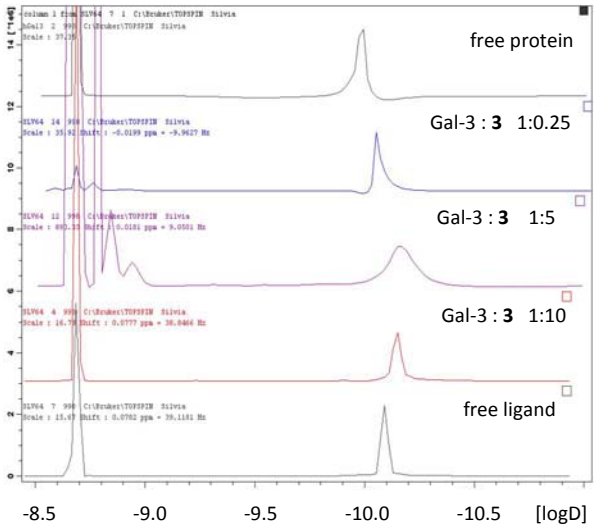


Figure 3.30 Sum projections of the diffusion dimension for *cone*-4LacNAc[4]Prop **3** and Gal-3 at different protein : ligand ratios.

3.2.4 Molecular Modeling (MM)

In many cases, not all the required information concerning the carbohydrate-protein interaction can be obtained directly by experimental studies even though, as previously shown, NMR experiments can be extremely informative. Therefore molecular modeling is most of the time necessary to support and supplement experimental data. Modeling protocols were here employed to provide a more detailed picture of the ligand-lectin interaction.

3.2.4.1 Modeling studies on the E/Z isomerism of the thiourea bond

The first challenge was to determine the structure of the ligands and in particular to find out which geometries for the thiourea moieties linking the sugars to the aromatic scaffolds are the most stable, and therefore most abundant, in water solution for the various compounds studied. In fact, the thiourea group can show ZZ, ZE, EZ or EE presentation, depending on the dihedral angles between N-Lac and C=S bonds and between N-aromatic and C=S bonds (Figure **3.31**).

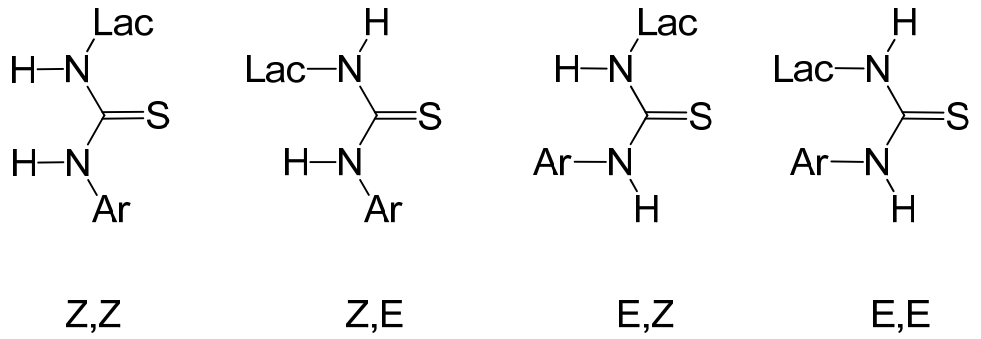


Figure 3.31 Schematic representation of the four possible geometries of an unsymmetrically substituted thiourea.

Most thioureido sugars present broad proton signals in ^{1}H NMR spectra, in particular for the anomeric and the NHs protons, even in solvents of high donicity such as DMSO-d_6[10,15]. In the case of our compounds, this broadening characterizes also the aromatic protons of the scaffold. This effect is due to the restricted rotation around the NH–CS bonds, and therefore the instauration of equilibrium processes among the different rotamers. In D_2O all the signals are even more broad and overlapped, and the NH protons completely disappear in the ^{1}H NMR spectrum recorded in standard conditions. To observe them for the compound **4**, it was necessary to record a ^{1}H NMR spectrum with a water-gate sequence in H_2O:D_2O 9:1 in phosphate buffer at pH=6 at 278K (figure **3.32**). Anyway the analyses of 1D and 2D spectra, including selective NOESY experiments, did not lead to conclusions on the issue of which rotamers are present in solution. Therefore molecular modeling protocols were employed to obtain information on this topic.

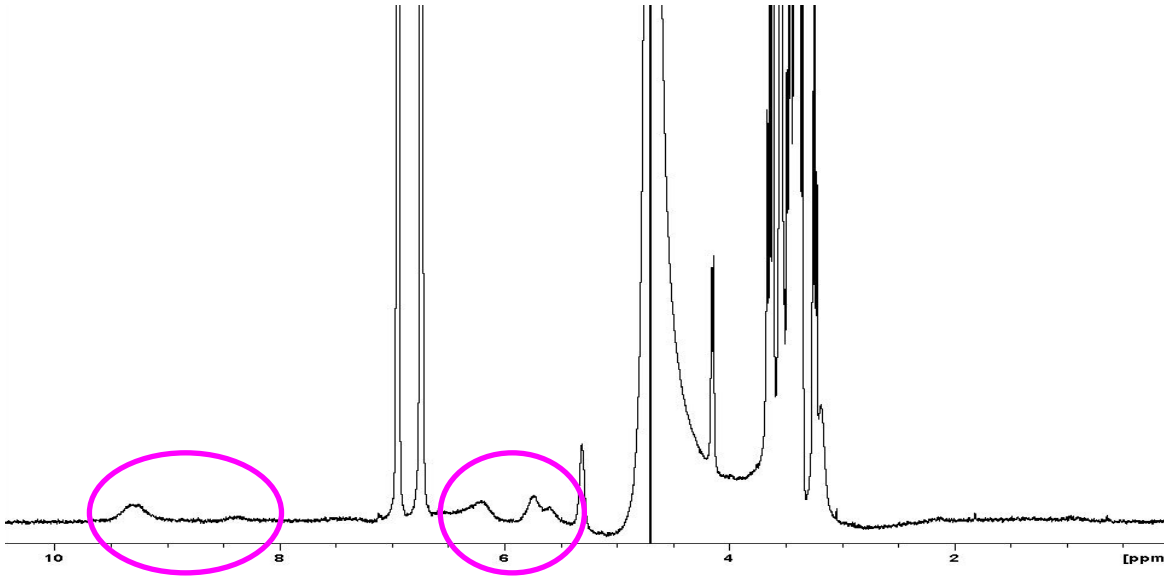


Figure 3.32 ^{1}H NMR spectrum of monomer **4** with water-gate (sequence p3919gp) in H_2O:D_2O 9:1 in phosphate buffer at pH=6 at 278K. In circles there are the NH (and probably also OH) signals.

The starting geometries of all the studied molecules **1**-**4**, generated with the MAESTRO suite of programmes, were minimized using the PRCG method and the OPLS_2005 force field. Molecules **1**, **2**, **4** were then submitted to molecular-dynamics and conformational-search simulations. The molecular-dynamics were run using the OPLS_2005 force field, at 300 K, with a total simulation time of 4 ns. The conformational searches were run using OPLS_2005 as force field, with the torsional sampling method (MCMM). Then, the thiourea dihedral angles for each molecule were analyzed. For the monomer **4** only EZ and ZZ conformers were found (figure **3.33**) and they present very similar potential energy values (figure **3.34**).

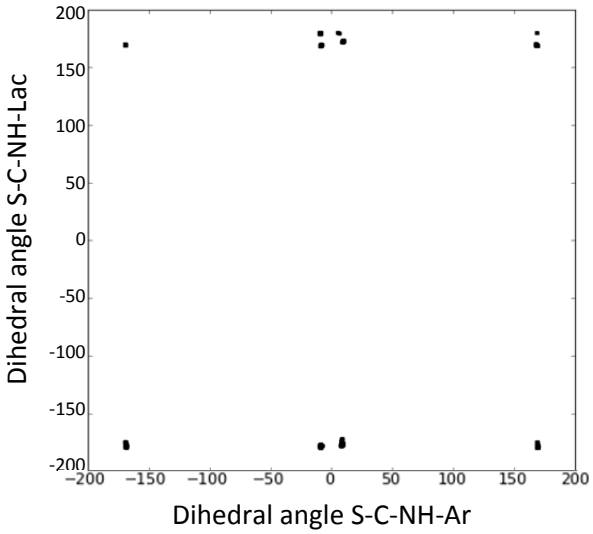


Figure 3.33 Schematic representation of the conformations found for thiourea bonds of monomer **4** after the conformational search. The *x*-axis represents the dihedral angles between N-aromatic ring and C=S, while the *y*-axis is for dihedral angles between N-Lac and C=S. When the value is at -180° or 180°, the angle is Z, while when it is close to 0° it is E. In this way it is possible to check that all the structures found are ZZ or EZ.

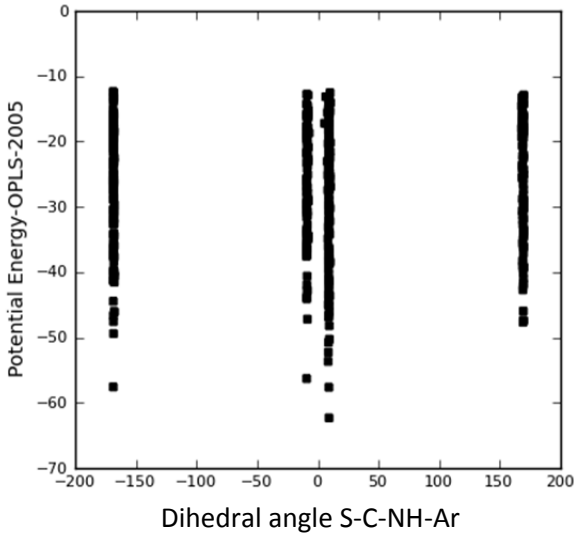


Figure 3.34 Potential energy for the different conformations found for monomer **4** after conformational-search analysis in function of the dihedral angle between N-aromatic ring and C=S. The other dihedral angle between N-Lac and C=S is not displayed here because it was found to be always in the Z conformation.

Subsequently, the obtained ZZ rotamer with the lowest potential energy value was used as template to build the other more complex calixarene compounds **1** and **2**.

For calixarene *alt*-4Lac[4]Prop **1**, after analyzing the eight dihedral angles together (two for each of the four thiourea present), it was found that almost half (49%) of the structures presented three ZZ and one EZ conformations for the thiourea moieties, one third (33%) of the structures presented a two ZZ and two EZ geometries, and another 13% presented all ZZ torsion angles (figure **3.35**). The potential energy graphic (figure **3.36**) indicated the structures with two or three thiourea moieties in the ZZ conformation to be the most stable geometries.

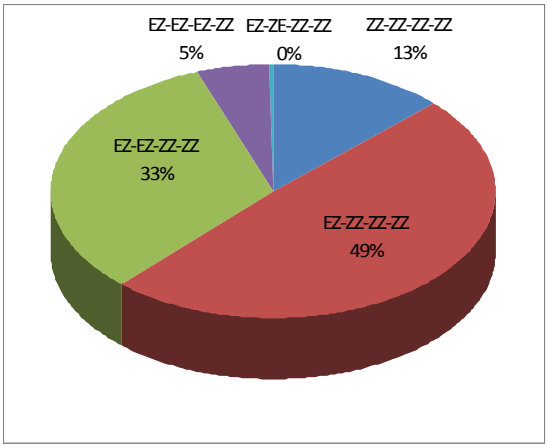


Figure 3.35 Composition of the conformations found for *alt*-4Lac[4]Prop **1** after conformational-search simulation.

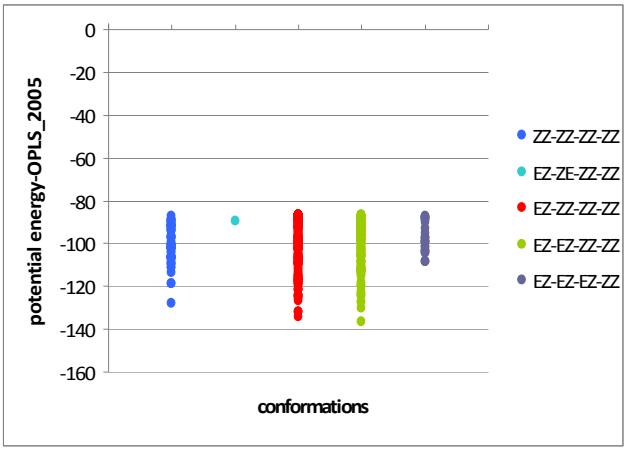


Figure 3.36 Potential energy values for the different conformations found for *alt*-4Lac[4]Prop **1** after conformational-search simulation.

Different results were found for *cone*-4Lac[4]Prop **2**, which showed almost all the structures (96%) with four EZ geometries for the thiourea moieties after running the conformational-search simulation (figures **3.37** and **3.38**).

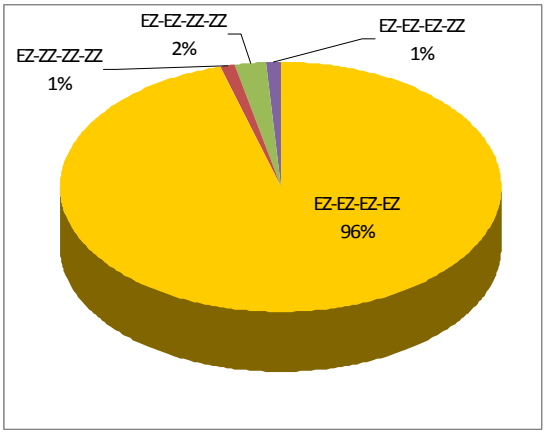


Figure 3.37 Composition of the conformations found for *cone*-4Lac[4]Prop **2** after conformational-search simulation.

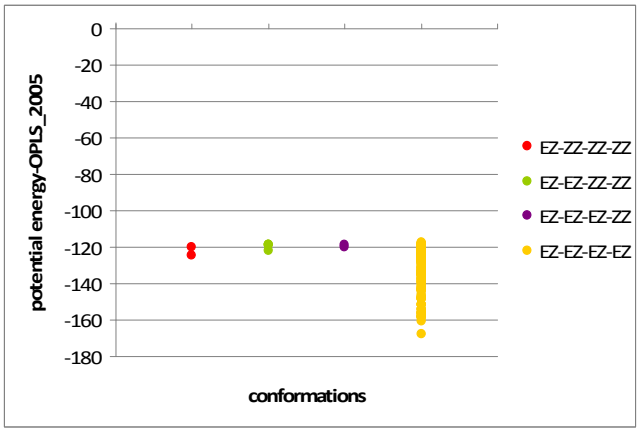


Figure 3.38 Potential energy values for the different conformations found for *cone*-4Lac[4]Prop **2** after conformational-search simulation.

This behaviour could be ascribed probably to steric reasons since compound **2** links all the four lactose moieties on the same side of the calixarene backbone. Inspection of the possible models showed that the resulting glycocluster geometries, when the four thiourea are all in

the EZ conformation, are much less constrained and displayed less destabilizing steric interactions than when ZZ geometries are present (figure **3.39**).

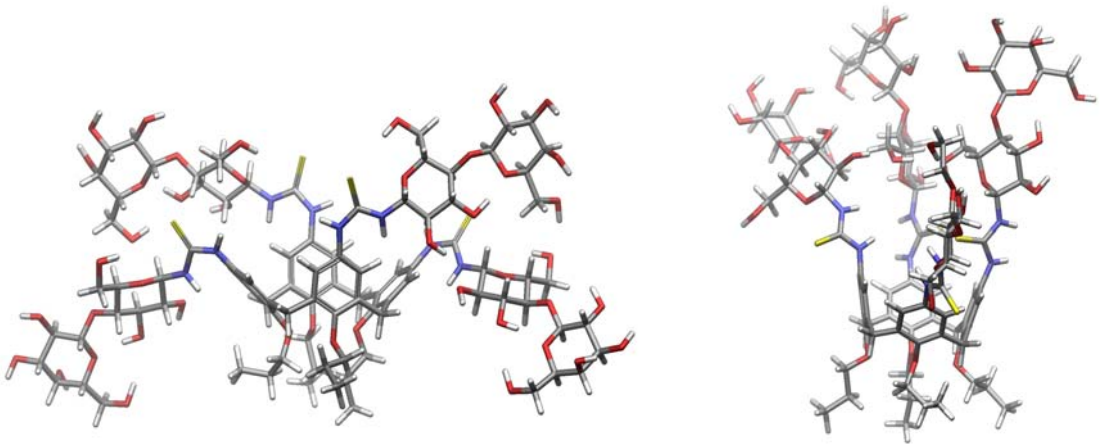


Figure 3.39 Compound **2** with all EZ thiourea (left) and with all ZZ thiourea (right).

It can be important to take into account, however, that in water solution the aggregation processes involving these clusters and verified by the NMR studies could force the *cone*-calixarenes in the all-ZZ geometry. In this way the amphiphilic character of the glycocalixarene would be enanched, making it more suitable for the participation to the formation of micelle-like aggregates.

3.2.4.2 Docking and dynamic simulation analyses

Finally, a docking protocol was adopted. Molecular modeling was performed for the two ligands **1** and **4** with Gal-3 to integrate and confirm the NMR data obtained. The starting geometry for docking was taken from the crystallographic structure of Gal-3 complexed with N-acetyllactosamine at 1.4 Å X-ray resolution (pdb: 1JKL). Initially, monomer **4** was studied. The ZZ lowest potential energy conformation of **4** was used as starting geometry of the ligand. Following the ligand preparation protocol of GLIDE within the MAESTRO suite, the docking process was performed. Upon docking, the analysis of the found structures showed the galactose residue of **4** to be in the Gal-3 binding pocket as expected, following the

lactose recognition pattern. Interestingly, upon binding a structural change in the thiourea linker of **4** towards the ZE conformation was observed. For the galactose moiety, four different H-bonds were detected. The endocyclic oxygen acts as H-bond acceptor of one NH of the side chain of the Arg 144 guanidinium group; O4 forms an H-bond with His 158; O6 accepts the Asn 174 amide proton and donates to the negatively charged carboxylate of Glu 184. Besides there are CH-π stacking interactions of the galactose H_3', H_4' and H_5' protons with the Trp 181 indole moiety. The glucose moiety also forms an array of hydrogen bonds: O6 forms two H-bonds with two NHs of Arg 144 guanidinium group, while H_6 is H-bonded to the Glu 184 carboxylate group (figure **3.40**). The same pattern of H-bonds was also found by Sörme et al.[16] for a LacNAc derivative bound to Gal-3. The NH-C(S)-NH moiety is not involved in H-bond, nor accepting nor donating, with the protein.

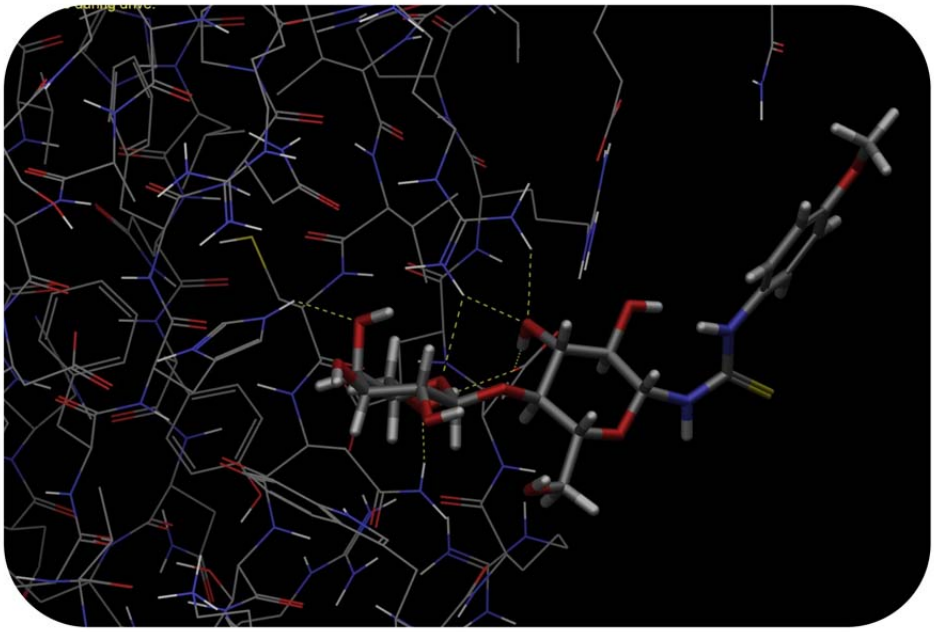


Figure 3.40 View of momovalent ligand **4** in the Gal-3 binding site and schematic representation of the H-bonds involved in the binding.

Docking of calixarene **1** resulted to be rather challenging, especially because the goal was to explore its interaction with more than one Gal-3 simultaneously. Attempts to dock compound **1** with just one Gal-3 did not lead to suitable results, since at the end of the docking process the calixarene scaffold (some aromatic rings and propyl chains) resulted

tightly connected to the protein surface, in disagreement with STD experimental data. Besides, the addition of a second Gal-3 moiety to check multivalency resulted impossible. Thus, a different strategy was adopted and molecular dynamics simulations were carried out. In the employed protocol, the docked structure of monomer **4** to Gal-3 was used as template to generate, using MAESTRO, a starting structure of *alt*-4Lac[4]Prop **1** bound to four Gal-3 modules. Then, molecular dynamics were carried out using MAESTRO (OPLS_2005 as force field) at 300 K, with a total simulation time of 1 ns. Just after the equilibration period, two of the four galectins disconnected from the ligand, while the resulting complex of **1** with two Gal-3, bound at the two opposite faces of the calixarene scaffold, remained stable for the complete simulation time (figure **3.41**).

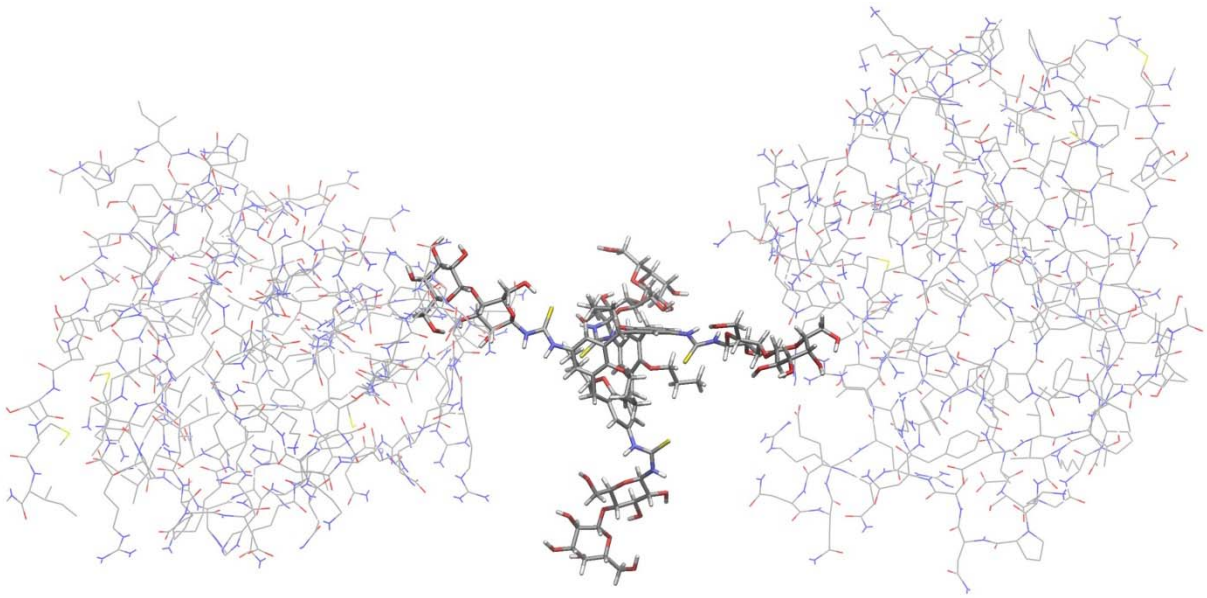


Figure 3.41 Calix[4]arene **1** binding to two Gal-3.

The final geometry of the 1:2 ligand:protein complex was modified and employed to run additional molecular dynamics (MD) analyses. The same simulation conditions were used, but different initial thiourea geometries for calixarene **1** were used as input: one with all-ZE thiourea, and the other one with two ZE and two ZZ geometries. In both cases the formed complex resulted stable for the whole simulation time, suggesting that the stability of the ligand:protein complex is not dependent on the particular initial E or Z thiourea conformation. For all MD runs, the energy plots were almost identical. From the molecular recognition perspective, the hydrogen bond pattern and the interactions formed between

the lactose moieties and the binding pocket of the protein was the same as that described above for the monomer **4**.

3.3 Conclusions

The binding of three glycocalixarenes (*alt*-4Lac[4]Prop **1**, *cone*-4Lac[4]Prop **2**, *cone*-4LacNAc[4]Prop **3**) and the monovalent compound **4** with three different lectins (VAA, Gal-3 and Gal-9N) was studied using STD- and DOSY-NMR techniques together with molecular modeling.

Through STD experiments it was possible to determine which parts of the molecules were responsible for binding to the lectins. The epitope map found for calixarene **1** and monomer **4** revealed that the binding occurs in different ways with the different proteins. Some differences found between the two ligands in the binding of the same lectin suggested that the aglycon part of the calixarene scaffold can affect the way of binding. In fact STD effects were observed also for the alkyl chains and the calixarene backbone of **1**, although of minor intensity compared to those of the sugar part.

In particular monomer **4** was found to have the same epitope of lactose, with H_2', H_3' and H_4' protons showing the greatest STD effect with VAA, and H_6', H_5' and H_4' protons interacting the most with Gal-3 and Gal-9N (the latter data collected at 0.81 ppm on-resonance frequency). *Alt*-4Lac[4]Prop calixarene **1** presented, with VAA and Gal-3, epitope maps for the lactoside protons similar to the ones found for compound **4**. Anyway the STD effects resulted to be quite intense for all the protons of its structure. The general conservation of the main epitope, consisting in the galactose unit, also for the calixarene-based ligand indicates that the saccharide maintains its propensity to the binding with these lectins despite the linkage to the rather hindered structure of the calixarene. On the basis of these results, the previously reported improved efficiency of these glycoclusters[2] in the lectin inhibition seems mainly due to the proper multivalent exposition of the epitopes by the calixarene scaffolds, rather than additional aspecific interactions of the macrocyclic moiety.

Standard ^{1}H NMR spectra for *cone*-calixarenes **2** and **3**, supported also by DOSY evidences, showed a strong tendency for these compounds to self-aggregate in water. Since their

resonances are completely overlapped, it was not possible to determine their epitope maps, even if it was qualitatively verified that they interact with the studied lectins.

DOSY experiments provided a method to unravel the presence of multivalent effects. Results showed that monomer **4**, as expected, can bind only one Gal-3. On the contrary, it was proved the ability of *alt*-4Lac[4]Prop calixarene **1** to interact with more than one Gal-3 at the same time. This result was also confirmed by molecular modeling simulations, which showed possible and stable the 1:2 glycocalix:protein complex. This suggests that in a multivalent environment simultaneous interactions can occur between the calixarene sugars and the protein recognition sites.

Molecular modeling was also employed to study the dihedral angles of the thiourea moieties for the various compounds, since no useful data could be gather from NMR experiments. In the free state monomer **4** was found to present either ZZ or EZ conformation, for 1,3-alternate calixarene **1** the structures with the lowest potential energies are either the EZ-ZZ-ZZ-ZZ or ZZ-ZZ-EZ-EZ geometries, while *cone*-calixarene **2** showed to be particularly stable with four EZ thiourea units. A docking protocol was finally adopted to study the interactions between compound **4** and Gal-3. Besides molecular dynamics simulations were employed to investigate the stability of **1** binding to more than one Gal-3, confirming the data obtained with DOSY experiments.

In conclusion STD- and DOSY-NMR experiments resulted to be a powerful tool to investigate the structure of the glycocalixarene-lectin complexes and the nature of their interaction. Besides through molecular modeling the experimental data were integrated and confirmed.

3.4 Experimental Part

Synthesis. For general information on instruments and chemicals used for synthesis, see Chapter 2.

[2,3,4,6-Tetra-*O*-acetyl-β-D-galactopyranosyl-(1→4)-2,3,6-tri-*O*-acetyl-β-D-glucopyranosy] -thioureido-4-methoxybenzene (6)

Lactosyl amine **5**[10] (0.84 g, 1.33 mmol) and isothiocyanato-4-methoxybenzene (0.20 g, 1.21 mmol) were dissolved in dry DMF (15 mL), then Et_3N (0.17 mL, 1.21 mmol) was added. The

mixture was stirred at 60 °C for 18 hours under nitrogen atmosphere. The progress of the reaction was monitored via TLC (eluent: hexane/AcOEt/CH_3OH 4:3:1). The reaction was quenched by adding water (20 mL) and extracting the product with AcOEt (5 x 20 mL). The combined organic phases were dried over anhydrous Na_2SO_4, filtered and the solvent removed *in vacuo*. The residue was purified via column chromatography (eluent: AcOEt/hexane 6:4) to give product **6** as a white solid. Yield: 62%. **^{1}H NMR** (300 MHz, $CDCl_3$): δ (ppm) 7.74 (s, 1H, ArNH); 7.07 (d, 2H, J = 8.8 Hz, Ar); 6.93 (d, 2H, J = 8.8 Hz, Ar); 6.35 (d, 1H, J = 8.6 Hz, ArNHCSNH); 5.73 (t, 1H, J = 9.0 Hz, H_1); 5.32 (m, 2H, H_3, H_4'); 5.09 (dd, 1H, $J_{1'-2'}$= 7.6 Hz, $J_{2'-3'}$ = 10.4 Hz, H_2'); 4.92 (dd, 1H, $J_{2'-3'}$ = 10.4 Hz, $J_{3'-4'}$ = 3.2 Hz, H_3'); 4.78 (t, 1H, J = 9.0 Hz, H_2); 4.45-4.49 (m, 2H, H_1', H_{6a}); 4.17-4.05 (m, 3H, H_{6b}, $H_6'_a$, $H_6'_b$); 3.82 (s, 3H, OCH_3); 3.86-3.71 (m, 3H, H_4, H_5, H_5'); 2.14, 2.10, 2.05, 2.04, 2.01, 1.95 (6s, 21H, CH_3CO). **^{13}C NMR** (75 MHz, $CDCl_3$): δ (ppm) 182.5 (CS); 170.9, 170.4, 170.3, 170.1, 170.0, 169.2, 168.9 (CO); 159.1, 127.6, 115.0 (Ar); 100.7 (C_1'); 82.9 (C_1); 75.9 (C_4); 74.4 (C_5); 72.2 (C_3); 70.9, 70.7, 70.6 (C_2, C_3', C_5'); 68.8 (C_2'); 66.6 (C_4'); 61.9 (C_6); 60.9 (C_6'); 54.4 (OCH_3); 20.8, 20.7, 20.6, 20.5, 20.4 ($COCH_3$). **ESI-MS(+)**: m/z 823.3 [100% $(M+Na)^+$]. **M.p.** : 146.0-147.0 °C.

[β-D-galactopyranosyl-(1→4)-β-D-glucopyranosyl]-thioureido-4-methoxybenzene (4)

Peracetylated lacto-derivative **6** (200 mg, 0.25 mmol) was dissolved in 10 mL of CH_3OH, and drops of a freshly prepared methanol solution of MeONa were added till pH 8-9. The mixture was stirred at room temperature for 1 hours. The progress of the reaction was monitored via TLC (eluent: CH_2Cl_2/CH_3OH/H_2O 6:3.5:0.5) and ESI-MS analysis. Since a precipitate was formed, drops of water were added to solubilize it. Amberlite resin IR 120/H^+ was subsequently added for quenching, and the mixture was gently stirred for 30 min till neutral pH. The resin was filtered off and the solvent was removed under reduced pressure to obtain product **4** as a white solid. Yield: 90%. **^{1}H NMR** (400 MHz, D_2O): δ (ppm) 7.15 (d, 2H, J = 7.6 Hz, Ar); 6.93 (d, 2H, J = 7.6 Hz, Ar); 5.47 (bs, 1H, H_1); 4.36 (d, 1H, J = 7.6 Hz, H_1'); 3.89-3.89 (m, 2H, H_{6a}, H_4'); 3.77-3.70 (m, 4H, H_{6b}, OCH_3); 3.69-3.66 (m, 2H, $H_6'_a$, $H_6'_b$); 3.65-3.52 (m, 5H, H_3', H_5', H_3, H_4, H_5), 3.48 (t, 1H, J = 8.8 Hz, H_2'); 3.41 (t, 1H, J = 8.0 Hz, H_2). **^{13}C NMR** (75 MHz, D_2O): δ (ppm) 185.8 (CS); 161.5, 131.5, 118.1 (Ar); 106.1 (C_1'); 87.3 (C_1); 81.0, 79.5, 78.6, 78.4, 75.8, 74.9, 74.2, 71.8 (C_2-C_5, C_2'-C_5'); 64.3, 63.1 (C_6, C_6'); 58.7 (OCH_3). **ESI-MS(+)**: m/z 529.4 [100% $(M+Na)^+$]. **M.p.** : 168.2-170.1 °C.

NMR spetroscopy. All NMR experiments were conducted at 298 K on a Bruker Advance DRX 500 MHz spectrometer. If not specified otherwise, all the samples were prepared in phosphate buffer, 100 mM NaCl, pD = 7.4, 100% D_2O. Three equivalents of DTT-d_{10} (dithiothreitol) for each cysteine present in the proteins were added to all prepared solutions to avoid oxidation processes.

STD-NMR experiments were performed for all compounds with protein-ligand molar ratios between 1:10 and 1:200. The off-resonance frequency was set at δ = 100 ppm, while the on-resonance frequency was centred at δ = -1.5 ppm. Only for selected experiments (as specified in the text) the on-resonance frequency was set at 0.81 ppm. A train of Gaussian pulses of 49 ms each spaced by 1 ms delays was chosen for saturation. Typical saturation times were 2 s. For DOSY experiments, all samples were analyzed with a ledbpg2s sequence pulse, with a linear gradient between 2% and 95%. Thirty-two 1D ^1H spectra were collected with a duration of δ = 2 ms and an echo delay of Δ = 200 ms for the ligands and of 230 ms for the proteins and the protein-ligand complexes. Spectral processing was performed on PC station using Topspin 2.1 (Bruker).

3.5 References

1 André, S.; Grandjean, C.; Gautier, F. M.; Bernardi, S.; Sansone, F.; Gabius, H. J.; Ungaro, R. *Chem. Commun.*, **2011**, *47*, 6126-6128.

2 Andrè, S.; Sansone, F.; Kaltner, H.; Casnati, A.; Kopitz, J.; Gabius, H. J.; Ungaro, R. *ChemBioChem*, **2008**, *9*, 1649-1661.

3 Solís, D.; Jiménez-Barbero, J.; Kaltner, H.; Romeo, A.; Siebert H. C.; von der Lieth, C. W. et al. *Cells Tissues Organs*, **2001**, *168*, 5-23.

4 Villalobo, A.; Nogales-Gonzáles, A.; Gabius, H. J. *Trends Glycosci. Glycotechnol.*, **2006**, *18*, 1-37.

5 Gabius H. J. *The Sugar Code. Fundamentals of Glycosciences*, Weinheim, Germany; Wiley-VCH, **2009**.

6 Meyer, B.; Peters, T. *Angew. Chem. Int. Ed.*, **2003**, *42* (8), 864-890.

7 Mayer, M.; Meyer, B. *Angew. Chem.*, **1999**, *111*, 1902-1906; *Angew. Chem. Int. Ed.*, **1999**, *38*, 1784-1788.

8 Peters, T.; Meyer, B. "*Methods for Detecting Biologically Active Compounds from Compound Libraries*", German Pat. No. 19649359, Swiss Pat. No. 690695, US Pat. No. 6214561, GB-Patent No. GB2321104 **1996**.

9 Stilbs, P. *Anal. Chem.*, **1981**, *53*, 2135-2137.

10 Sansone, F.; Chierici, E.; Casnati, A.; Ungaro, R. *Org. Biomol. Chem.,***2003**, *1*, 1802-1809.

11 Zemplén, G.; Pascu, E. *Ber. Dtsch. Chem. Ges.*, **1929**, *62*, 1613-1618.

12 Ribeiro, J. P.; Andrè, S.; Cañada J. F.; Gabius H. J.; Butera, A. P.; Alves, R. J.; Jiménez-Barbero, J. *ChemMedChem*, **2010**, *5*, 415-419.

13 Sorland, G. H.; Aksnes, D. *Magn. Reson. Chem.*, **2002**, *40*, S139-S146.

14 Groves, P.; Ohsten Rasmussen, M.; Molero M. S.; Samain, E.; Cañada, F. J.; Driguez, H.; Jiménez-Barbero, J. *Glycobiology*, **2004**, *14* (5), 451-456.

15 (*a*) García Fernández, J. M.; Jiménez Blanco, J. L.; Ortiz Mellet, C.; Fuentes, J. *J. Chem. Soc., Chem. Commun.*, **1995**, 57-58. (*b*) Jiménez Blanco, J. L.; Benito, J. M.; Ortiz Mellet, C.; García Fernández, J. M. *Org. Lett.*, **1999**, *1*, 1217-1220. (*c*) Díaz Pérez, V. M.; Ortiz Mellet, C.; Fuentes, J.; García Fernández, J. M.; *Carbohydr. Res.*, **2000**, 326, 161–175. (*d*) García Fernández J. M.; Ortiz Mellet, C. *Adv. Carbohydr. Chem. Biochem.*, **2000**, 55, 35-135.

16 Sörme et al *J. Am. Chem. Soc.* **2005***, 127*(6), 1737-1743.

Chapter 4

New glycocalix[4]arenes via Cu-catalyzed azide-alkyne cycloaddition

Abstract

The Cu-catalyzed azide-alkyne cycloaddition (CuAAC) is one of the most used reactions of the "click" chemistry to connect carbohydrates to a multivalent core. In this chapter it is presented the synthesis of calix[4]arene-based glycoclusters where the saccharide units are conjugated to the macrocyclic scaffold via the CuAAC reaction. Initially two galactosyl-calix[4]arnes were prepared, differing in the relative dispositions of the alkyne and azido groups on the moieties undergoing the CuAAC, to explore the best methodology to obtain these compounds. Once found the most convenient pathway, also two lactosyl-calix[4]arenes, one in the *cone* the other one in the 1,3-alternate structure, were prepared. All these compounds present in their structures ethylene glycol linkers between the calixarene core and the sugar units. The study of the interactions of these novel glycocalixarenes with galectins is currently under investigation and the results will hopefully allow also to disclose the possible roles of the long and mobile linkers in the inhibition processes.

4.1 Introduction

Efficient carbohydrate conjugation to scaffold molecules is an important goal in bio-organic chemistry. Till now we have focused our attention on calixarenes presenting carbohydrate moieties covalently linked to the macrocycle via a thiourea linker[1,2], that originates from an highly efficient amine-isothiocyanate reaction and allows the potential formation of hydrogen bonding[3,4] during the recognition processes. In general there are many different strategies for the conjugation of sugar units to a multivalent core. During the past years, in this research group, also classical O-glycosylation reactions[5,6], Mitsunobu reaction[5,6] and amide bond formation[7] were exploited. Reductive amination[8], thiol photoaddition onto allylic systems[9] and other methods are also commonly used. A method which can be currently considered one of the most classical and widely used strategies, is the 1,3-dipolar cycloaddition reaction between an alkyne and an azide, known as Huisgen reactions[10], and it represents one the most important examples of the so called "click" chemistry[11]. In particular the Copper(I)-catalyzed azide-alkyne 1,3-dipolar cycloaddition (CuAAC)[12], that

leads to the selective formation of 1,4 disubstituted triazole rings, has been established as one of the most reliable procedure for the covalent assembly of complex molecules and specifically of multivalent and multiglycosylated molecules[13,14,15,16]. In addition to the synthetic advantages offered in the conjugation processes by this reaction, the presence of triazole moieties in the assembled molecules confers some attractive properties. First of all they show high chemical stability and inertness to severe hydrolytic, oxidizing, reducing conditions, even at high temperature, and they are also relatively stable to metabolic degradation[17]. Besides, they have a strong dipole moment, aromatic character and are capable of forming hydrogen bonds[18,19] which can be favorable in the binding of biomolecular targets and for solubility in water solution[20]. The availability of very efficient and selective reactions is particularly crucial when the synthesis of multivalent ligands has to be carried out, to allow a fast, complete, stereo- and regioselective functionalization of the scaffold[21]. For these reasons we have focused our attention on the synthesis of new calix[4]arene based glycoclusters that could interact with target galectins thanks to the presence of four saccharide units connected to the scaffold through ethylene glycol linkers of proper length and triazole units generated by CuAAC reactions.

4.2 Results and discussion

There are two possibilities for the use of the Cu(I)-catalyzed Huisgen cycloaddition for the synthesis of glycocalixarenes. An alkyne functionalized calixarene scaffold can be reacted with a sugar bearing an azido group (method A in figure **4.1**) or, alternatively, a macrocycle functionalized with azido units is “clicked” to alkyne containing carbohydrates (method B in figure **4.1**). Both methodologies were explored in this work and they are illustrated and discussed in this chapter.

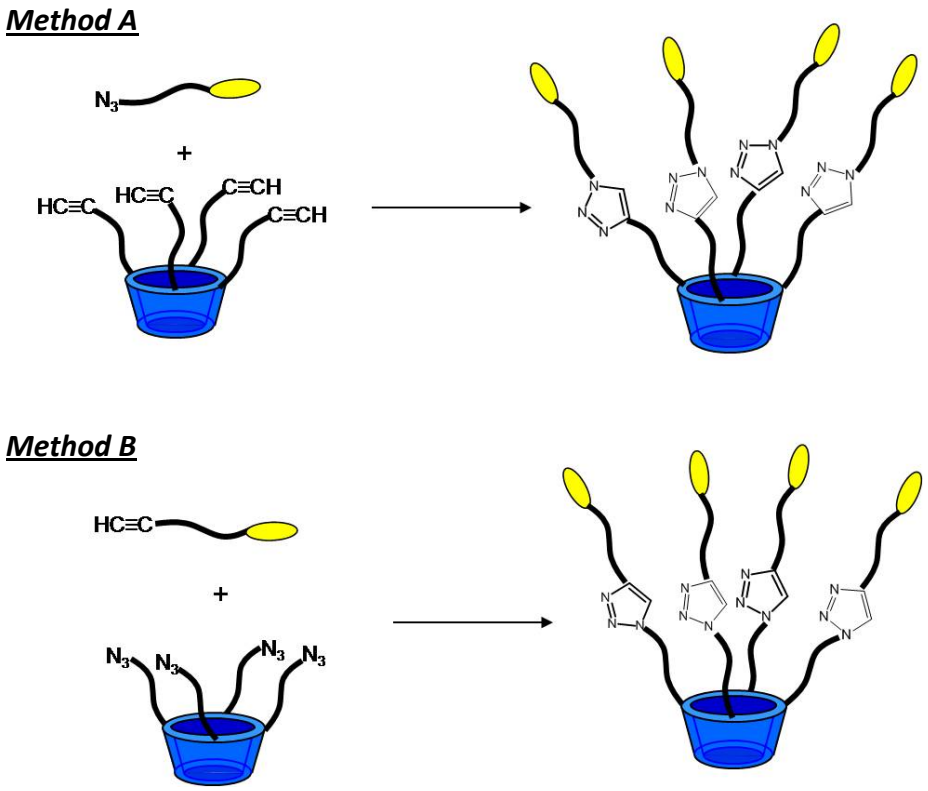


Figure 4.1 The two possible pathways to connect carbohydrate units to a calixarene core via the CuAAC reaction.

The first two target glycoclusters synthesized are compounds **1** and **2** (figure **4.2**), obtained by following method A and method B, respectively. The polyethylene glycol spacer, present in both molecules between the glycoside units and the triazole ring, was chosen also to ensure a good hydrophilicity to the clusters and then possibly to increase their water solubility. Moreover the presence of different O atoms along the linkers could allow the formation of additional hydrogen bonds with the protein surface, potentially enhancing the efficiency of the recognition processes.

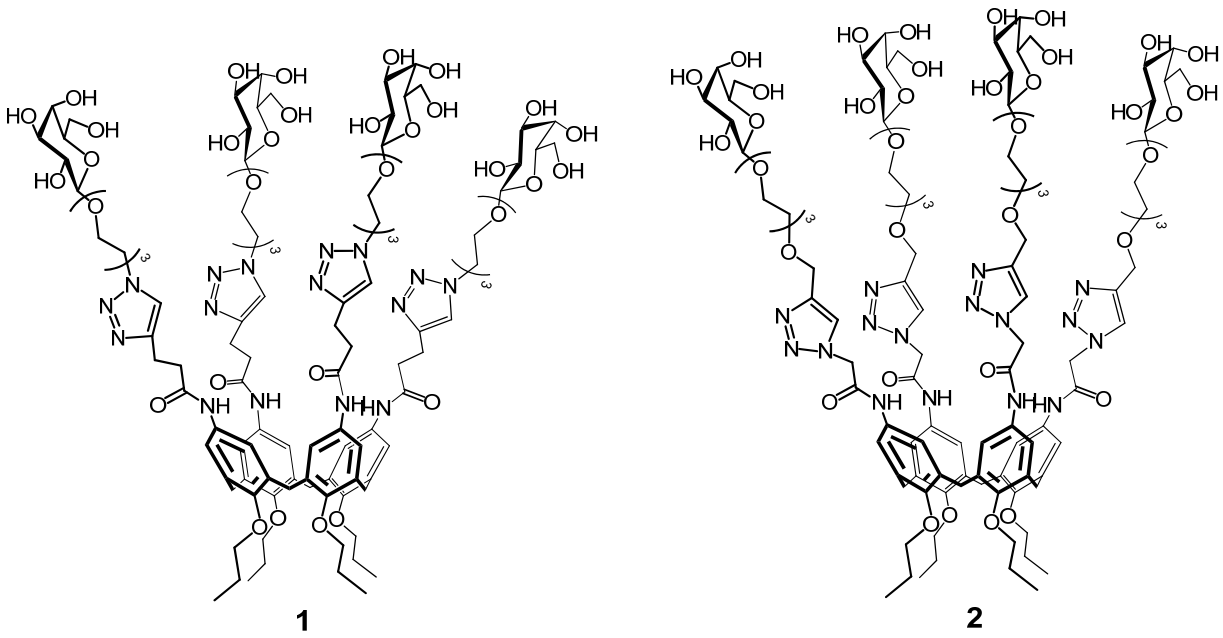


Figure 4.2 The two target glycoclusters **1** and **2** synthesized with method A and B respectively.

The choice of synthesizing calixarenes blocked in the cone conformation and functionalized with galactose units, constituted only the starting point towards the preparation of lactoclusters and, in perspective to build a wider library of ligands for galectin targeting and inhibition. Once selected the best strategy and conditions, according to the number of steps, yields, simplicity of the synthetic and purification procedures, other analogous compounds, differing in the conformation and type of linked carbohydrates, could be prepared.

4.2.1 Synthesis of triazole-containing galactosylcalixarenes

The first route explored was method A applied to the preparation of the multivalent compound **1.** This glycocluster could be synthesized by exploiting a convergent synthetic approach, where the azido-terminating glycoside **6** was connected via the Copper-catalyzed Azido-Alkyne Cycloaddition (CuAAC) reaction to calix[4]arene **10**, decorated at the upper rim with alkyne terminating chains (*vide infra* figure **4.9**).

Initially, the azido glycoside **6** was synthesized (figure **4.3**). Since glycosylation reactions usually show a moderate stereoselectivity, the reaction of the peracetylated carbohydrate with the glycol chain was performed in one of the first steps of the whole synthetic

pathway. Once the β-anomer is isolated pure from the reaction, the sugar moiety can be linked to the calixarene scaffold via the "click" CuAAC reaction without affecting anymore the stereochemistry of the anomeric carbon. The starting step was therefore the synthesis of 1-β-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)glycoside **6** (figure **4.3**).

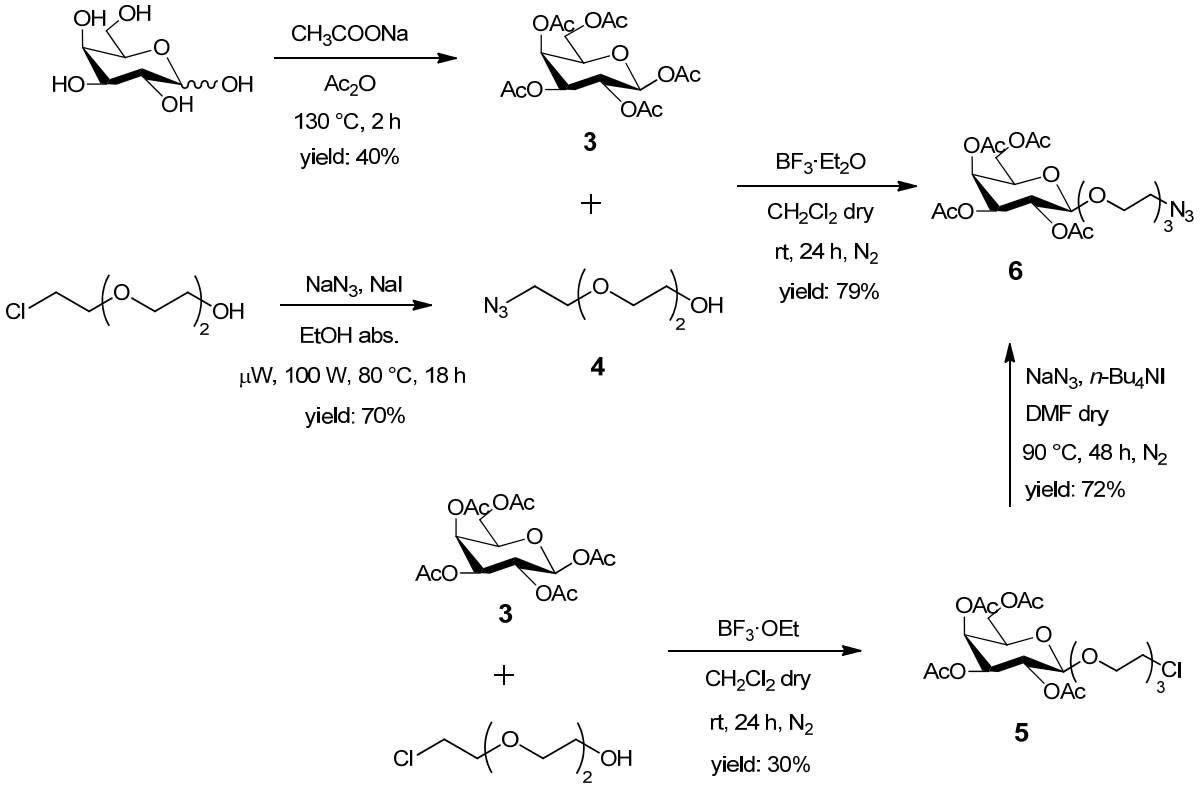


Figure 4.3 The two alternative pathways to obtain azido galactoside **6**.

According to a literature procedure[22], the chlorine atom on 2-(2-(2-chloroethoxy)ethoxy) ethanol was replaced with an azido group through the reaction with NaN_3 to give compound **4**. 2-(2-(2-chloroethoxy)ethoxy)ethanol is commercially available and constitutes an important building block since it allows to start the synthesis from a triethylene glycol already desymmetrized. Microwave irradiation was used to speed up the substitution reaction. The azido derivative **4** was in this way obtained in 70% yield and in only 18 hours, instead of the five days reported[22]. By treatment of peracetylated-galactose **3** with azido derivative **4** in presence of $BF_3{\cdot}Et_2O$ as promoter, a mixture of β- and α-anomers of compound **6** was formed. Attempts to separate the two anomers via flash chromatography

and preparative TLC on plates failed and therefore we decided to follow a different strategy[23] (figure **4.3**). The galactoside **3** was directly reacted with 2-(2-(2-chloroethoxy) ethoxy)ethanol again using $BF_3{\cdot}Et_2O$ as Lewis acid to promote the glycosylation reaction. To minimize the formation of the α-anomer the temperature was initially kept at 0 °C and then slowly allowed to reach room temperature. Nevertheless, a mixture of the two α- and β-anomers (1.5 : 8.5 ratio) was obtained also in this case. However, separation of the two anomers of compound **5** via column chromatography yielded, without particular difficulties, the pure β-compound **5**. The ^1H NMR spectrum proved the success of the glycosylation reaction and of the purification step. Diagnostic is the coupling constant of 8.0 Hz for the signal of H_1 proton at 4.55 ppm, that unambiguously confirmed the presence of the product with the desired β-stereochemistry. Once obtained, compound **5** was reacted with NaN_3 to give the corresponding azido derivative **6** in satisfactory yield (72%).

To connect the carbohydrate moieties to the calixarene scaffold via the "click" reaction, it was necessary to functionalize the macrocycle upper rim with alkyne containing moieties (figure **4.4**).

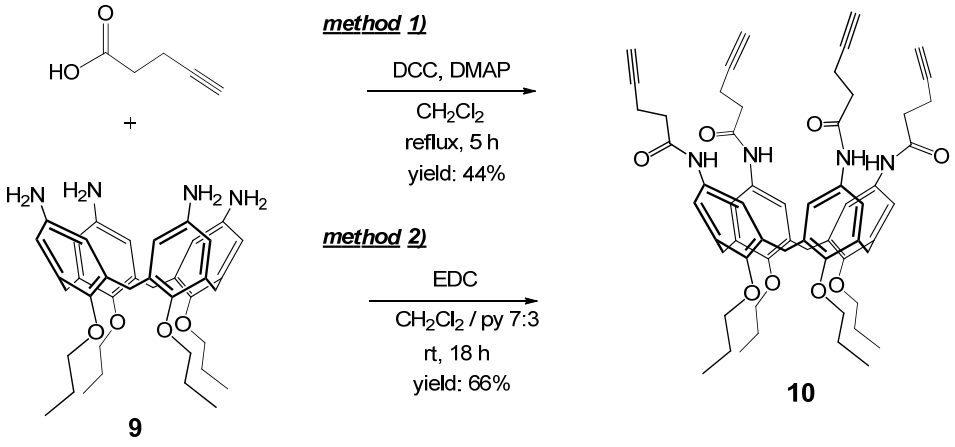


Figure 4.4 Synthesis of tetra-alkyne calix[4]arene **10**.

Amino-calix[4]arene **9** was synthesized according to literature procedures[24] as already described in chapter 2. Coupling reaction between amino-calix[4]arene **9** and 4-pentynoic acid in presence of dicyclohexylcarbodiimide (DCC) led to compound **10** in 44% yield. Due to the concurrent formation of dicyclohexylurea (DCU), several purification steps were necessary to obtain pure calix[4]arene **10**. The use of 1-ethyl-3-(3-dimethylaminopropyl)

carbodiimide (EDC) as alternative coupling agent allowed on the contrary to obtain pure target compound **10** in a more straightforward way (only one column chromatography) and also in higher yield (66%). Calixarene **10** resulted to be very soluble in methanol, while it showed almost no solubility in CH_2Cl_2.

Attempts to connect the alkyne functionality in closer proximity to the calixarene core, by decreasing the number of carbon atoms between the carboxylic group and the triple bond, did not give fruitful results. Reactions between amino-calix[4]arene **9** and propiolic acid (figure **4.5**) were carried out with different coupling agents. In presence of DCC it was possible to obtain product **11**, as confirmed by ESI-MS analyses. Anyway due to the high amount of non-identified byproducts formed during the reaction, it was not possible to isolate it pure from the crude mixture. The use of HBTU as coupling reagent could not lead to the tetra-functionalized product **11** even after a three days stirring of the reaction mixture.

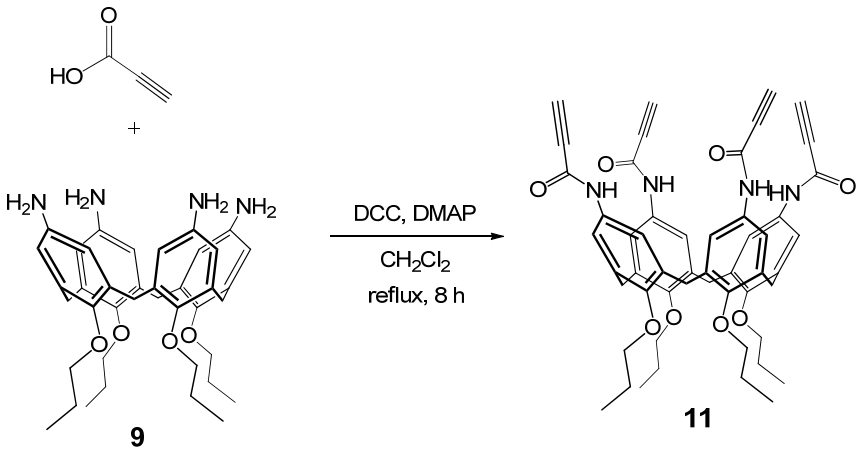


Figure 4.5 Synthesis of tetra-alkyne calix[4]arene **11**.

The second strategy explored was aimed at obtaining the glycocluster **2** through the complementary method B (figures **4.1** and **4.2**). As said above, in this case the alkyne functionality lays on the carbohydrate unit **8**, while the calixarene core **13** is tetra-functionalized with azido terminating moieties. Also in this pathway the two reagents (**8** and **13**) obtained from a convergent approach, were connected together via a CuAAC reaction (*vide infra*, figure **4.10**).

The synthesis of the galactoside moiety bearing an alkyne group was achieved as shown in figure **4.6**.

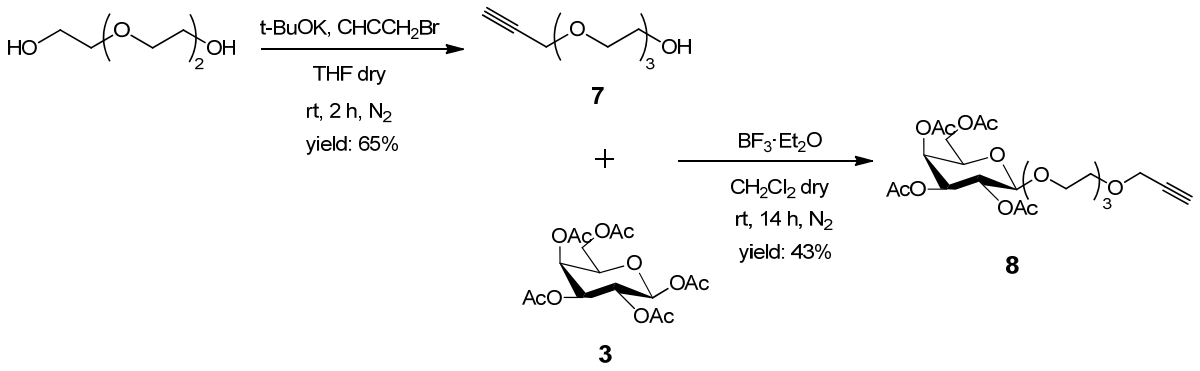


Figure 4.6 Synthetic pathway to obtain alkyne galactoside **8**.

Triethylene glycol was alkylated with propargyl bromide in presence of t-BuOK in dry THF to give compound **7**[25]. Gycosylation of peracetylated-β-galactose **3** with **7** in presence of $BF_3{\cdot}Et_2O$ as catalyst gave derivative **8**[25]. Also in this case, as for the analogous glycosylation reaction leading to compound **5**, it was necessary to initially perform the reaction at low temperature to favor the selective formation of the β-anomer. Nevertheless, also the α-amomer was obtained. Purification via flash chromatography allowed anyway to get pure β-galactoside **8**, as confirmed by NMR data.

Azido-calix[4]arene **13**, necessary as counterpart of the alkyne carbohydrate **8** in the CuAAC reaction, was synthesized in two steps starting from amino derivative **9** (figure **4.7**).

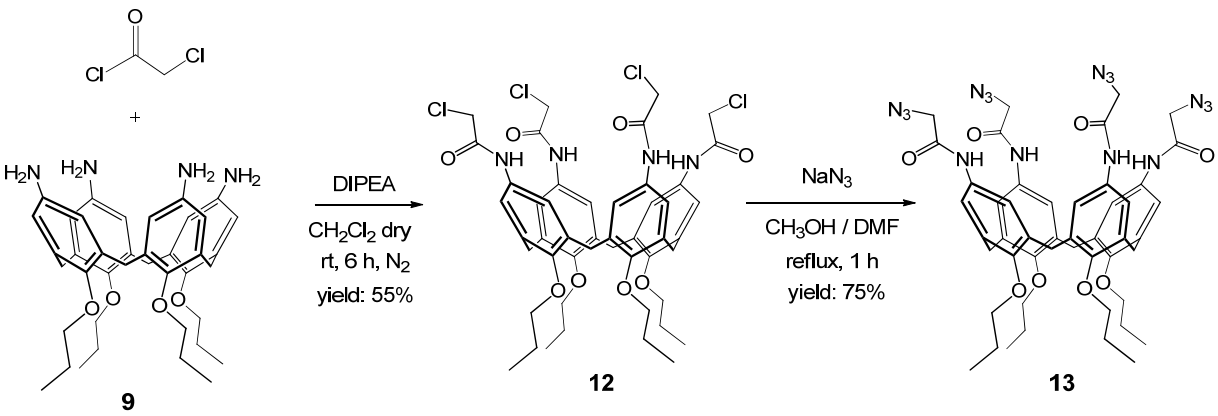


Figure 4.7 Synthesis of tetra-azido calix[4]arene **13**.

Amino-calix[4]arene **9** was reacted with chloroacetyl chloride to give compound **12** which presents an amide group with a chlorine atom in α-position. Reaction with NaN_3 led to the formation of calixarene derivative **13** in 1 hour, as confirmed by ESI-MS profiles, and in high yield (75%).

Once obtained the calixarene cores and the carbohydrate derivatives properly functionalized with azido and alkyne functions, the next step was their conjugation via CuAAC reaction with formation of a 1,2,3-triazole. The uncatalyzed thermically induced reaction between azides and acetylenes was systematically described by Huisgen in the early 1960s[26], therefore it is also called "Huisgen reaction". This reaction often leads to a mixture of 1,4- and 1,5-regioisomers, in general difficult to be separated and therefore not so convenient, especially when multivalent scaffolds are involved. The discovery that catalytic use of Cu(I) increases the reaction rate and controls the regioselectivity to give exclusively 1,4-di-substituted triazoles was made independently in recent years by Sharpless[12] and Meldal[27] (figure **4.8**).

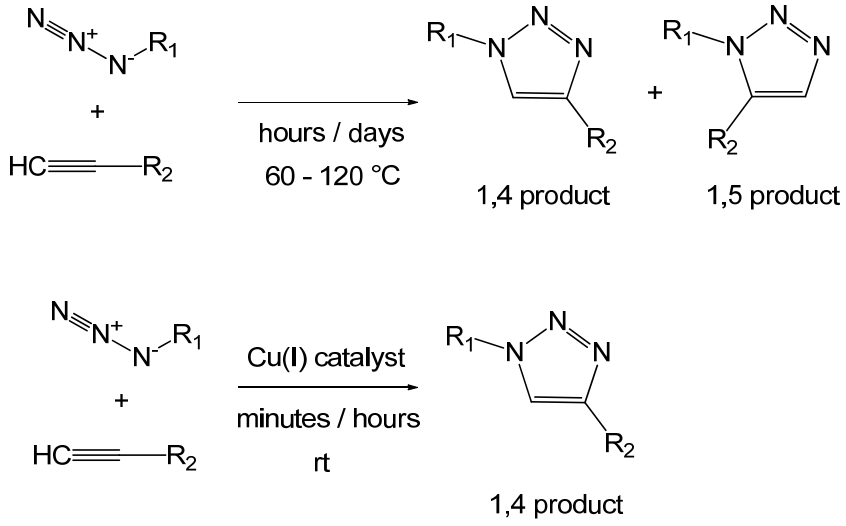


Figure 4.8 Top: Huisgen 1,3-dipolar cycloaddition reaction. Bottom: the CuAAC reaction.

For the copper catalysis a number of Cu(I) sources can be used, like CuI and CuBr, or coordination complexes like $[Cu(CH_3CN)_4]PF_6$ [12], $(EtO)_3P{\cdot}CuI$ [28], and $[Cu(PPh_3)_3]Br$ [29,30]. However Cu(I) is thermodynamically unstable and can be easily oxidized to inactive Cu(II),

therefore it is preferred to generate it *in situ* by reduction of Cu(II) salts by sodium ascorbate for example, that proved to be excellent. The mechanism of the CuAAC (copper-catalysed azido-alkyne cycloaddition) reaction does not retain the original concerted nature of thermically induced reactions, but proceeds through a multistep process involving the formation of alkyne-copper and azide-copper complexes[31,32] that leads to the final formation of the aromatic heterocyclic 1,4-disubstituted triazole ring.
The CuAAC reaction has become one of the most widely used reactions in organic chemistry, molecular biology, material and surface sciences because it readily transforms the reagents into a single product in high yield, at room temperature and with the maximum level of atom economy, even on multivalent molecules.

The CuAAC reactions between the tetra-alkyne calix[4]arene **10** and azido-galactoside **6** to give glycocluster **14** (figure **4.9**) and between tetra-azido calix[4]arene **13** and alkyne-galactoside **8** to afford glycocluster **15** (figure **4.10**) were performed in a rather similar way. In both cases the reaction was performed in DMF and H_2O. $CuSO_4$ was used as copper source and sodium ascorbate was added as reducing agent. The use of microwave-assisted procedures[33] allowed the rapid preparation of glycoclusters **14** and **15**, that were obtained in 20 minutes at 150 W and 80 °C in about 80% yield. No partially functionalized compounds or other byproducts were detected in the crude mixtures. Purification by column chromatography was necessary just to remove the excess of starting galactoside reagents used in the reactions.

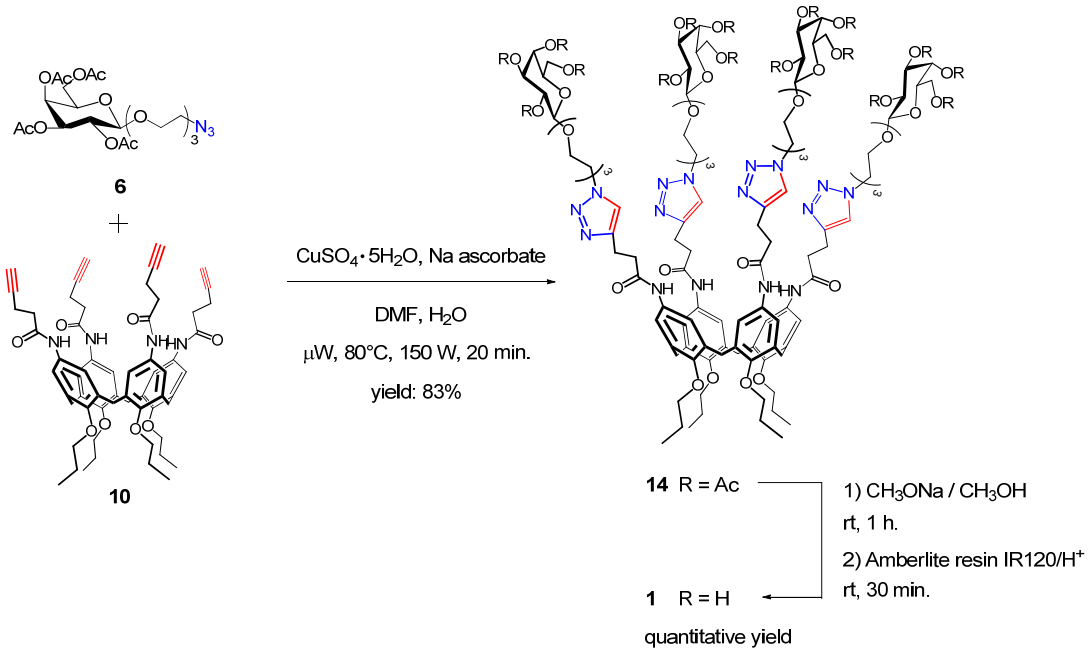


Figure 4.9 Synthesis of tetravalent calix[4]arenes **14** via CuAAC reaction and subsequent deprotection to give compound **1**.

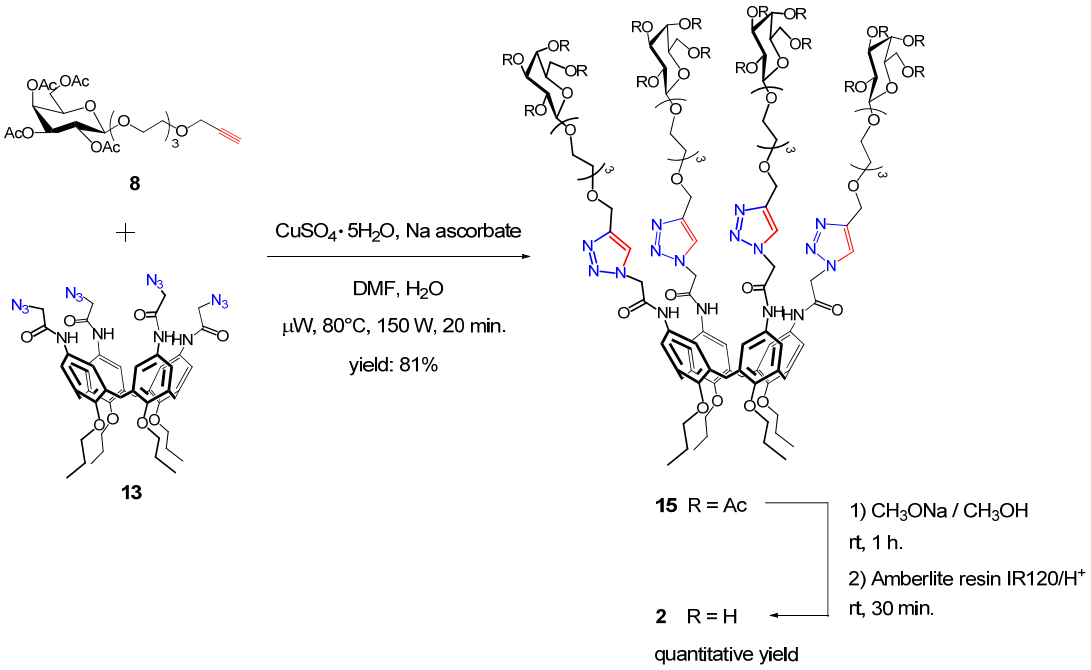


Figure 4.10 Synthesis of tetravalent calix[4]arenes **15** via CuAAC reaction and subsequent deprotection to give compound **2**.

All compounds were fully characterized via NMR (mono- and two-dimensional techniques) and ESI-MS analyses. The ^{1}H NMR spectrum of galactosyl-calix[4]arene **14** (figure **4.11**) presents a broad signal at 7.75 ppm for the proton of the triazole ring and complete absence at 2.27 ppm of the peak for the alkyne group present at the upper rim of the reagent **10**.

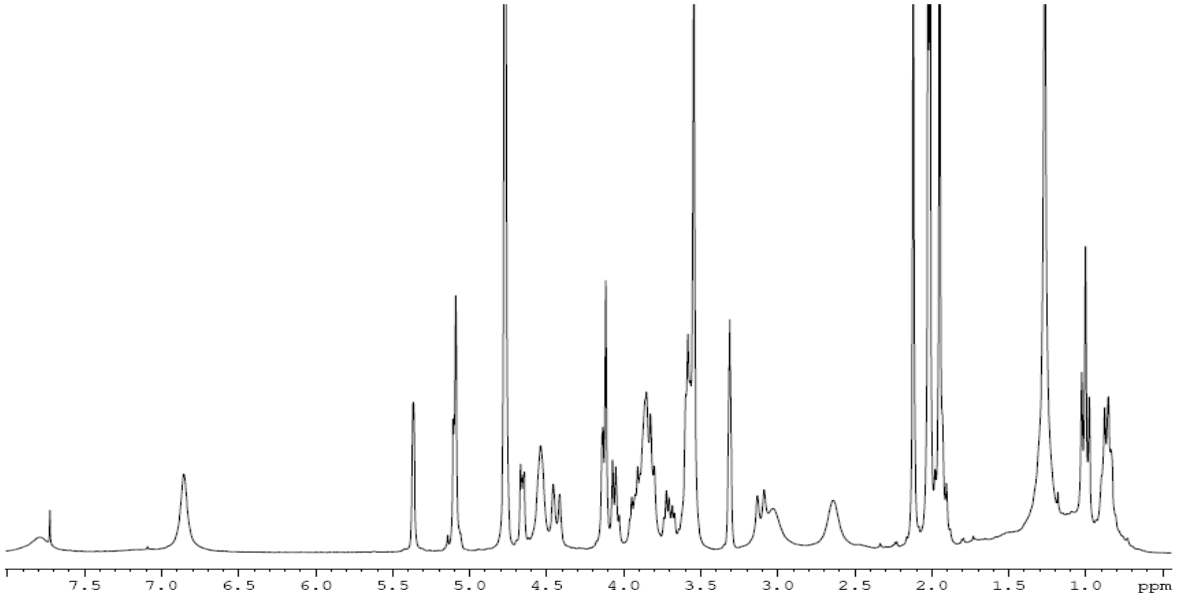


Figure 4.11 ^{1}H NMR spectrum (400 MHz, $CD_3OD/CDCl_3$ 4:1) of the acetylated glycocluster compound **14**.

Moreover, it is possible to recognize at 4.11 and 3.09 ppm two doublets for the axial and equatorial proton of the methylene bridge of the calixarene, respectively, typical for the symmetrical *cone*-structure. The complete tetra-functionalization of compound **14** was confirmed also by ESI-MS analyses which showed peaks at m/z 1520.5 and 1021.1 for the $(M+2Na)^{2+}$ and $(M+3Na)^{3+}$ adducts, respectively.

In figure **4.12** is reported the ^{1}H NMR spectrum of galactosyl-calix[4]arene **15**. Also in these case the peak at 7.89 ppm for the triazole proton and the two signals at 4.46 and 3.12 ppm for the methylene protons of the calixarene bridge constitute a proof of the formation of the triazole moiety and of the symmetrical geometry of the molecule. We could therefore conclude that it was completely tetra-functionalized. ESI-MS peaks at m/z 1042.8 and at m/z 1552.5 for the $(M+3Na)^{3+}$ and the $(M+2Na)^{2+}$ adducts, respectively, gave a further confirmation of the success of the "click" reaction.

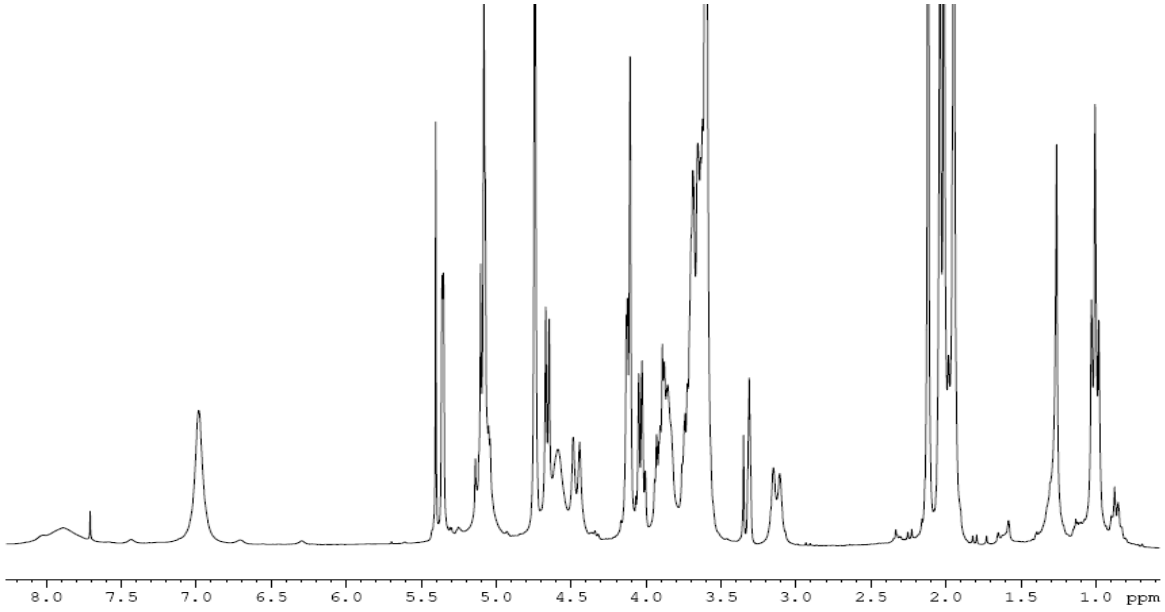


Figure 4.12 ^{1}H NMR spectrum (300 MHz, $CD_3OD/CDCl_3$ = 3:1) of the acetylated glycocluster compound **15**.

The deprotection from the acetyl groups was made via transesterification reaction in presence of CH_3ONa in CH_3OH at room temperature according to the standard Zemplén procedure[34] (figures **4.9** and **4.10**). Complete deacetylation was achieved in 1 hour, as confirmed by ^{1}H NMR spectra of the obtained compounds **1** and **2** (figure **4.13** for **1** and **4.14** for **2**).

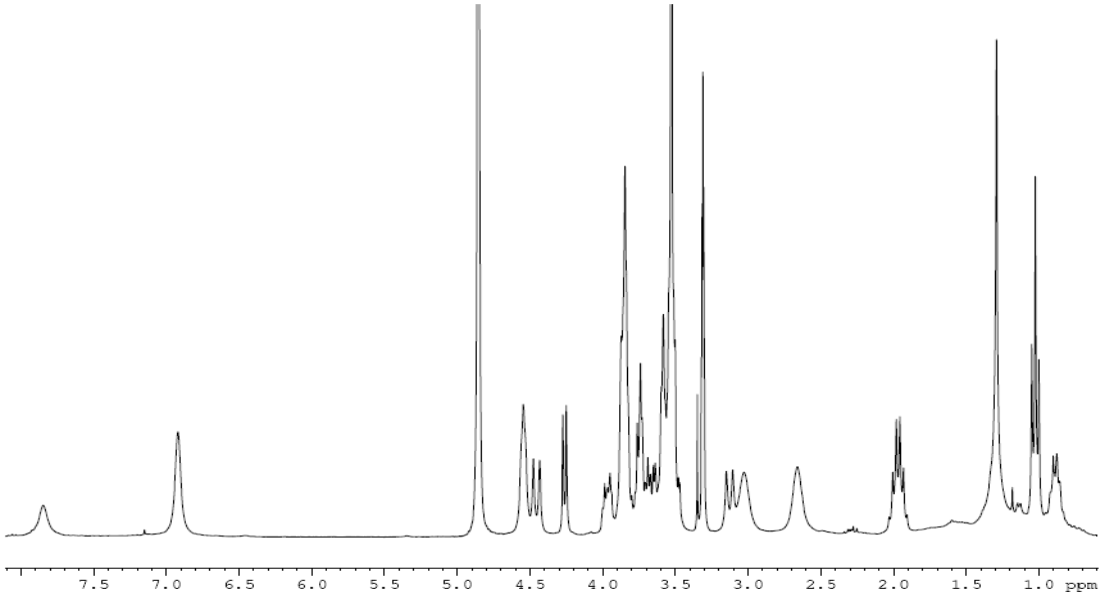


Figure 4.13 ^{1}H NMR spectrum (300 MHz, CD_3OD) of target compound **1**.

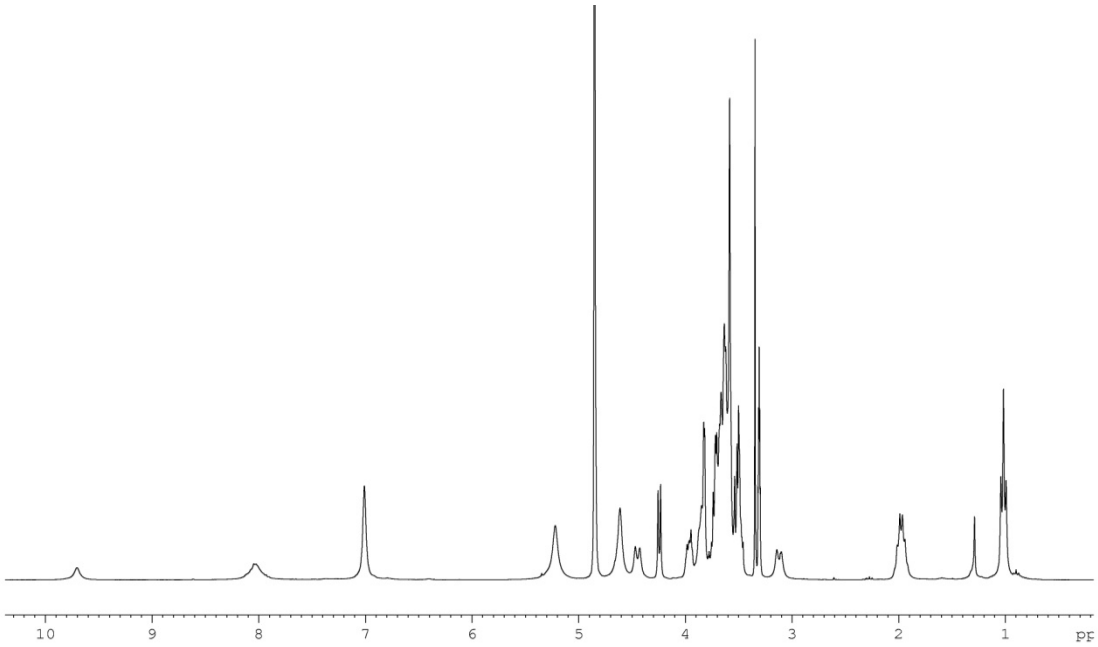


Figure 4.14 ^{1}H NMR spectrum (300 MHz, CD_3OD) of target compound **2**.

In both cases it is possible to observe the complete absence of the acetyl signals between 2.15 and 1.90 ppm in the ^{1}H NMR spectra, and of the signals at 171-169 ppm and 21-19 ppm respectively for the carbonyl and methyl groups of the acetyl functionalities in the ^{13}C NMR spectra. Confirmation of the complete deprotection came also from ESI-MS characterization. It is interesting to notice that while compound **1** showed high stability under Zemplén conditions also when the reaction was continued overnight, compound **2** instead started to decompose after 18 hours. ESI-MS profiles showed the presence of compounds cleaved at the amide position with loss of the entire glycosylated chain and formation of an amine group at the calixarene upper rim.

On the basis of all the data and information in our hands relative to the synthetic availability of intermediates and final products and to their stability under the reaction conditions, especially at the last deprotection step, we decided to synthesize triazole-containing lactosylcalixarenes according to method A. Indeed, the differences between method A and B on the galactoside clusters **1** and **2** are not so important, particularly in terms of yields and purification procedures. The choice for method A was mainly based on the numbers of synthetic steps (the alkyne calixarene **10** being obtained in just one step instead of the two

necessary to synthesize azido calixarene **13**) and for the higher stability under Zemplén conditions of the final product **1** compared to **2**. Therefore, the new glycoclusters were designed and planned to be synthesized exploiting the CuAAC reactions involving azido moieties on the carbohydrate units and alkyne groups on the calixarene scaffold.

4.2.2 Synthesis of triazole-containing lactosylcalixarenes

Once established the most convenient way to obtain this class of compounds, two new glycoclusters were designed and synthesized. *Cone*-compound **16**, similar to compound **1** but functionalized with lactose units, and compound **17**, that presents lactose moieties connected to a 1,3-alternate calixarene core (figure **4.15**) were the new target molecules.

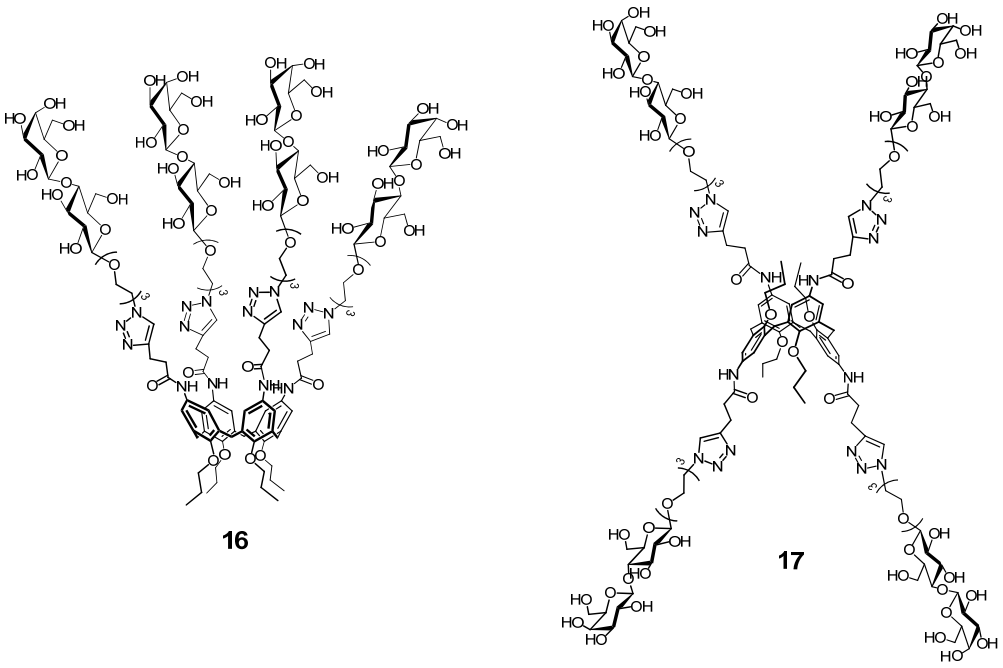


Figure 4.15 The two lactosyl-clusters: *cone*-lactosyl-calix[4]arene **16** and *alt*-lactosyl-calix[4]arene **17**.

The first step of their synthesis consisted in the preparation of lactoside derivative **20** (figure **4.16**). Attempts to synthesize it according to the procedure already used for the corresponding galactoside derivative **6** could not lead to satisfactory results. When

peracetylated-lactose **18** was reacted with 2-(2-(2-chloroethoxy)ethoxy)ethanol in presence of $BF_3{\cdot}Et_2O$ to give compound **19**, a mixture of α and β anomers (α/β ratio 2:3) was obtained. Differently from galacto-compound **6**, the anomers mixture in this case resulted very difficult to be separated with standard chromatographic methodologies. Therefore it was necessary to explore different methods to obtain compound **19** as pure β-anomer. Xue and coworkers[35] in a recent article reported a series of glycosylation reactions of sugar peracetates with various alcohols exploiting $SnCl_4$ and CF_3CO_2Ag as promoters. In particular, under the studied conditions, they obtained β-glycosides as major products and in high yields (between 45 and 80%) when sufficiently bulky alcohols (such as trifluoroethanol, propargylic alcohol, chlorotriethylene glycol) were used. When octa-*O*-acetyl-lactose **18** was reacted with 2-(2-(2-chloroethoxy)ethoxy)ethanol in presence of $SnCl_4$ and CF_3CO_2Ag, compound **19** was obtained mainly, but not exclusively, as β-anomer (α/β ratio 1:4). Column chromatography purification could anyway yield to pure β-compound in 74% yield.

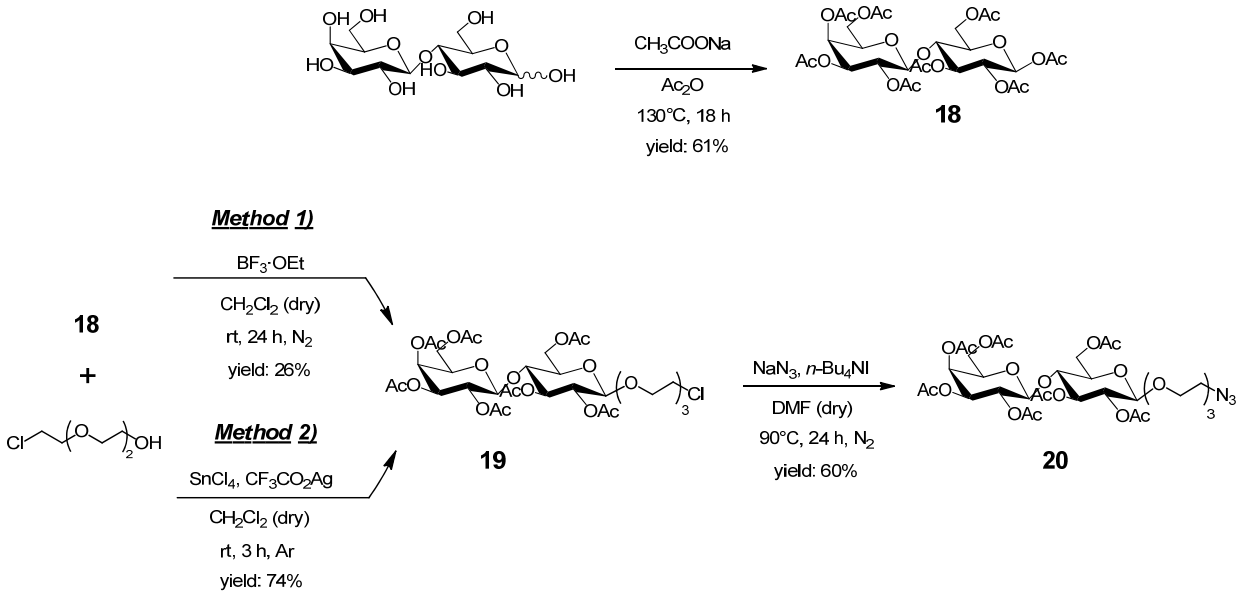


Figure 4.16 Synthetic scheme to obtain azido-lacto-derivative **20**.

The subsequent reaction of chlorinated compound **19** with NaN_3 led to the corresponding azido derivative **20** in 60% yield after 24 hours.

Alt-tetra-propargyl-calix[4]arene **25** was obtained from the corresponding *alt*-amino-calix[4]arene **24** (figure **4.17**) via reaction with EDC in CH_2Cl_2 and pyridine 7:3 as previously described for compound **10**. Amino-*alt*-calixarene **24** was synthesized following a three steps procedure from *p-tert*-butyl-calix[4]arene **21,** that was alkylated at the lower rim with propyl bromide in presence of Cs_2CO_3 in dry CH_3CN to fix the macrocycle in the 1,3-alternate conformation[36,37]. Compound **22** was therefore obtained. Subsequent ipso-nitration allowed the complete substitution of the *p-tert*-butyl groups with nitro groups, yielding nitro derivative **23**[37], that was then treated with hydrazine and Pd/C (10%) to obtain amino-product **24**[38].

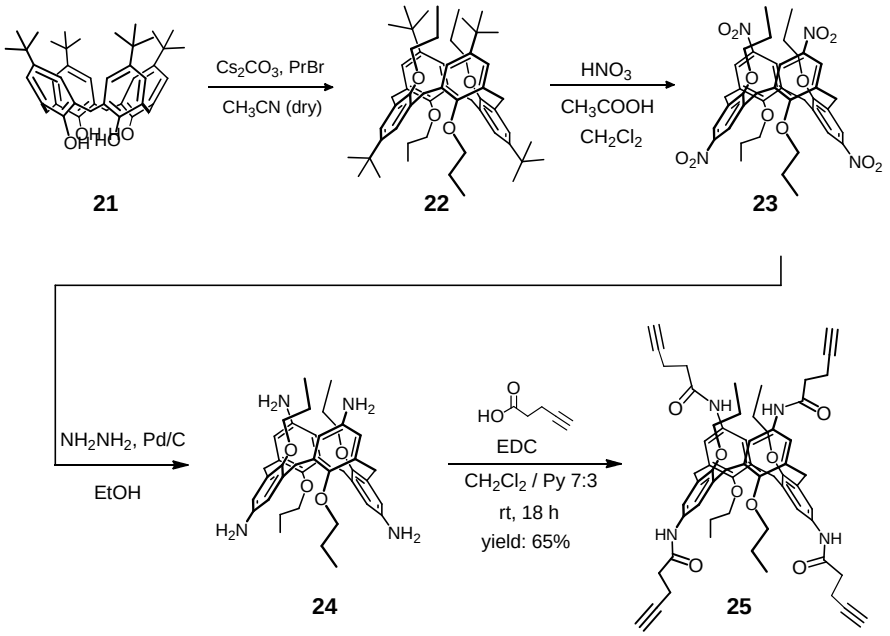


Figure 4.17 Synthetic scheme to obtain tetra-alkyne 1,3-alternate calixarene **25**.

Azido-lactoside **20** was then reacted with the *cone*- and 1,3-alternete-tetra-alkyne-calix[4]arenes (compound **10** and **25**, respectively), via CuAAC "click" reaction, as previously described for the analogous galacto-clusters, to give respectively *cone*-calix[4]arene **26** (figure **4.18**) and *alt*-calix[4]arene **27** (figure **4.19**) in 45-50% yield. Microwave irradiation (150 W, 80 °C) allowed to obtain complete tetra-functionalization in only 40 minutes reaction. Subsequent deacetylation with Zemplén method led to target compounds **16** and **17**. All products were completely characterized via 1D- and 2D- NMR techniques and ESI-MS analyses.

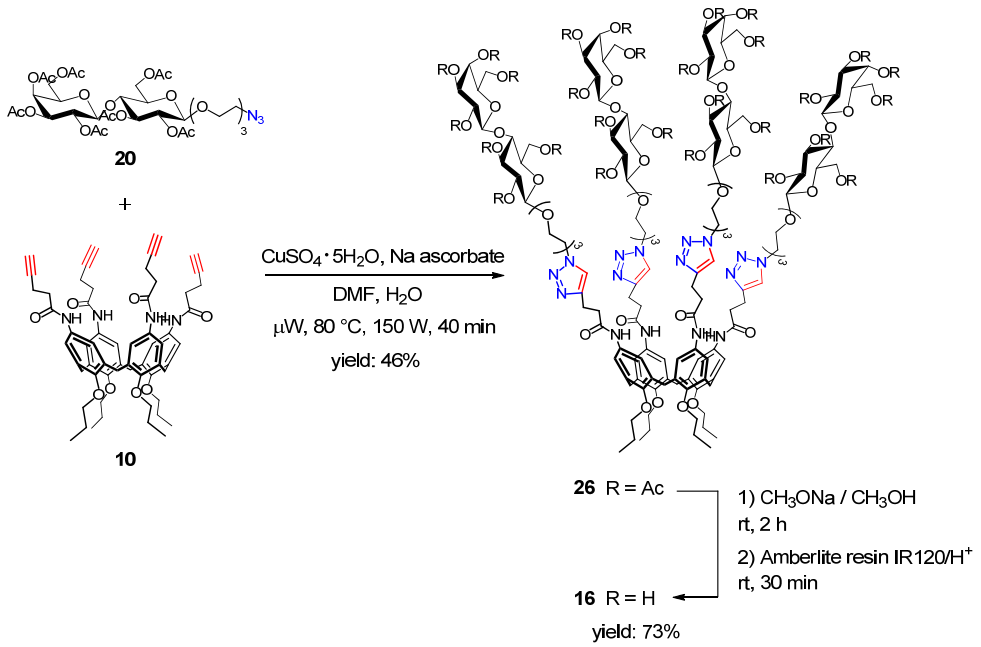


Figure 4.18 CuAAC conjugation reaction between lactosylazide **20** and alkynecalix[4]arene **10** followed by deacetylation to give *cone*-lactosyl-calix[4]arene **16**.

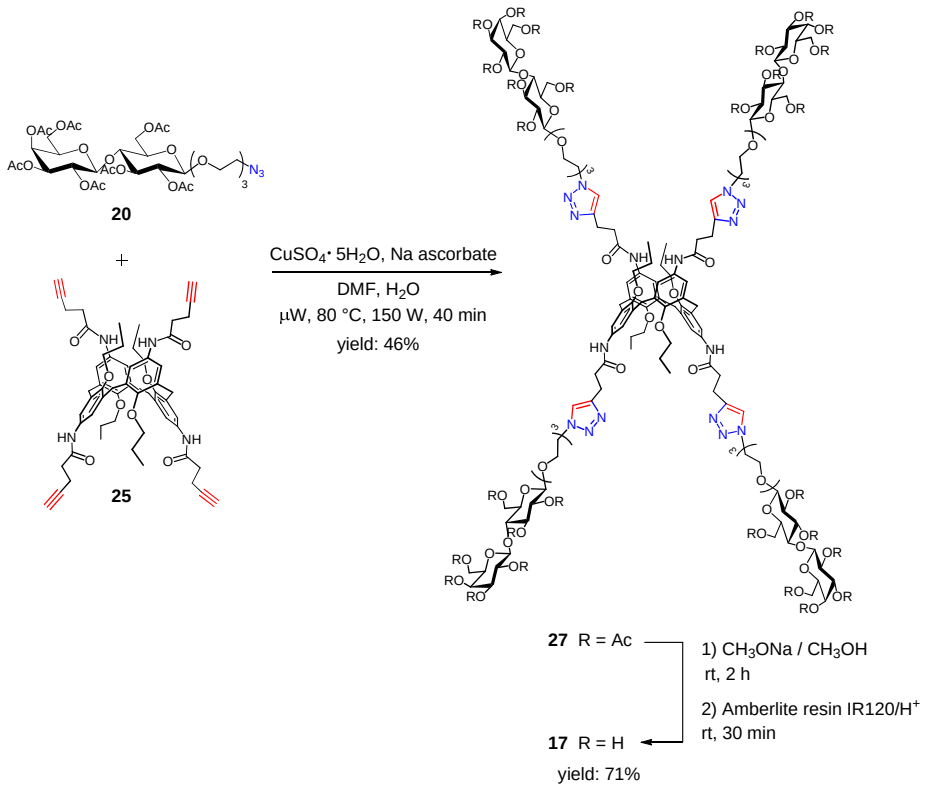


Figure 4.19 CuAAC conjugation reaction between lactosylazide **20** and alkynecalix[4]arene **26** followed by deacetylation to give *alt*-lactosyl-calix[4]arenes **17**.

4.3 Conclusions

Multivalency is a key factor to ensure an high affinity and selectivity in the lectin-carbohydrate recognition. Finding methods to efficiently connect carbohydrates to a multivalent scaffold is a matter of primary importance in the synthesis of glycoclusters. This is why we focused our attention on the development of new methods to covalently connect carbohydrates moieties to calixarenes backbones. The formation of 1,4-disubstituted triazole rings via the so called Copper(I)-cataliyzed Azido-Alkyne Cycloaddition (CuAAC) was the selected methodology to properly link functionalized sugars to multivalent ligands. Two alternative pathways were explored that differ for the relative dispositions of the alkyne and azido groups onto the sugar and calixarene moieties undergoing the 1,3-dipolar cycloaddition reaction. Both pathways have been explored and discussed in this chapter. Initially a reaction between galactoside derivatives and calix[4]arenes blocked in the *cone* conformation has been explored. For the lower number of synthetic steps and the higher stability of the final product under the studied reaction conditions, the route that led to compound **1**, from a tetra-alkyne calixarene scaffold and an azido sugar, was preferred to the other leading to compound **2** from an alkyne galactoside and a tetra-azide calixarene. Once the optimal reaction conditions were established, other two glycoclusters were synthesized: glycocluster **16**, bearing lactoside residues connected via triazole ring to a cone-calix[4]arene, and glycoclusters **17**, which instead is characterized by a 1,3-alternate presentation of the lactoside units. Interestingly, these compounds present a long, mobile, hydrophilic linker between the calixarene core and the carbohydrate units, whose role in the galectin inhibition processes is currently under study.

4.4 Experimental Part

General Information. All moisture sensitive reactions were carried out under nitrogen atmosphere, using previously oven-dried glassware. All dry solvents were prepared according to standard procedures, distilled before use and stored over 3 Å or 4Å molecular sieves. Most of the solvents and reagents were obtained from commercial sources and used

without further purification. Analytical TLC was performed using prepared plates of silica gel (Merck 60 F-254 on aluminum) and then, according to the functional groups present on the molecules, revealed with UV light or using staining reagents: $FeCl_3$ (1% in H_2O/CH_3OH 1:1), H_2SO_4 (5% in EtOH), ninhydrin (5% in EtOH), basic solution of $KMnO_4$ (0.75% in H_2O). Merck silica gel 60 (70-230 mesh) was used for flash chromatography and for preparative TLC plates. ^1H NMR and ^{13}C NMR spectra were recorded on Bruker AV300 and Bruker AV400 spectrometers (observation of ^1H nucleus at 300 MHz and 400 MHz respectively, and of ^{13}C nucleus at 75 MHz and 100 MHz respectively). All chemical shifts are reported in part per million (ppm) using the residual peak of the deuterated solvent, whose values are referred to tetramethylsilane (TMS, $\delta_{TMS} = 0$), as internal standard. All ^{13}CNMR spectra were performed with proton decoupling. Electrospray ionization (ESI) mass analyses were performed with a Waters spectrometer. Gas chromatography mass analyses (GC-MS, electronic impact 70 eV) were recorded on a HP 6890 Series GC System apparatus, equipped with capillary column DB5 and quadrupolar mass selector HP 5973 Mass Selective Detector. Melting points were determined on an Electrothermal apparatus in closed capillaries. Microwave reactions were performed using a CEM Discovery System reactor.

General procedure to obtain peracetylated carbohydrates (compounds 3 and 18)

Sodium acetate (7 eq) was suspended in acetic anhydride (25 eq) and heated to 130 °C. At this temperature D-galactose (1 eq) was added batchwise. The mixture was stirred at 130 °C and checked via TLC (eluent: AcOEt/hexane 1:1). When the reaction was completed, it was poured into a flask containing ice and water (5 times the volume of acetic anhydride). The mixture was then stirred at room temperature overnight. The solid formed was recovered by filtration on Buchner funnel and then purified via trituration in EtOH.

1,2,3,4,6-Penta-*O*-acetyl-β-D-galactopyranoside (3)

Sodium acetate (32.0 g, 0.39 mol), acetic anhydride (134 mL, 1.38 mol), D-galactose (10.0 g, 55.5 mmol) Reaction time: 2 hours. Product **3** was obtained as white solid in 40% yield. **^1H NMR** (300 MHz, $CDCl_3$): δ (ppm) 5.69 (d, 1H, $J_{1\text{-}2}$ = 8.2 Hz, H_1); 5.42 (dd, 1H, $J_{4\text{-}5}$ = 0.9 Hz, $J_{4\text{-}3}$ = 3.4, H_4); 5.33 (dd, 1H, $J_{1\text{-}2}$ = 8.2 Hz, $J_{2\text{-}3}$ = 10.4 Hz, H_2); 5.07 (dd, 1H, $J_{2\text{-}3}$ = 10.4 Hz, $J_{3\text{-}4}$ = 3.4 Hz, H_3); 4.17-4.11 (m, 2H, H_{6a}, H_{6b}); 4.09-4.00 (m, 1H, H_5); 2.16 (s, 3H, Ac); 2.12 (s, 3H, Ac); 2.04 (s, 6H, 2Ac); 1.99 (s, 3H, Ac). The product shows the same physical and spectroscopic properties reported in the literature[39].

2,3,4,6-Tetra-*O*-acetyl-β-D-galactopyranosyl-(1→4)-2,3,6-tri-*O*-acetyl-β-D-glucopyranoside (18)

Sodium acetate (16.1 g, 0.20 mol), acetic anhydride (68 mL, 0.70 mol), D-lactose (10.0 g, 28.0 mmol) Reaction time: 18 hours. Product **18** was obtained as white solid in 61% yield. **^{1}H NMR** (300 MHz, $CDCl_3$): δ(ppm) 5.67 (d, 1H, J = 8.4 Hz, H_1); 5.34 (d, 1H, J = 3.3 Hz, H_4'); 5.24 (t, 1H, J = 9.2 Hz, H_3); 5.15-4.98 (m, 2H, H_2, H_2'); 4.93 (dd, 1H, $J_{2'-3'}$ = 10.4 Hz , $J_{3'-4'}$ = 3.3 Hz, H_3'); 4.51-4.39 (m, 2H, H_1', H_{6a}); 4.17-4.01 (m, 3H, H_{6b}, $H_6'_a$, $H_6'_b$); 3.90-3.78 (m, 2H, H_4, H_5'); 3.77-3.70 (m, 1H, H_5); 2.18 (s, 3H, Ac); 2.11 (s, 3H, Ac); 2.08 (s, 3H, Ac); 2.07-1.99 (m, 12H, Ac); 1.95 (s, 3H, Ac). The product shows the same physical and spectroscopic properties reported in the literature[40].

2-(2-(2-Azidoethoxy)ethoxy)ethanol (4)

2-(2-(2-chloroethoxy)ethoxy)ethanol (0.75 g, 4.44 mmol) was added to a suspension of NaN_3 (0.35 g, 5.38 mmol) and NaI (0.07 g, 4.67 mmol) in absolute ethanol (4 mL). The mixture was heated at 80 °C by microwave irradiation (100 W) for 18 h. The progress of the reaction was checked via TLC (eluent: CH_2Cl_2/CH_3OH 9:1) and GC-MS analyses. The insoluble salts were then filtered off on a Celite pad and the solvent was removed under reduced pressure. The residue was redissolved in CH_2Cl_2 and kept at 4°C overnight to favor impurities precipitation that were then filtered off giving compound **4** as a brownish oil. Yield: 70%. **^{1}H NMR** (300 MHz, $CDCl_3$): δ (ppm) 3.70-3.58 (m, 10H, OCH_2); 3.38 (t, 2H, J = 5.1 Hz, $C\underline{H}_2N_3$); 2.88 (bs, 1H, OH). **EI-MS** (70 eV) m/z: 119 [100%, $(M\text{-}CH_2N_3)^+$]; 89 [20%, $(M\text{-}OCH_2CH_2N_3)^+$]. The product shows the same physical and spectroscopic properties reported in the literature[22].

2-(2-(2-Chloroethoxy)ethoxy)ethoxy-2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranoside (5)

Penta-*O*-acetyl-β-galactose **3** (0.50 g, 1.28 mmol) and 2-(2-(2-chloroethoxy)ethoxy)ethanol (0.28 mL, 1.92 mmol) were dissolved in dry CH_2Cl_2 (5 mL). The mixture was cooled to 0 °C with an ice-water bath and $BF_3{\cdot}Et_2O$ (0.80 mL, 6.48 mmol) was added dropwise. The mixture, kept under N_2, was allowed to slowly reach room temperature and it was stirred for 24 h. The reaction was checked via TLC (eluent: AcOEt/petroleum ether 1:1) and once finished it was quenched by adding $NaHCO_3$ saturated aqueous solution (20 mL). The aqueous phase was extracted with CH_2Cl_2 (3 x 20 mL). The combined organic phases were

washed with water (20 mL), dried over anhydrous Na_2SO_4, filtered and the solvent removed under reduced pressure. The residue was purified by flash chromatography (eluent: AcOEt/petroleum ether 1:1). Product **5** was obtained pure as a yellow oil. Yield: 30%. **^{1}H NMR** (300 MHz, $CDCl_3$): δ (ppm) 5.36 (dd, 1H,$J_{4\text{-}5}$ = 1.0 Hz, $J_{3\text{-}4}$ = 3.3 Hz, H_4); 5.19 (dd, 1H, $J_{1\text{-}2}$ = 8.0 Hz, $J_{2\text{-}3}$ = 10.5 Hz, H_2); 5.01 (dd, 1H, $J_{2\text{-}3}$ = 10.5 Hz, $J_{3\text{-}4}$ = 3.3 Hz, H_3); 4.55 (d, 1H, J = 8.0 Hz, H_1); 4.17-4.07 (m, 2H, H_{6a} H_{6b}); 3.98-3.87 (m, 2H, H_5, β-OCH_a); 3.77-3.59 (m, 11H, β-OCH_b, CH_2 ethylene glycol chain); 2.13, 2.04, 2.03, 1.96 (4s, 12H, Ac). **ESI-MS(+)** m/z : 521.7 [100%, $(M+Na)^+$]. The product shows the same physical and spectroscopic properties reported in the literature[23].

2-(2-(2-Azidoethoxy)ethoxy)ethoxy-2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranoside (6)

Method 1) A solution of penta-*O*-acetyl-β-galactose **3** (0.97 g, 2.50 mmol) in dry CH_2Cl_2 (5 mL) was cooled to 0 °C with an ice-water bath. At this temperature 2-(2-(2-azidoethoxy)ethoxy)ethanol **4** (0.49 g, 2.80 mmol) and $BF_3{\cdot}Et_2O$ (1.1 mL, 8.75 mmol) were added. The mixture was stirred at room temperature under N_2 for 24 h. The reaction was monitored via TLC (eluent: AcOEt/hexane 7:3). The reaction was then quenched by adding water (5 mL) and brine (2 mL) and then it was extracted with CH_2Cl_2 (3 x 10 mL). The combined organic phases were washed with water (20 mL) and brine (20 mL), dried over anhydrous Na_2SO_4, filtered and the solvent removed under reduced pressure. Attempts to purify the residue were made by using flash chromatography (elution gradient: AcOEt/hexane 1:1 → 7:3) and preparative TLC plates (eluent: AcOEt/hexane 7:3), but it was not possible to obtain **6** as a pure compound. Yield: 79% (crude).

Method 2) NaN_3 (0.23 g, 3.60 mmol) and tetrabutyl ammonium iodide (0.53 g, 1.44 mmol) were added to a solution of galactoside derivative **5** (0.36 g, 0.72 mmol) in dry DMF (5 mL). The mixture was stirred at 90 °C for 48 hours under nitrogen atmosphere. ESI-MS analyses were performed to monitor the reaction. The solvent was subsequently removed *in vacuo* and the residue redissolved in AcOEt (30 mL). The organic phase was washed with water (2 x 30 mL), dried over anhydrous Na_2SO_4, filtered and the solvent removed under reduced pressure. Purification via column chromatography (eluent: AcOEt/petroleum ether 1:1) gave pure product **6** as a light yellow oil. Yield: 72%.

^{1}H NMR (300 MHz, $CDCl_3$): δ (ppm) 5.37 (dd, 1H, $J_{4\text{-}5}$ = 1.2 Hz, $J_{4\text{-}3}$ = 4.0 Hz, H_4); 5.19 (dd, 1H, $J_{1\text{-}2}$ = 8.1 Hz, $J_{2\text{-}3}$ = 10.5 Hz, H_2); 5.00 (dd, 1H, $J_{3\text{-}4}$ = 4.0 Hz, $J_{2\text{-}3}$ = 10.5 Hz, H_3); 4.56 (d, 1H, $J_{1,2}$ =

8.1 Hz, H_1); 4.19-4.08 (m, 2H, H_{6a}, H_{6b}); 3.98-3.87 (m, 2H, β-OCH_a, H_5); 3.81-3.71 (m, 1H, β-OCH_b); 3.70-3.59 (m, 8H, CH_2 ethylene glycol chain); 3.87 (t, 2H, J = 5.1 Hz, CH_2N_3); 2.14, 2.05, 2.04, 1.97 (4s, 12H, Ac). **ESI-MS(+)** m/z: 528.2 [100, $(M+Na)^+$]. The product shows the same physical and spectroscopic properties reported in the literature[41].

2-(2-(2-Propargyloxyethoxy)ethoxy)ethanol (7)

t-BuOK (2.92 g, 26.0 mmol) and triethylene glycol (3.47 mL, 26.0 mmol) were solubilized in 150 mL dry THF. The mixture was stirred at room temperature for 30 min with molecular sieves 4 Å under nitrogen atmosphere. Subsequently a propargyl bromide solution (80% in toluene, 2.90 mL, 26.0 mmol) in dry THF (50 mL) was slowly added to the reaction mixture (drop time: 1 hour), that was stirred at room temperature for 2 hours under N_2. The progress of the reaction was monitored via TLC (eluent: AcOEt 100%). The reaction was then quenched by filtration through a Celite pad and the solvent was removed under reduced pressure. The residue was purified via column chromatography (eluent: AcOEt 100%) giving compound **7** as a pale yellow oil. Yield: 65%. **^{1}H NMR** (300 MHz, $CDCl_3$): δ (ppm) 4.15 (d, 2H, J = 2.4 Hz, C<u>H</u>$_2$CCH); 3.64-3.62 (m, 10H, CH_2); 3.55 (t, 2H, J = 4.5 Hz, CH_2OH); 2.74 (bs, 1H, OH); 2.41 (t, 1H, J = 2.4 Hz, CH_2CC<u>H</u>). **ESI-MS(+)** m/z: 211.1 [100%, $(M+Na)^+$]; 189.1 [25%, $(M+H)^+$]. **EI-MS(+)** m/z: 89 [100%, $(CH_2CH_2OCH_2CH_2OH)^+$]; 99 [33%, $(CHCCH_2OCH_2CH_2O)^+$]. The product shows the same physical and spectroscopic properties reported in the literature[25].

2-(2-(2-Propargyloxyethoxy)ethoxy)ethoxy-2,3,4,6-tetra-*O*-acetyl-β-D-galactopiranoside (8)

A solution of propargyl functionalized triethylene glycol **7** (0.68 g, 3.60 mmol) in 5 mL of dry CH_2Cl_2 was added to a solution of peracetylated galactose **3** (1.40 g, 3.60 mmol) in 10 mL of dry CH_2Cl_2. The mixture was cooled to 0 °C and then $BF_3{\cdot}Et_2O$ (2.28 mL, 18.0 mmol) was slowly added. The reaction was stirred at 0 °C for 1 h and then allowed to slowly reach room temperature and further stirred for 14 h under nitrogen atmosphere. The reaction was monitored via TLC (eluent: CH_2Cl_2/AcOEt 7:3). The mixture was then poured into a $NaHCO_3$ aqueous saturated solution (20 mL) and it was stirred till the formation of gas was not observed any longer. The aqueous phase was extracted with CH_2Cl_2 (3 x 20 mL). The combined organic phases were washed with water (20 mL), dried over anhydrous Na_2SO_4,

filtered and the solvent removed under reduced pressure. The crude was purified by flash chromatography, first with an elution gradient: CH_2Cl_2/AcOEt 9:1 → 7:3 and then with petroleum ether/AcOEt 1:1. Product **8** was obtained pure as a yellow oil. Yield: 43%. **^{1}H NMR** (400 MHz, $CDCl_3$): δ (ppm) 5.34 (d, 1H, J = 3.2 Hz, H_4); 5.16 (dd, 1H, $J_{1\text{-}2}$ = 8.0 Hz, $J_{2\text{-}3}$ = 10.4 Hz, H_2); 4.97 (dd, 1H, $J_{3\text{-}4}$= 3.2 Hz, $J_{2\text{-}3}$ = 10.4 Hz, H_3); 4.54 (d, 1H, J = 8.0 Hz, H_1); 4.16 (d, 2H, J = 2.4 Hz, C<u>H</u>$_2$CCH); 4.14–4.07 (m, 2H, H_{6a}, H_{6b}); 3.92–3.86 (m, 2H, H-5, β-OCH_a); 3.74–3.60 (m, 11H, β-OCH_b, ethylene glycol chain); 2.41 (t, 2H, J = 2.4 Hz, C≡C<u>H</u>); 2.11, 2.02, 2.01, 1.94 (4s, 12H, $COCH_3$). **ESI-MS(+)** m/z: 541.8 [100%, $(M+Na)^+$]. The product shows the same physical and spectroscopic properties reported in the literature[25].

***Cone*-5,11,17,23-tetrakis[penta-4-ynoyl-amido]-25,26,27,28-tetrapropoxy-calix[4]arene (10)**

Method 1) 4-Pentynoic acid (0.18 g, 1.84 mmol) and DCC (0.38 g, 1.84 mmol) were dissolved in 3 mL CH_2Cl_2. The solution was stirred for 10 min and then amino-calix[4]arene **9**[24] (0.20 g, 0.31 mmol) and a catalytic amount of DMAP were added. The reaction was refluxed for 5 hours, checking the progress via TLC (CH_2Cl_2/CH_3OH 9:1). The formed precipitate was then filtered off and the filters were washed with CH_2Cl_2 and CH_3OH. The organic phase was collected and the solvent removed in vacuum. The residue was purified by flash chromatography (eluent: CH_2Cl_2/CH_3OH 97:3) and preparative TLC plates (eluent: CH_2Cl_2/CH_3OH 95:5) giving compound **10** as a white solid. Yield: 44%.

Method 2) 4-Pentynoic acid (0.18 g, 1.84 mmol) and EDC (0.35 mg, 1.84 mmol) were dissolved in 14 mL CH_2Cl_2 and 6 mL pyridine. The solution was stirred for 30 min at room temperature and then amino-calix[4]arene **9**[24] (0.20 g, 0.31 mmol) were added. The mixture turned immediately red. The reaction was stirred at room temperature for 18 hours and it was monitored via TLC (eluent: CH_2Cl_2/CH_3OH 9:1). The mixture was then diluted with CH_2Cl_2 (20 mL) and washed with 1M HCl (2 x 20 mL) and water (20 mL). The combined aqueous phases were extracted with CH_2Cl_2 (30 mL). The organic layers were collected together and evaporated to dryness. The residue was purified via column chromatography (eluent: CH_2Cl_2/CH_3OH 97:3) to give product **10** as a white solid. Yield: 66%

^{1}H NMR (300 MHz, CD_3OD): δ (ppm) 6.88 (s, 8H, Ar); 4.46 (d, 4H, J = 13.2 Hz, *ax*-ArCH_2Ar); 3.85 (t, 8H, J = 7.5 Hz, OC<u>H</u>$_2$$CH_2CH_3$); 3.11 (d, 4H, J = 13.2 Hz, *eq*-ArCH_2Ar); 2.49–2.46 (m, 16H, COC<u>H</u>$_2$$CH_2$, $COCH_2$C<u>H</u>$_2$); 2.27 (bs, 4H, C≡CH), 2.00–1.93 (m, 8H, OCH_2C<u>H</u>$_2$$CH_3$); 1.03 (t,

12H, J = 7.2 Hz, $OCH_2CH_2C\underline{H}_3$). **^{13}C NMR** (100 MHz, CD_3OD): δ (ppm) 170.4 (CO); 153.1 (Ar ipso); 134.8 (Ar ortho); 132.1 (Ar para); 120.5 (Ar meta); 82.1 ($\underline{C}\equiv CH$); 76.7 ($O\underline{C}H_2CH_2CH_3$); 69.0 ($C\equiv\underline{C}H$); 35.3 ($CO\underline{C}H_2CH_2$); 30.7 ($ArCH_2Ar$); 22.9 ($OCH_2\underline{C}H_2CH_3$); 14.1 ($COCH_2\underline{C}H_2$); 9.4 ($OCH_2CH_2\underline{C}H_3$). **ESI-MS(+)** m/z: 995.7 [100%, $(M+Na)^+$].

***Cone*-5,11,17,23-tetrakis[propynoyl-amido]-25,26,27,28-tetrapropoxy-calix[4]arene (11)**

Propiolic acid (0.28 mL, 4.50 mmol) and DCC (0.93 g, 4.50 mmol) were dissolved in 5 mL CH_2Cl_2 and the solution was stirred for 10 min at room temperature. Subsequently, amino-calix[4]arene **9** (0.50 g, 0.76 mmol) and a catalytic amount of DMAP were added. The reaction was refluxed for 8 h and checked via TLC (eluent: CH_2Cl_2/CH_3OH 9:1). The precipitate was then filtered off, and the solvent was removed under reduced pressure. ESI-MS analyses revealed the formation of product **11**, but purification by preparative TLC plates (eluent: CH_2Cl_2/CH_3OH 9:1 or CH_2Cl_2/CH_3OH 98:2) and trituration in AcOEt could not yield to pure compound **11**. **ESI-MS(+)** m/z: 884 [100%, $(M+Na)^+$].

***Cone*-5,11,17,23-tetrakis[α-chloroacetamido]-25,26,27,28-tetrapropoxy-calix[4]arene (12)**

N,*N*-Diisopropylethylamine (0.53 mL, 3.06 mmol) and chloroacetyl chloride (0.24 mL, 3.06 mmol) were added to a solution of amino-calix[4]arene **9** (200 mg, 0.306 mmol) in dry CH_2Cl_2 (10 mL). The mixture was stirred at room temperature for 6 h under nitrogen atmosphere. The progress of the reaction was monitored via TLC (eluent: AcOEt/hexane 7:3). The reaction was quenched by stirring the mixture for 10 min with 0.5M HCl (10 mL). The aqueous phase was then extracted with CH_2Cl_2 (2 x 20 mL). The combined organic phases were washed with water till neutral pH, dried over anhydrous Na_2SO_4, filtered and the solvent removed under reduced pressure. The residue was recrystallized from Et_2O/AcOEt 95:5. Product **12** was obtained as a light brown solid. Yield: 55%. **^{1}H NMR** (300 MHz, $CDCl_3$/CD_3OD 9:1): δ (ppm) 6.75 (s, 8H, Ar); 4.31 (d, 4H, J = 13.2 Hz, *ax*-$ArCH_2Ar$); 3.94 (s, 8H, CH_2Cl); 3.70 (t, 8H, J = 7.4 Hz, $O\underline{C}H_2CH_2CH_3$); 3.01 (d, 4H, J = 13.2 Hz, *eq*-$ArCH_2Ar$); 1.88-1.73 (m, 8H, $OCH_2\underline{C}H_2CH_3$); 0.87 (t, 12H, J = 7.4 Hz, $OCH_2CH_2C\underline{H}_3$). **ESI-MS (+)**m/z: 981 [100%, $(M+Na)^+$]. **M.p.** : 208-210 °C. The spectroscopic data found are in agreement with those reported in literature[42].

***Cone*-5,11,17,23-tetrakis[α-azidoacetamido]-25,26,27,28-tetrapropoxy-calix[4]arene (13)**

Calix[4]arene **12** (290 mg, 0.30 mmol) was dissolved in 3 mL DMF and diluted to 30 mL with CH_3OH. NaN_3 (312 mg, 4.80 mmol) was then added and the mixture was refluxed for 1 hour. The reaction was monitored via TLC (eluent: hexane/THF 6:4) and ESI-MS analyses. When complete, the solvent was removed under reduced pressure. The residue was redissolved in CH_2Cl_2 and filtered to remove the inorganic salts. The organic phase was collected, brought to dryness and the crude purified via recrystallization from hexane. Product **13** was obtained as a yellow solid. Yield: 75%. **^{1}H NMR** (300 MHz, CD_3OD): δ (ppm) 6.95 (s, 8H, ArH); 4.50 (d, 4H, J = 13.2 Hz, *ax*-$ArCH_2Ar$); 3.95-3.85 (m, 16H, $OC\underline{H}_2CH_2CH_3$, CH_2N_3); 3.16 (d, 4H, J = 13.2 Hz, *eq*-$ArCH_2Ar$); 2.06-1.90 (m, 8H, $OCH_2C\underline{H}_2CH_3$); 1.04 (t, 12H, J = 7.4 Hz, $OCH_2CH_2C\underline{H}_3$). **^{13}C NMR** (300 MHz, CD_3OD): δ (ppm) 158.8 (CONH); 145.2 (Ar ipso); 136.6 (Ar ortho); 133.4 (Ar para); 122.1 (Ar meta); 78.5 ($O\underline{C}H_2CH_2CH_3$); 53.4 ($CH_2N_3$); 32.4 ($ArCH_2Ar$); 24.7 ($OCH_2\underline{C}H_2CH_3$); 11.1 ($OCH_2CH_2\underline{C}H_3$). **ESI-MS(+)** m/z: 1007.4 [100%, $(M+Na)^+$]. **M.p.** : 201-203 °C.

2-(2-(2-Chloroethoxy)ethoxy)ethoxy-2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl-(1→4)-2,3,6-tri-*O*-acetyl-β-D-glucopyranoside (19)

Method 1) Octa-*O*-acetyl-lactose **18** (2.03 g, 3.00 mmol) and 2-(2-(2-chloroethoxy)ethoxy) ethanol (0.65 mL, 4.50 mmol) were dissolved in 10 mL dry CH_2Cl_2. The mixture was cooled to 0°C with an ice-water bath and $BF_3{\cdot}Et_2O$ (1.90 mL, 15.0 mmol) was slowly added. The mixture, kept under N_2, was allowed to slowly reach room temperature and it was stirred for 24 h. The reaction was checked via TLC (eluent: AcOEt/petroleum ether 7:3). Subsequently, $NaHCO_3$ aqueous saturated solution (30 mL) was added for quenching. The aqueous phase was extracted with CH_2Cl_2 (3 x 30 mL). The combined organic phases were washed with water (30 mL), dried over anhydrous Na_2SO_4, filtered and the solvent removed under reduced pressure. The residue was purified by flash chromatography (elution gradient: AcOEt/petroleum ether 65:35 → 8:2). Product **19** (light yellow oil) was obtained as a mixture of α and β anomers. Yield: 26%.

Method 2) Octa-*O*-acetyl-lactose **18** (2.03 g, 3.00 mmol) and 2-(2-(2-chloroethoxy)ethoxy) ethanol (0.65 mL, 4.50 mmol) were dissolved in 10 mL dry CH_2Cl_2. CF_3COOAg (0.99 g, 4.50 mmol) and a 1M solution of $SnCl_4$ (0.91 mL, 9.00 mmol) in dry CH_2Cl_2 was added dropwise under argon atmosphere. The reaction was stirred at room temperature under Ar for 3

hours. The reaction was monitored via TLC (AcOEt/petroleum ether 7:3). $NaHCO_3$ saturated aqueous solution (30 mL) was subsequently added for quenching. The precipitate was filtered off through a Celite pad. The organic phase was collected and the aqueous phase was extracted with CH_2Cl_2 (4 x 30 mL). The combined organic phases were washed with water and brine till neutral pH, dried over anhydrous Na_2SO_4, filtered and the solvent removed under reduced pressure. The residue was purified by flash chromatography (elution gradient: AcOEt/hexane 65:35 →7:3). Product **19** was obtained pure as a yellow oil. Yield: 74%

^{1}H NMR (300 MHz, $CDCl_3$): δ(ppm) 5.32 (d, 1H, J = 3.3 Hz, H_4'); 5.17 (t, 1H, J = 9.3 Hz, H_3); 5.08 (dd, 1H, $J_{1'-2'}$ = 7.9 Hz, $J_{2'-3'}$ = 10.4 Hz, H_2'); 4.93 (dd, 1H, $J_{2'-3'}$ = 10.4 Hz, $J_{3'-4'}$ = 3.3Hz, H_3'); 4.87 (dd, 1H, J_{1-2} = 7.8 Hz, J_{2-3} = 9.3 Hz, H_2); 4.55 (d, 1H, J = 7.8 Hz, H_1); 4.51-4.38 (m, 2H, H_1', H_{6a}); 4.20-4.00 (m, 3H, H_{6b}, $H_6'_a$, $H_6'_b$); 3.93-3.79 (m, 2H, H_5', β-OCH_a); 3.78-3.68 (m, 4H, H_4, β-OCH_b, CH_2 ethylene glycol chain); 3.67-3.53 (m, 9H, H_5, $3CH_2$ ethylen glycol chain, CH_2Cl), 2.13, 2.12, 2.10, 2.09, 2.04, 2.02, 1.94 (7s, 21H, $COCH_3$). **^{13}C NMR** (300 MHz, $CDCl_3$): δ(ppm) 170.3, 170.1, 170.0, 169.7, 169.6, 169.5, 169.0 ($CH_3\underline{C}$=O); 101.1 (C_1'); 100.6 (C_1); 76.3 (C_4); 72.8 (C_3); 72.6 (C_5); 71.7 (C_2); 71.4 (CH_2 ethylene glycol chain); 71.0 (C_3'), 70.7, 70.6 (C_5', $2CH_2$ ethylene glycol chain); 70.3,70.1 ($2CH_2$ ethylene glycol chain); 69.1 (C_2', β-OCH_2); 66.6 (C_4'); 62.0 (C_6); 60.8 (C_6'); 42.8 (CH_2Cl); 20.9, 20.8, 20.7, 20.6, 20.5 ($\underline{C}H_3C$=O). **ESI-MS(+)** m/z: 809.2 [100%, $(M+Na)^+$]. The product shows the same physical and spectroscopic characteristics reported in literature[43].

2-(2-(2-Azidoethoxy)ethoxy)ethoxy-2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl-(1→4)-2,3,6-tri-*O*-acetyl-β-D-glucopyranoside (20)

NaN_3 (0.15 g, 2.35 mmol) and tetrabutyl ammonium iodide (0.35 g, 0.94 mmol) were added to a solution of lactoside derivative **19** (0.37 g, 0.47 mmol) in dry DMF (10 mL). The mixture was stirred at 90 °C for 24 hours under nitrogen atmosphere. ESI-MS analyses were performed to monitor the reaction. The solvent was subsequently removed *in vacuo* and the residue redissolved in AcOEt (40 mL). The organic phase was washed with water (2 x 30 mL), dried over anhydrous Na_2SO_4, filtered and the solvent removed under reduced pressure. Purification via column chromatography (elution gradient: AcOEt/toluene 65:35 → 7:3) gave pure product **20** as a light yellow oil. Yield: 60%. **^{1}H NMR** (300 MHz, $CDCl_3$): δ (ppm) 5.25 (d, 1H, J = 3.3 Hz, H_4'); 5.10 (t, 1H, J = 9.3Hz, H_3); 5.01 (dd, 1H, $J_{1'-2'}$ = 7.9 Hz, $J_{2',3'}$ = 10.4 Hz, H_2');

4.87 (dd, 1H, $J_{2'-3'}$ = 10.4 Hz, $J_{3'-4'}$ = 3.3 Hz, H_3'); 4.87 (dd, 1H, J_{1-2} = 7.9 Hz, J_{2-3} = 9.3 Hz, H_2); 4.49 (d, 1H, J = 7.9 Hz, H_1); 4.44-4.36 (m, 2H, H_1', H_{6a}); 4.09-3.94 (m, 3H, H_{6b}, $H_6'_a$, $H_6'_b$); 3.88-3.76 (m, 2H, H_5', β-OCH_a); 3.71 (t, 1H, J = 9.3 Hz, H_4); 3.67-3.47 (m, 10H, H_5, β-OCH_b, CH_2 ethylene glycol chain); 3.31 (t, 2H, J = 5.0 Hz, CH_2N_3); 2.06 (s, 3H, $COCH_3$); 2.03 (s, 3H, $COCH_3$); 2.00-1.92 (m, 12H, $COCH_3$); 1.88 (s, 3H, $COCH_3$). **^{13}C NMR** (300 MHz, $CDCl_3$): δ (ppm) 170.3, 170.2, 170.1 , 170.0, 169.7, 169.6, 169.0 ($CH_3\underline{C}O$); 101.0 (C_1'); 100.5 (C_1); 76.2 (C_4); 72.8 (C_3); 72.5 (C_5); 71.6 (C_2); 70.9 (C_3'); 70.7, 70.6 (C_5', $2CH_2$ ethylene glycol chain); 70.3,70.0 ($2CH_2$ ethylene glycol chain); 69.0 (C_2', β-OCH_2); 66.6 (C_4'); 62.0 (C_6); 60.8 (C_6'); 50.6 (CH_2N_3); 20.8, 20.7, 20.6, 20.5, 20.4 ($\underline{C}H_3CO$). **ESI-MS(+)** m/z: 816.2 [100%, $(M+Na)^+$]. The product shows the same physical and spectroscopic properties reported in the literature[43].

***Alt*-5,11,17,23-tetrakis[penta-4-ynoyl-amido]-25,26,27,28-tetrapropoxy-calix[4]arene (25)**

4-Pentynoic acid (90.0 mg, 0.92 mmol) and EDC (176.4 mg, 0.92 mmol) were dissolved in 7 mL CH_2Cl_2 and 3 mL pyridine. The solution was stirred for 30 min at room temperature and then *alt*-tetra-amino calix[4]arene **24**[38] (98.1 mg, 0.15 mmol) was added. The mixture turned immediately red. The reaction was stirred at room temperature for 18 hours and it was monitored via TLC (eluent: CH_2Cl_2/CH_3OH 9:1). The mixture was then diluted with CH_2Cl_2 (10 mL) and washed with 1M HCl (2 x 10 mL) and water (2 x 10 mL). The combined aqueous phases were extracted with CH_2Cl_2 (30 mL). The organic layers were collected together and evaporated to dryness. The residue was purified via column chromatography (eluent: CH_2Cl_2/CH_3OH 97:3) to give product **25** as a white solid. Yield: 65%. **^{1}H NMR** (300 MHz, CD_3OD/$CDCl_3$ 9:1): δ (ppm) 7.31 (s, 8H, Ar); 3.73 (s, 8H, $ArCH_2Ar$); 3.24 (t, 8H, J = 7.4 Hz, $O\underline{H}_2CH_2CH_3$); 2.57-2.45 (m, 16H, $COC\underline{H}_2CH_2$, $COCH_2C\underline{H}_2$); 2.18-2.14 (m, 4H, C≡CH), 1.34-1.23 (m, 8H, $OCH_2C\underline{H}_2CH_3$); 0.67 (t, 12H, J = 7.4 Hz, $OCH_2CH_2C\underline{H}_3$). **^{13}C NMR** (100 MHz, CD_3OD/$CDCl_3$ 9:1): δ (ppm) 171.8 (CO); 154.9 (Ar ipso); 135.7 (Ar ortho); 134.1 (Ar para); 122.5 (Ar meta); 83.8 ($\underline{C}$≡CH); 73.9 ($O\underline{C}H_2CH_2CH_3$); 70.5 (C≡$\underline{C}$H); 39.6 ($Ar\underline{C}H_2Ar$); 37.2 ($CO\underline{C}H_2CH_2$); 24.1 ($OCH_2\underline{C}H_2CH_3$); 16.2 ($COCH_2\underline{C}H_2$); 10.9 ($OCH_2CH_2\underline{C}H_3$). **ESI-MS(+)** m/z: 996.3 [100%, $(M+Na)^+$].

General procedure for "click" reactions (compounds 14, 15, 26, 27)

Calix[4]arene derivative (1 eq) and β-galactoside derivative (6 eq) were dissolved in 2.5 mL of DMF in a microwave tube. $CuSO_4 \cdot 5H_2O$ (0.3 or 0.6 eq), sodium ascorbate (0.6 or 1.2 eq) and 0.5 mL H_2O were then added. The mixture was heated at 80 °C by microwave irradiation

(150 W) for a reaction time between 20 and 40 minutes. When the reaction was completed (checked via TLC and ESI-MS), it was quenched by addition of water (15 mL) and extracted with AcOEt (5 x 15 mL). The combined organic layers were dried over anhydrous Na_2SO_4, filtered and the solvent removed under reduced pressure. The crude was purified by flash chromatography to afford pure the peracetylated glycosyl calix[4]arene.

***Cone*-peracetylated-galactosyl-calix[4]arene (14)**

Cone-tetra-alkyne calix[4]arene **10** (38.9 mg, 0.04 mmol), azido galactoside **6** (121.3 mg, 0.24 mmol), $CuSO_4 \cdot 5H_2O$ (3.0 mg 12.0 µmol), sodium ascorbate (4.7 mg, 24.0 µmol). Reaction time: 20 min. TLC (eluent: CH_2Cl_2/CH_3OH 95:5). Flash chromatography (elution gradient: CH_2Cl_2/CH_3OH 95:5 → 93:7). Product **14** was obtained as a light yellow solid. Yield: 83%. **^{1}H NMR** (400 MHz, $CD_3OD/CDCl_3$ 4:1): δ(ppm) 7.75 (bs, 4H, C<u>H</u> triazole); 6.83 (s, 8H, Ar); 5.35 (s, 4H, H_4); 5.09-5.07 (m, 8H, H_2, H_3); 4.64 (d, 4H, J = 6.8 Hz, H_1); 4.52 (bs, 8H, NC<u>H</u>$_2$CH$_2$); 4.41 (d, 4H, J = 13.2 Hz, *ax*-ArC<u>H</u>$_2$Ar); 4.12-4.10 (m, 8H, H_{6a}, H_{6b}); 4.05-4.02 (m, 4H, H_5); 3.93-3.90 (m, 4H, β-OCH$_a$); 3.84-3.79 (m, 16H, NCH$_2$C<u>H</u>$_2$, OC<u>H</u>$_2$CH$_2$CH$_3$); 3.70-3.67 (m, 4H, β-OCH$_b$); 3.58-3.53 (m, 24H, ethylene glycol chain); 3.09 (d, 4H, J = 13.2 Hz, *eq*-ArC<u>H</u>$_2$Ar); 3.02 (bs, 8H, COCH$_2$C<u>H</u>$_2$triazole); 2.62 (bs, 8H, COC<u>H</u>$_2$CH$_2$triazole); 2.10, 2.01, 1.99 (3s, 36H, COCH$_3$); 1.93 (s, 20H, COCH$_3$, OCH$_2$C<u>H</u>$_2$CH$_3$); 0.98 (t, 12H, J = 7.2 Hz, OCH$_2$CH$_2$C<u>H</u>$_3$). **^{13}C NMR** (100 MHz, $CD_3OD/CDCl_3$ 4:1): δ(ppm) 170.7, 170.6, 170.2, 170.0 (CH$_3$<u>C</u>O); 153.1 (Ar ipso); 146.3 (C$_q$ triazole); 134.9 (Ar ortho); 132.1 (Ar para); 123.3 (CH triazole); 120.6 (Ar meta); 100.9 (C_1); 76.8 (O<u>C</u>H$_2$CH$_2$CH$_3$); 71.0, 69.0 (C_2, C_3); 70.4 (C_5); 70.3, 70.2, 70.0 (CH_2 ethylene glycol chain); 69.0 (NCH$_2$<u>C</u>H$_2$, β-OCH$_2$); 67.4 (C_4); 61.2 (C_6); 50.2 (NCH$_2$); 35.8 (CO<u>C</u>H$_2$CH$_2$triazole); 30.9 (Ar<u>C</u>H$_2$Ar); 23.0 (OCH$_2$<u>C</u>H$_2$CH$_3$); 21.2 (COCH$_2$<u>C</u>H$_2$triazole); 19.9, 19.7 (CO<u>C</u>H$_3$); 9.7 (OCH$_2$CH$_2$<u>C</u>H$_3$). **ESI-MS(+)** m/z: 1520.5 [100%, $(M+2Na)^{2+}$]; 1021.1 [50%, $(M+3Na)^{3+}$].

***Cone*-peracetylated-galactosyl-calix[4]arene (15)**

Cone-tetra-azido calix[4]arene **13** (39.4 mg, 0.04 mmol), alkyne galactoside **8** (124.4 mg, 0.24 mmol), $CuSO_4 \cdot 5H_2O$ (3.0 mg 12.0 µmol), sodium ascorbate (4.7 mg, 24.0 µmol). Reaction time: 20 min. TLC (eluent: CH_2Cl_2/CH_3OH 95:5). Flash chromatography (elution gradient: CH_2Cl_2/CH_3OH 95:5 → 93:7). Product **15** was obtained as a light yellow solid. Yield: 81%. **^{1}H NMR** (300 MHz, $CD_3OD/CDCl_3$ 3:1): δ (ppm) 7.89 (bs, 4H, CH triazole); 6.98 (bs, 8H, Ar); 5.35 (d, 4H, J = 2.1 Hz, H_4); 5.13-5.08 (m, 16H, H_2, H_3, COC<u>H</u>$_2$triazole); 4.65 (d, 4H, J = 7.2 Hz, H_1);

4.58 (bs, 8H, OC$\underline{H}_2$triazole); 4.46 (d, 4H, J = 12.9 Hz, *ax*-ArCH_2Ar); 4.13-4.10 (m, 8H, H_{6a}, H_{6b}); 4.05-4.00 (m, 4H, H_5); 3.93-3.85 (m, 12H, β-OCH_a, OC$\underline{H}_2CH_2CH_3$); 3.74-3.59 (m, 44H, β-OCH_b, ethylene glycol chain); 3.12 (d, 4H, J = 12.9 Hz, *eq*-ArCH_2Ar); 2.12 (s, 12H, CH_3CO); 2.03-1.92 (m, 44H, CH_3CO, $OCH_2C\underline{H}_2CH_3$); 1.00 (t, 12H, J = 7.2 Hz, $OCH_2CH_2C\underline{H}_3$). **^{13}C NMR** (75 MHz, $CD_3OD/CDCl_3$ 3:1): 170.7, 170.6, 170.2, 170.1 ($CH_3\underline{C}O$); 163.6 (CONH); 153.3 (Ar ipso); 144.8 (C_q triazole); 135.1 (Ar ortho); 131.9 (Ar para); 125.6 (CH triazole); 120.1 (Ar meta); 100.9 (C_1); 77.0 ($O\underline{C}H_2CH_2CH_3$); 71.0, 69.0 ($C_2$, C_3); 70.4 (C_5); 70.3, 70.2, 70.1 (ethylene glycol chain); 69.5 (β-$OCH_2\underline{C}H_2$); 68.9 (β-$O\underline{C}H_2$); 67.4 (C_4); 63.7 ($O\underline{C}H_2$triazole); 61.3 (C_6); 52.4 ($CO\underline{C}H_2$triazole); 31.0 (Ar$\underline{C}H_2$Ar); 23.0 ($OCH_2\underline{C}H_2CH_3$); 20.0, 19.8, 19.7 ($CO\underline{C}H_3$); 9.7 ($OCH_2CH_2\underline{C}H_3$). **ESI-MS(+)** m/z: 1042.8 [100%, $(M+3Na)^{3+}$]; 1552.5 [50%, $(M+2Na)^{2+}$].

***Cone*-peracetylated-lactosyl-calix[4]arene (26)**

Cone-tetra-alkyne calix[4]arene **10** (23.5 mg, 24.1 μmol), azido lactoside **20** (119.8 mg, 0.15 mmol), $CuSO_4{\cdot}5H_2O$ (3.6 mg 14.5 μmol), sodium ascorbate (5.7 mg, 28.9 μmol). Reaction time: 40 min. TLC (eluent: CH_2Cl_2/CH_3OH 95:5). Flash chromatography (elution gradient: CH_2Cl_2/CH_3OH 97:3 → 94:6). Product **26** was obtained as a yellow oil. Yield: 46%.

^{1}H NMR (300 MHz, CD_3OD): δ (ppm) 7.78 (bs, 4H, CH triazole); 6.89 (bs, 4H, Ar); 6.87 (bs, 4H, Ar); 5.33 (d, 4H, J = 3.3 Hz, H_4'); 5.16 (t, 4H, J = 9.2 Hz, H_3); 5.10 (dd, 4H, $J_{2'-3'}$ = 10.4 Hz, $J_{3'-4'}$ = 3.3 Hz, H_3'); 4.99 (dd, 4H, $J_{1'-2'}$ = 7.8 Hz, $J_{2',3'}$ = 10.4 Hz, H_2'); 4.82 (dd, 4H, J_{1-2} = 7.8 Hz, J_{2-3} = 9.2 Hz, H_2); 4.73-4.63 (m, 8H, H_1, H_1'); 4.62-4.40 (m, 16H, H_{6a}, *ax*-ArC$\underline{H}_2$Ar, NCH_2); 4.21-4.06 (m, 16H, H_5', H_{6b}, $H_6'_a$, $H_6'_b$); 3.94-3.78 (m, 24H, H_4, β-OCH_a, $NCH_2C\underline{H}_2$, OC$\underline{H}_2CH_2CH_3$); 3.77-3.61 (m, 8H, H_5, β-OCH_b); 3.60-3.43 (m, 24H, CH_2 ethylene glycol chain); 3.12 (d, 4H, J = 13.0 Hz, *eq*-ArC$\underline{H}_2$Ar); 3.01 (t, 8H, J = 6.8 Hz, $COCH_2C\underline{H}_2$triazole); 2.63 (t, 8H, J = 6.8 Hz, $COC\underline{H}_2CH_2$triazole); 2.14, 2.11 (2s, 24H, $COCH_3$); 2.06-1.87 (m, 68H, $COCH_3$, $OCH_2C\underline{H}_2CH_3$); 1.02 (t, 12H, J = 7.4, $OCH_2CH_2C\underline{H}_3$). **^{13}C NMR** (75 MHz, CD_3OD): δ (ppm) 170.9, 170.7, 170.6, 170.5, 170.3, 170.0, 169.9, 169.7 ($CH_3\underline{C}O$, CONH); 153.0 (Ar ipso); 146.2 (C_q triazole); 134.8 (Ar ortho); 132.3 (Ar para); 122.9 (CH triazole); 120.4 (Ar meta); 100.6, 100.4 (C_1, C_1'); 76.8 ($O\underline{C}H_2CH_2CH_3$); 76.3 ($C_4$); 73.1 ($C_3$); 72.6 ($C_5$); 71.7 ($C_2$); 71.1 ($C_3$'); 70.4 ($C_5$); 70.3, 70.1, 70.0, 69.9, 69.3, 69.1, 68.9 (C_5', C_2', β-OCH_2, $3CH_2$ ethylene glycol chain, $NCH_2\underline{C}H_2$,); 67.2 (C_4'); 62.3 (C_6); 60.8 (C_6'); 50.0 (NCH_2); 35.6 ($CO\underline{C}H_2CH_2$triazole); 30.7 (Ar$\underline{C}H_2$Ar); 23.0 ($OCH_2\underline{C}H_2CH_3$); 20.8 ($COCH_2\underline{C}H_2$triazole); 19.8, 19.5, 19.4, 19.3, 19.1 ($CO\underline{C}H_3$); 9.4 ($OCH_2CH_2\underline{C}H_3$). **ESI-MS(+)** m/z: 1405.3 [100% $(M+3Na)^{3+}$]; 1059.8 [10% $(M+4Na)^{4+}$].

***Alt*-peracetylated-lactosyl-calix[4]arene (27)**

Alt-tetra-alkyne calix[4]arene **25** (11.9 mg, 12.2 µmol), azido lactoside **20** (58.1 mg, 73.2 µmol), $CuSO_4 \cdot 5H_2O$ (1.8 mg 7.3 µmol), sodium ascorbate (2.9 mg, 14.6 µmol). Reaction time: 40 min. TLC (eluent: CH_2Cl_2/CH_3OH 95:5). Flash chromatography (elution gradient: CH_2Cl_2/CH_3OH 97:3 → 94:6). Product **27** was obtained as a yellow oil. Yield: 46%. **^{1}H NMR** (300 MHz, CD_3OD): δ (ppm) 7.80 (bs, 4H, CH triazole); 7.33 (bs, 4H, Ar); 7.32 (bs, 4H, Ar); 5.33 (d, 4H, J = 3.3 Hz, H_4'); 5.16 (t, 4H, J = 9.2 Hz, H_3); 5.10 (dd, 4H, $J_{2'-3'}$ = 10.4 Hz, $J_{3'-4'}$ = 3.3 Hz, H_3'); 4.99 (dd, 4H, $J_{1'-2'}$ = 7.8 Hz, $J_{2',3'}$ = 10.4 Hz, H_2'); 4.82 (dd, 4H, J_{1-2} = 7.8 Hz, J_{2-3}= 9.2 Hz, H_2); 4.73-4.63 (m, 8H, H_1, H_1'); 4.58-4.43 (m, 12H, H_{6a}, NCH_2); 4.19-4.05 (m, 16H, H_5', H_{6b}, $H_6'_a$, $H_6'_b$); 3.94-3.80 (m, 16H, H_4, β-OCH_a, $NCH_2C\underline{H}_2$); 3.79-3.63 (m, 16H, H_5, β-OCH_b, $ArC\underline{H}_2Ar$); 3.60-3.43 (m, 24H, $3CH_2$ ethylene glycol chain); 3.22 (t, 8H, J = 6.9 Hz, $OC\underline{H}_2CH_2CH_3$); 3.05 (t, 8H, J = 7.2 Hz, $COCH_2C\underline{H}_2$triazole); 2.71 (t, 8H, J = 7.2 Hz, $COC\underline{H}_2CH_2$triazole); 2.11, 2.08, 2.03, 2.02, 2.01, 1.99, 1.91 (7s, 84H, $COCH_3$); 1.38-1.18 (m, 8H, $OCH_2C\underline{H}_2CH_3$); 0.64 (t, 12H, J = 7.4 Hz, $OCH_2CH_2C\underline{H}_3$). **^{13}C NMR** (75 MHz, CD_3OD): δ (ppm) 170.9, 170.7, 170.6, 170.5, 170.3, 170.0, 169.9, 169.7 ($CH_3\underline{C}O$, CONH); 153.3 (Ar ipso); 146.2 (C_q triazole); 134.2 (Ar ortho); 132.9 (Ar para); 123.0 (CH triazole); 120.8 (Ar meta); 100.6, 100.3 (C_1, C_1'); 76.3 (C_4); 73.1 (C_3); 72.7 (C_5); 72.3 ($O\underline{C}H_2CH_2CH_3$); 71.7 ($C_2$); 71.1 ($C_3$'); 70.3, 70.2, 70.2, 69.9, 69.3, 69.1, 68.8 (C_5', C_2', β-OCH_2, $3CH_2$ ethylene glycol chain, $NCH_2\underline{C}H_2$,); 67.2 (C_4'); 62.3 (C_6); 60.9 (C_6'); 50.0 (NCH_2); 38.0 ($Ar\underline{C}H_2Ar$); 35.9 ($CO\underline{C}H_2CH_2$triazole); 22.5 ($OCH_2\underline{C}H_2CH_3$); 21.2 ($COCH_2\underline{C}H_2$triazole); 19.8, 19.5, 19.4, 19.3, 19.1 ($CO\underline{C}H_3$); 9.2 ($OCH_2CH_2\underline{C}H_3$). **ESI-MS(+)** m/z: 1405.3 [100% $(M+3Na)^{3+}$]; 1060.3 [80%, $(M+4Na)^{4+}$].

General procedure for deacetylation reactions (compounds 1, 2, 16, 17)

Peracetylated glyco-clusters were dissolved in MeOH and drops of a freshly prepared methanol solution of MeONa were added till pH 8-9. The mixture was stirred at room temperature for 1 or 2 hours. The progress of the reaction was monitored via TLC and/or ESI-MS. When a precipitate was observed, H_2O was added to help complete solubilisation. Amberlite resin IR 120/H^+ was subsequently added for quenching and gently stirred for 30 min till neutral pH. The resin was then filtered off and the solvent removed under vacuum to give pure product.

***Cone*-galactosyl-calix[4]arene (1)**

Cone-peracetylated-galactosyl-calix[4]arene **14** (30.0 mg, 10.0 µmol) in 4 mL MeOH. Reaction time: 1 hour. TLC (eluent: butanol/acetone/H_2O 35:35:15). Product **1** was obtained as a white solid. Yield: quantitative. **^{1}H NMR** (300 MHz, CD_3OD): δ (ppm) 7.81 (s, 4H, CH triazole); 6.92 (s, 8H, Ar); 4.52 (t, 8H, J = 4.8 Hz, NCH_2); 4.45 (d, 4H, J = 13.2 Hz, *ax*-$ArCH_2Ar$); 4.25 (d, 4H, J = 7.2 Hz, H_1); 4.00-3.91 (m, 4H, β-OCH_a); 3.90-3.78 (m, 20H, H_4, $OCH_2CH_2CH_3$, NCH_2CH_2); 3.77-3.71 (m, 8H, H_{6a}, H_{6b}); 3.70-3.61 (m, 4H, β-OCH_b); 3.60-3.44 (m, 36H, H_2, H_3, H_5, CH_2 ethylene glycol chain); 3.11 (d, 4H, J = 13.2 Hz, *eq*-$ArCH_2Ar$); 3.01 (t, 8H, J = 7.4 Hz, $COCH_2CH_2$triazole); 2.64 (t, 8H, J =7.4 Hz, $COCH_2CH_2$triazole); 2.03-1.87 (m, 8H, $OCH_2CH_2CH_3$); 1.01 (t, 12H, J = 7.4 Hz, $OCH_2CH_2CH_3$). **^{13}C NMR** (100 MHz, CD_3OD): δ (ppm) 171.0 (NHCO); 153.0 (Ar ipso); 146.2 (C_q triazole); 134.8 (Ar ortho); 132.3 (Ar para); 123.2 (CH triazole); 120.5 (Ar meta); 103.7 (C_1); 76.8 ($OCH_2CH_2CH_3$); 75.3, 73.5, 71.2 (C_2, C_3, C_4); 70.0 (ethylene glycol chain); 68.9 (C_4, NCH_2CH_2); 68.2 (β-OCH_2); 61.2 (C_6); 50.1 (NCH_2); 35.7 ($COCH_2CH_2$triazole); 30.7 ($ArCH_2Ar$); 22.9 ($OCH_2CH_2CH_3$); 20.9 ($COCH_2CH_2$triazole); 9.4 ($OCH_2CH_2CH_3$). **ESI-MS(+)** m/z:797.0 [100%, $(M+3Na)^{3+}$]; 1184.0 [25%, $(M+2Na)^{2+}$].

***Cone*-galactosyl-calix[4]arene (2)**

Cone-peracetylated-galactosyl-calix[4]arene **15** (60.0 mg, 19.6 µmol) in 10 mL MeOH. Reaction time: 1 hour. TLC (eluent: AcOEt/isopropanol/H_2O 5:2:1). Product **2** was obtained as a white solid. Yield: quantitative. **^{1}H NMR** (300 MHz, CD_3OD): δ (ppm) 9.70 (bs, 4H, NHCO); 8.04 (bs, 4H, CH triazole); 7.01 (s, 8H, Ar); 5.22 (bs, 8H, $COCH_2$triazole); 4.62 (bs, 8H, OCH_2triazole); 4.45 (d, 4H, J = 12.6 Hz, *ax*-$ArCH_2Ar$); 4.24 (d, 4H, J = 7.2 Hz, H_1); 4.02-3.92 (m, 4H, β-OCH_a); 3.91-3.80 (m, 12H, H_4, $OCH_2CH_2CH_3$); 3.79-3.56 (m, 52H, H_{6a}, H_{6b}, β-OCH_b, CH_2 ethylene glycol chain); 3.55-3.44 (m, 12H, H_2, H_3, H_5); 3.12 (d, 4H, J = 12.6 Hz, *eq*-$ArCH_2Ar$); 2.01-1.87 (m, 8H, $OCH_2CH_2CH_3$); 1.02 (t, 12H, J = 7.3 Hz, $OCH_2CH_2CH_3$). **^{13}C NMR** (75 MHz, CD_3OD): δ (ppm) 164.0 (NHCO); 153.2 (Ar ipso); 144.5 (C_q triazole); 135.0 (Ar ortho); 132.0 (Ar para); 125.9 (CH triazole); 120.1 (Ar meta); 103.7 (C_1); 76.9 ($OCH_2CH_2CH_3$); 75.3, 73.5 (C_3, C_5), 71.2 (C_2); 70.1, 69.4 (ethylene glycol chain); 68.9 (C_4); 68.2 (β-OCH_2); 63.5 (OCH_2triazole); 61.2 (C_6); 52.3($COCH_2$triazole); 30.7 ($ArCH_2Ar$); 23.0 ($OCH_2CH_2CH_3$); 9.4 ($OCH_2CH_2CH_3$). **ESI-MS(+)** m/z: 818.3 [100%, $(M+3Na)^{3+}$]; 1215.9 [40%, $(M+2Na)^{2+}$].

***Cone*-lactosyl-calix[4]arene (16)**

Cone-peracetylated-lactosyl-calix[4]arene **26** (46.0 mg, 11.1 µmol) in 10 mL MeOH. Reaction time: 2 hours. Product **16** was obtained as a white solid. Yield: 73%. **^{1}H NMR** (300 MHz, CD_3OD): δ (ppm) 7.82 (s, 4H, CH triazole); 6.93 (bs, 4H, Ar); 6.91 (bs, 4H, Ar); 4.54 (t, 8H, J = 4.5 Hz, NCH_2); 4.45 (d, 4H, J = 13.0 Hz, *ax*-Ar$C\underline{H}_2$Ar); 4.40-4.29 (m, 8H, H_1, H_1'); 4.00-3.64 (m, 44H, H_4', H_{6a}, H_{6b}, H_6'$_a$, H_6'$_b$, β-OCH_2, $NCH_2C\underline{H}_2$, $OC\underline{H}_2CH_2CH_3$); 3.63-3.39 (m, 48H, H_2', H_3', H_5', H_3, H_4, H_5, $3CH_2$ ethylene glycol chain); 3.26 (t, 4H, J = 8.3 Hz, H_2); 3.13 (d, 4H, J = 13.0, *eq*-Ar$C\underline{H}_2$Ar); 3.02 (t, 8H, J = 7.2 Hz, $COCH_2C\underline{H}_2$triazole); 2.65 (t, 8H, J = 7.2 Hz, $CO\underline{CH}_2CH_2$triazole); 2.08-1.85 (m, 8H, $OCH_2C\underline{H}_2CH_3$); 1.02 (t, 12H, J = 7.4 Hz, $OCH_2CH_2C\underline{H}_3$). **^{13}C NMR** (75 MHz, CD_3OD): δ (ppm) 170.9 (CONH); 152.9 (Ar ipso); 146.2 (C_q triazole); 134.8 (Ar ortho); 132.4 (Ar para); 123.1 (CH triazole); 120.4 (Ar meta); 103.7 (C_1'); 102.9 (C_1); 79.3 (C_4); 76.8 ($O\underline{C}H_2CH_2CH_3$); 75.7, 75.1, 74.9, 73.4, 73.3, 71.1 (C_2, C_3, C_5, C_2', C_3', C_5'); 70.0 ($3CH_2$ ethylene glycol chain); 69.1 ($NCH_2\underline{C}H_2$); 68.9 (C_4'); 68.4 (β-OCH_2); 61.1, 60.6 (C_6, C_6'); 49.9 (NCH_2); 35.7 ($CO\underline{C}H_2CH_2$triazole); 30.7 (Ar$\underline{C}H_2$Ar); 23.0 ($OCH_2\underline{C}H_2CH_3$); 20.9 ($COCH_2\underline{C}H_2$triazole); 9.5 ($OCH_2CH_2\underline{C}H_3$). **ESI-MS(+)** m/z: 1013.6 [100% $(M+3Na)^{3+}$]; 1508.5 [50%, $(M+2Na)^{2+}$].

A*lt*-lactosyl-calix[4]arene (17)

Alt-peracetylated-lactosyl-calix[4]arene **27** (23.0 mg, 5.5 µmol) in 5 mL MeOH. Reaction time: 2 hours. Product **17** was obtained as a white solid. Yield: 71%. **^{1}H NMR** (400 MHz, CD_3OD): δ (ppm) 8.07 (bs, 4H, CH triazole); 7.35 (bs, 8H, Ar); 4.64 (t, 8H, J = 4.7 Hz, NCH_2); 4.37 (d, 4H, J = 7.7 Hz, H_1'); 4.35 (d, 4H, J = 8.0 Hz, H_1); 4.04-3.95 (m, 4H, β-OCH_a); 3.94-3.68 (m, 40H, H_4', H_{6a}, H_{6b}, H_6'$_a$, H_6'$_b$, $NCH_2C\underline{H}_2$, β-OCH_b, Ar$C\underline{H}_2$Ar); 3.67-3.39 (m, 48H, $3CH_2$ ethylene glycol chain, H_2', H_3', H_5', H_3, H_4, H_5); 3.31-3,21 (m, 12H, H_2, $OC\underline{H}_2CH_2CH_3$); 3.13 (t, 8H, J = 7.1 Hz, $COCH_2C\underline{H}_2$triazole); 2.79 (t, 8H, J = 7.1 Hz, $CO\underline{CH}_2CH_2$triazole); 1.38-1.19 (m, 8H, $OCH_2C\underline{H}_2CH_3$); 0.67 (t, 12H, J = 7.4, $OCH_2CH_2C\underline{H}_3$). **^{13}C NMR** (100 MHz, CD_3OD): δ(ppm) 170.7 (CONH); 153.3 (Ar ipso); 145.7 (C_q triazole); 134.3 (Ar ortho); 132.8 (Ar para); 124.1 (CH triazole); 120.9 (Ar meta); 103.8 (C_1'); 102.9 (C_1); 79.4 (C_4); 75.7, 75.1, 74.9, 73.4, 71.2, (C_3, C_5, C_2', C_3', C_5'); 73.3 (C_2); 72.3 ($O\underline{C}H_2CH_2CH_3$); 70.0 ($H_2$, $3CH_2$ ethylene glycol chain); 68.9, 68.8 ($NCH_2\underline{C}H_2$, C_4'); 68.4 (β-OCH_2); 61.1, 60.6 (C_6, C_6'); 50.6 (NCH_2); 38.0 (Ar$\underline{C}H_2$Ar); 35.4 ($CO\underline{C}H_2CH_2$triazole); 22.5 ($OCH_2\underline{C}H_2CH_3$); 20.7 ($COCH_2\underline{C}H_2$triazole); 9.2 ($OCH_2CH_2\underline{C}H_3$). **ESI-MS(+)** m/z: 1013.6 [100%, $(M+3Na)^{3+}$]; 1508.6 [20%, $(M+2Na)^{2+}$].

4.5 References

1 Andrè, S.; Sansone, F.; Kaltner, H.; Casnati, A.; Kopitz, J.; Gabius, H. J.; Ungaro, R. *ChemBioChem*, **2008**, *9*, 1649-1661.

2 André, S.; Grandjean, C.; Gautier, F. M.; Bernardi, S.; Sansone, F.; Gabius, H. J.; Ungaro, R. *Chem. Commun.*, **2011**, *47*, 6126-6128.

3 Garcia Fernández J. M.; Ortiz Mellet C. *Adv. Carbohydr. Chem. Biochem.*, **2000**, *55*, 35-135.

4 Sansone, F.; Baldini, L.; Casnati, A.; Ungaro, R. *New J. Chem.,* **2010**, *34*, 2715-2728.

5 Marra, A.; Scherrmann, M. C.; Dondoni, A.; Casnati, A.; Minari, P.; Ungaro, R. *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 2479-2481.

6 Dondoni, A.; Marra, A.; Scherrmann, M. C.; Casnati, A.; Sansone, F.; Ungaro, R. *Chemistry Eur. J.* **1997**, *3*, 1774-1782.

7 Schädel, U.; Sansone, F.; Casnati, A.; Ungaro, R. *Tetrahedron* **2005**, *61*, 1149-1154.

8 M. Dubber, T. K. Lindhorst, *Synthesis* **2001**, 327-330.

9 A. Nelson, J. F. Stoddart, *Carbohydr. Res.* **2004**, *339*, 2069-2075.

10 Huisgen, R. *Angew. Chem., Int. Ed. Engl.* **1963**, 2, 565.

11 Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem. Int. Ed.*, **2001**, *40*, 2004-2021.

12 Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem.* **2002,** 114, 2708-2711; *Angew. Chem. Int. Ed.*, **2002**, *41* (14), 2596-2599.

13 Chabre, Y. M.; Giguère, D.; Blanchard, B.; Rodrigue, J.; Rocheleau, S.; Neault, M.; Rauthu, S.; Papadopoulos, A.; Arnold, A. A.; Imberty, A.; Roy, R. *Chem. Eur. J.,* **2011**, *17*, 6545-6562.

14 Cecioni, S.; Faure, S.; Darbost, U.; Bonnamour, I.; Parrot-Lopez, H.; Roy, O.; Taillefumier, C.; Wimmerová, M.; Praly, J. P.; Imberty, A.; Vidal, S. *Chem. Eur. J.*, **2011**, *17*, 2146-2159.

15 Sánchez-Navarro, M.; Muñoz, A.; Illescas, B. M.; Javier Rojo, J.; Martín, N. *Chem. Eur. J.*, **2011**, *17*, 766-769.

16 Pukin, A. V.; Branderhorst, H. M.; Sisu, C.; Weijers, C. A. G. M.; Gilbert, M.; Liskamp, R. M. J.; Visser, G. M.; Zuilhof, H.; Pieters, R. J. *ChemBioChem*, **2007**, *8*, 1500-1503.

17 Dalvie, D. K.; Kalgutkar, A. S.; Khojasteh-Bakht, S. C.; Obach, R. S.; O'Donnell, J. P. *Chem. Res. Toxicol.* **2002**, *15*, 269-299.

18 Tome, A. C.; Storr, R. C.; Gilchrist, T. L. *Science of Synthesis* Eds. Thieme: New York, **2004,** vol. 13, pp. 415-601.

19 Krivopalov, V. P.; Shkurko, O. P. *Russ. Chem. Rev.* **2005,** 74, 339.

20 Horne, W. S.; Yadav, M. K.; Stout, C. D.; Ghadiri, M. R. *J. Am. Chem. Soc.* **2004**, *126*, 15366-15367.

21 Sansone, F.; Rispoli, G.; Casnati, A.; Ungaro, R. *Multivalent Glycocalixarenes, in "Synthesis and Biological Applications of Glycoconjugates"*, Eds. O. Renaudet and N. Spinelli, Bentham Science Publishers, **2011**, p. 36-63. doi: 10.2174/978160805277611101010036.

22 Calvé, G.; Boutal, H.; Hoang, A.; Perraut, F.; Volland, H.; Renard, P. Y.; Romieu, A. *Org. Biomol. Chem.*, **2008**, *6*, 3065-3078.

23 Sasaki, A.; Murahashi, N.; Yamada, H.; Morikawa, A. *Biol. Pharm. Bull.*, **1994**, *17*, 680-685.

24 Sansone, F.; Chierici, E.; Casnati, A.; Ungaro, R. *Org. Biomol. Chem.,* **2003**, *1*, 1802-1809.

25 Michel, O.; Ravoo, B. J. *Langmuir*, **2008**, *24*, 12116-12118.

26 Huisgen, R. in *1,3-Dipolar Cycloaddition Chemistry*; (Ed.: A. Padwa), Wiley, New York, **1984**, Vol. 1, pp. 1-176.

27 Tornoe, C. W.; Christensen, C.; Meldal, M. *J. Org. Chem.* **2002,** 67, 3057-3062.

28 Perez-Balderas, F.; Ortega-Munoz, M.; Morales-Sanfrutos, J.; Hernandez-Mateo, F.; Calvo-Flores, F. G.; Calvo-Asin, J. A.; Isac-Garcia, J.; Santoyo-Gonzalez, F. *Org. Lett.,* **2003**, *5* (11), 1951-1954.

29 Malkoch, M.; Schleicher, K.; Drocknmuller, E.; Hawker, C. J.; Russel, T. P.; Wu, P.; Fokin, V. V. *Macromolecules* **2005,** 38, 3663.

30 Wu, P.; Feldman, A. K.; Nugent, A. K.; Hawker, C. J.; Scheel, A.; Voit, B.; Pyun, J.; Frèket, J. M. J.; Sharpless, K. B.; Fokin, V. V. *Angew. Chem. Int. Ed.* **2004,** 43, 3928.

31 Rodionov, V. O.; Fokin, V. V.; Finn, M. G. *Angew. Chem. Int. Ed.,* **2005**, *44*, 2211-2215.

32 Bock, V. D.; Hiemstra, H.; van Maarseeven, J. H. *Eur. J. Org. Chem.*, **2006**, 51-68.

33 Joosten, J. A. F.; Tholen, N. T. H.; Ait El Maate, F.; Brouwer, A. J.; van Esse, G. W.; Rijkers, D. T. S.; Liskamp, R. M. J.; Pieters, R. J. *Eur. J. Org. Chem.* **2005**, 3182-3185

34 Zemplen, G.; Pascu, E.; *Ber. Dtsch. Chem. Ges.*, **1929**, *62*, 1613-1618.

35 Xue, J. L.; Cecioni, S.; He, L.; Vidal, S.; Praly, J. P. *Carbohydrate Research,* **2009**, *344*, 1646-1653.

36 Iwamoto, K.; Araki, K.; Shinkai, S. *J. Org. Chem.* **1991***, 56,* 4955-4962*.*

37 Sansone, F.; Dudič, M.; Donofrio, G.; Rivetti, C.; Baldini, L.; Casnati, A.; Cellai, S.; Ungaro*, R. J. Am. Chem. Soc.* **2006***, 128,* 14528-14536*.*

38 Mislin, G.; Graf, E.; Hosseini, M. W. *Tetrahedron Lett.* **1996***, 37,* 4503-4506*.*

39 Liu, Z.; Byun H. S.; Bittman, R. *Org. Lett.*, **2010**, *12* (13), 2974-2977.

40 Šardzík, R.; Noble, G. T.; Weissenborn, M. J.; Martin, A.; Webb S. J.; Flitsch S. L. *Beilstein J. Org. Chem.* **2010,** *6,* 699-703.

41 Bouillon, C.; Mayer, A.; Vidal, S.; Jochum, A.; Chevolot, Y.; Colarec, J. P.; Praly, J. P.; Vasseur, J. J.; Morvant, F. *J. Org. Chem.*, **2006**, *71*, 4700-4702.

42 Alyapyshev, M. Y.; Babain, V. A.; Boyko, V. I.; Eliseev, I. I.; Kirsanov, D. O.; Klimchuk, O. V.; Legin, A. V.; Mikhailina, E. S.; Rodik, R. V.; Smirnov, I. V. *J. Incl. Phenom. Macrocycl. Chem*., **2010**, *67*, 117-126.

43 Kato, H.; Uzawa, H.; Nagatsuka, T.; Kondo, S.; Sato, K.; Ohsawa, I.; Kanamori-Kataoka, M.; Takei, Y.; Ota, S.; Furuno, M.; Dohi, H.; Nishida, Y.; Seto, Y. *Carbohydrate Research*, **2011**, 346, 1820-1826.

Chapter 5

Pentavalent glycocalix[5]arenes for cholera toxin inhibition

Abstract[§]

One of the most interesting and studied cases of glycoside cluster effect concerns the interactions between Cholera Toxin (CT) and its natural receptor the ganglioside GM1. The synthesis of four pentavalent glycocalix[5]arenes functionalized, respectively, with five residues of galactose, lactose, GM2os or GM1os is presented in this chapter. Having the same valency as the B pentameric subunit of the target protein, these compounds could be ideal inhibitors for cholera toxin, potentially forming 1:1 multivalent complexes instead of intermolecular aggregates of variable stoichiometry. Also a monovalent compound, analogous in structure to the calixarene based ligands, was synthesized to be used as reference in the ELISA experiments in order to disclose possible multivalent effects in the inhibition processes.

5.1 Introduction

5.1.1 Cholera toxin structure and mechanism of action

Cholera Toxin (CT) is produced by the *Vibrio cholerae* bacterium (also called *Kommabacillus*). It is a gram negative, facultative[1], comma-shaped bacterium (figure **5.1**) that causes cholera in humans[2,3]. It was first identified and isolated by the Italian anatomist Filippo Pacini in 1859 and studied in detail by the German physician Robert Koch starting from 1886. This bacterium is responsible for the gastrointestinal life-threatening disease of the cholera, which is an infection of the small intestine. The main symptoms are profuse watery diarrhea and vomiting. Transmission occurs primarily through drinking water or eating food contaminated by feces from an infected person. The severity of the diarrhea and vomiting can lead to rapid dehydration and electrolyte imbalance, and thus to death if not promptly treated. Worldwide it affects 3-5 million people and causes 100,000-130,000 deaths per year[4].

[§] *Part of the research presented in this chapter was carried out in the group of Prof. Han Zuilhof at the Wageningen University (NL), during a three weeks work exchange within an European COST programme.*

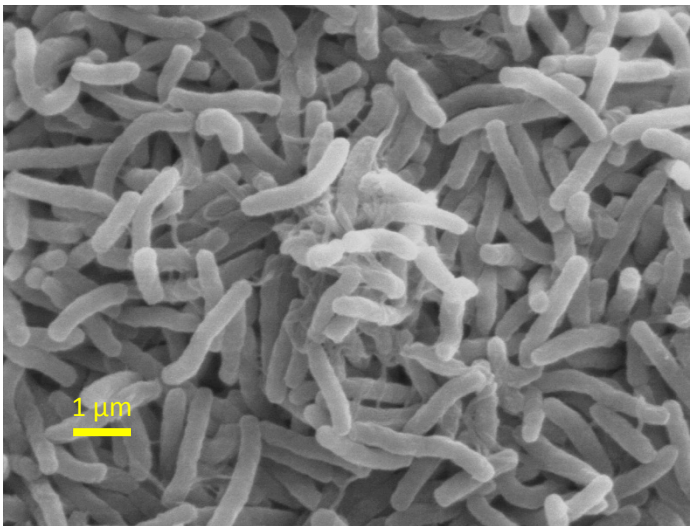


Figure 5.1 SEM (Scanning Electron Microscope) image of *Vibrio cholerae* bacteria, the pathogen that infects the digestive system during the development of the cholera ailment. Image adapted from ref. 5.

Currently there are no effective prophylactics nor vaccines with long lasting protective effects to prevent this disease. Oral rehydration can significantly reduce the fatality rate, but it requires an intense labor and supplies of clean water. Since cholera epidemics usually outbreak in regions with contaminated water sources, the spread control of this disease and the cure of infected people can result very difficult and sometimes even impossible. Therefore the development of an effective treatment can lead to worldwide benefits.

Cholera toxin belongs to the family of AB toxins[6]. These toxins are characterized by having an enzymatically active A subunit, responsible for inducing toxicity, and a B subunit, responsible for binding to cell surface. In particular cholera toxin is a member of the subset AB_5 family of toxins. It presents a heterohexameric structure, with a single copy of the A subunit (CTA) connected to a homopentameric ring-shaped B subunit (CTB), as shown in figure **5.2**.

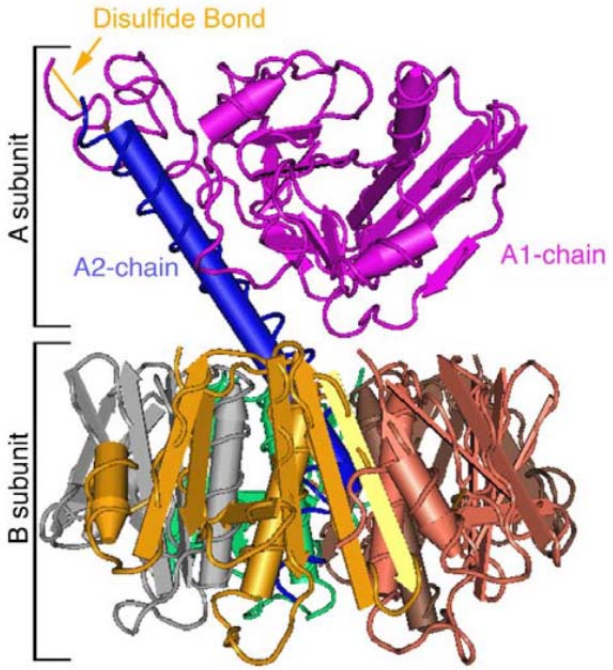


Figure 5.2 The three-dimensional structure of cholera toxin (side view). It is possible to distinguish the heterohexameric structure AB_5, with the toxic enzymatically active A subunit further divided into the A1- (in purple) and the A2-chain (in blue), and the pentameric B domain (in five different colors) responsible for binding to cell surface. Image from ref. 7.

The crystal structure of cholera toxin was first reported in 1991 by Sixma and coworkers[8]. In 1995, Zhang et al. have solved the three-dimensional crystal structure of CT[9] and compared it with a closely related toxin belonging to the same AB_5 family, the heat labile enterotoxin (LT) that shares over 80% sequence identity with CT in both A and B subunits. Other typical members of this family of toxins include the shiga toxin, the shiga-like toxins, and pertussis toxin. In cholera toxin, the 27 kDa A part is divided into two components, part A1 and A2, held together by extensive non-covalent forces and a single covalent disulfide bond. The wedge-shaped A1 portion is an enzyme that is able to ADP-ribosylate G proteins, while the elongated A2 domain consists of an α-helix plus a "tail" that fits into the central pore of the B-pentamer ring via non-covalent interactions. The five identical B subunits are tightly linked together in a pentameric toroid-like structure with a combined molecular weight of approximately 58 kDa. Each B subunit can bind to the oligosaccharide portion of the GM1 ganglioside (GM1os, figure **5.3**)[10] that is located on the surface of the intestinal epithelium cells, giving an overall very strong association due to multivalent effect.

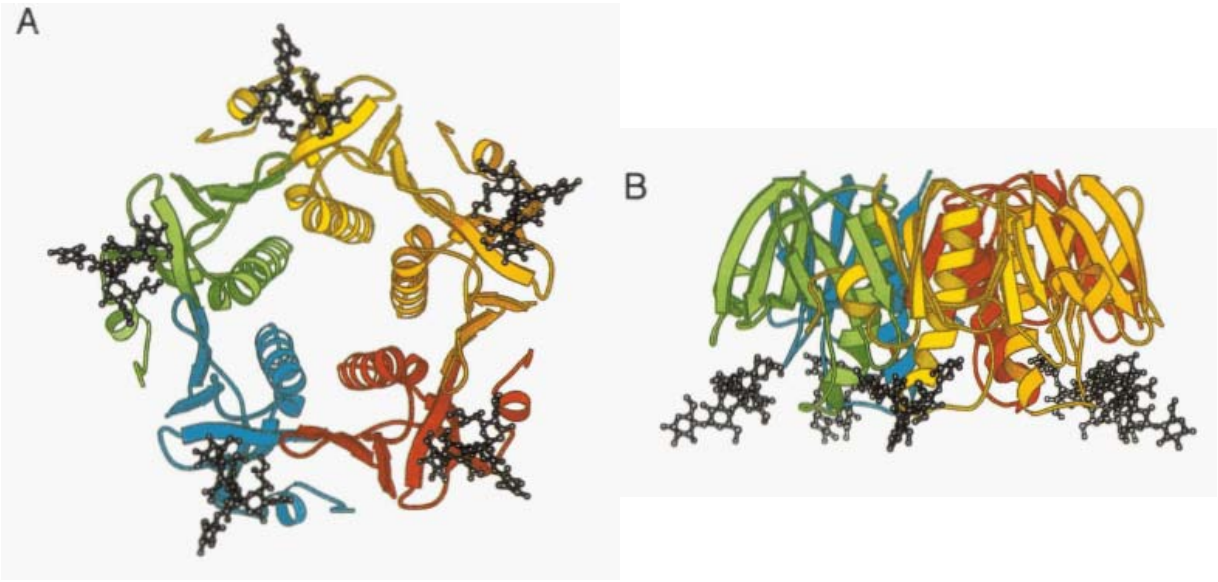


Figure 5.3 A snapshot of the structure of cholera toxin B-pentamer complexed with five GM1 pentasaccharides. A) apical view. B) lateral view. Image from ref. 10.

Upon binding of CT to the cell surface, thanks to the pentameric ring of B subunits, the A unit is taken up into the cell via receptor-mediated endocytosis. Once inside the cell, the A1 subunit induces a series of events that end up with massive cAMP production, which in turn leads to secretion of H_2O, Na^+, K^+, Cl^-, and HCO_3^- into the lumen of the small intestine. The consequence is the rapid dehydration that constitutes the main symptom in people affected by cholera disease. To contrast this mechanism, it is possible to design potential inhibitors able to interfere at three different points. The first method is to use molecules able to block the enzyme active site located in the A1 subunit, the second method is to interrupt the A2-B interaction, the third method is to block the adhesion of CT to cell membrane gangliosides. Although the first two strategies have been followed by several groups, we decided, as other authors, to focus on the third approach, by designing and synthesizing receptor-binding antagonists, that is to say inhibitors able to prevent the binding of the B unit to the cell surface. The design of strong inhibitors is nowadays of great interest, mainly for the obtainment of drugs able to prevent or treat the cholera disease, but also for detection of CT in water, food, patients' samples[11] or in material of suspected terroristic origin[12]. Last but not least, it is also of clearly basic interest to prove once again that the multivalent approach can be successful in the design of inhibitors for multivalent proteins.

5.1.2 Gangliosides

Glycolipids are compounds ubiquitously found in animals, plants and microorganisms. They consist of a lipidic part connected to a carbohydrate residue. Due to the complexity of the structures they may form, this class of glycoconjugates comprises a large variety of structurally heterogeneous compounds. One of the most important classes of glycolipids are the glycosphingolipids. and a particular subtype of glycosphingolipids are gangliosides. Gangliosides present a quite complicate structure. They consist of a ceramide unit connected to a oligo-saccharide moiety which presents at least a sialic acid moiety attached to a galactose or to a hexosamine residue (figure **5.4**). Considering that the IUPAC nomenclature is quite long and complicated, the Svennerholm nomenclature[13,14] is the most convenient and frequently used method to rapidly identify gangliosides. According to this system gangliosides are identified by combining letter G (indicating that the compound belongs to the ganglioside class) with M, D, T, Q, P that gives the degree of presence of sialic acid moieties in the molecule (mono, di, tri, etc. respectively). To this label a number (sometimes also accompanied by a letter) is added, to univocally identify the molecule. In this thesis our attention is mainly focused on GM1 ganglioside, whose oligosaccharide portion is the natural cholera toxin ligand. The GM1 oligosaccharide (GM1os) is a branched pentasaccharide with the Galβ1-3GalNAcβ1-4(Neu5Acα2-3)Galβ1-4Glc sequence. All the sugars are in the D-configuration and the glucose moiety is covalently linked to the ceramide residue (fig. **5.4**).

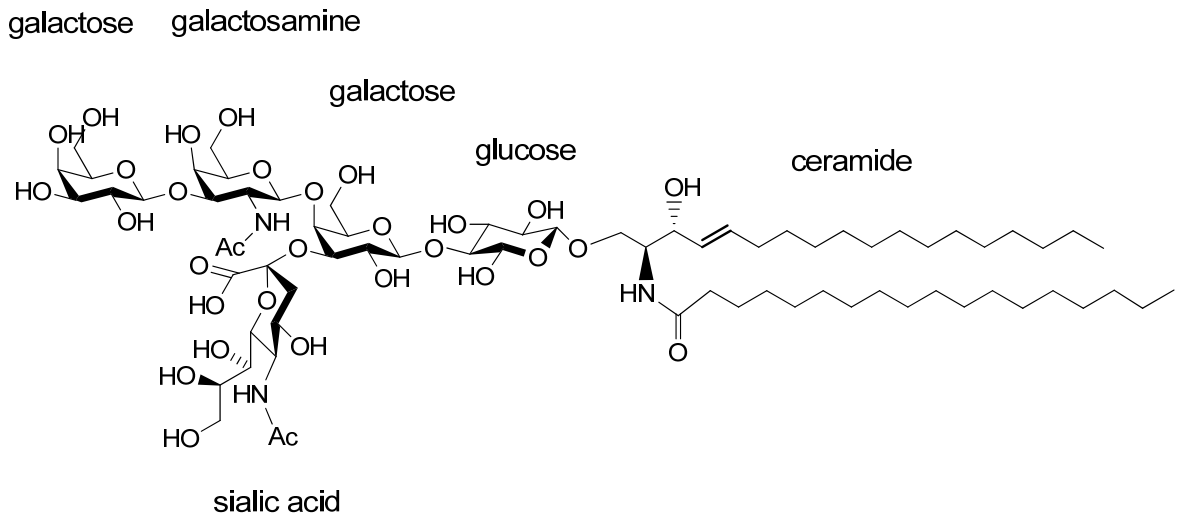


Figure 5.4 Structure of ganglioside GM1.

In cells, the ceramide portion is firmly embedded into the outer plasma membrane, exposing the carbohydrate part from the cell surface where the interaction with the CT may take place.

5.1.3 Cholera toxin – GM1os binding

The monovalent interaction of the GM1 oligosaccharide with one CTB binding site is reported to have a K_d (dissociation constant) of 43 nM[15], which places it among the highest affinity carbohydrate-protein interactions known so far. In a biological multivalent environment, such as the cell surface, the strength of this interaction might be amplified thanks to a positive glycoside cluster effect between the B units of CT and the multiple copies of GM1. To deeply understand how the binding occurs, crystallographic studies have been carried out. In 1994 Meritt *et al*. solved the structure of cholera toxin B-pentamer complexed with GM1 pentasaccharide at 2.2 Å resolution[10] (figure **5.5**), refined in 1998 to 1.25 Å resolution[16].

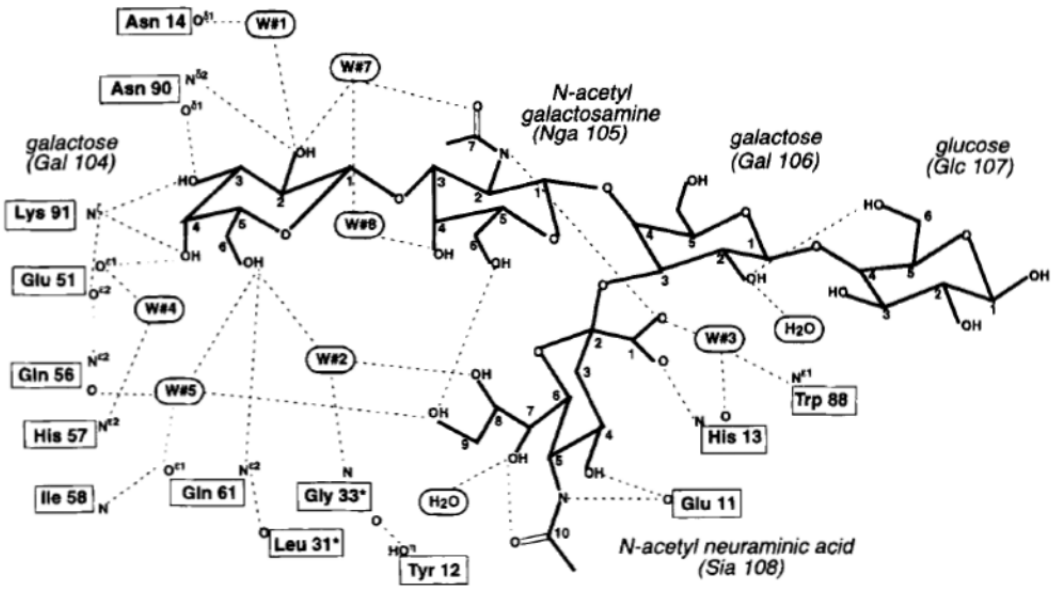


Figure 5.5 Schematic representation of the binding sites of cholera toxin B subunit in complex with GM1 oligosaccharide. Image from ref. 10.

From these structures it could be determined that GM1os interacts with CT mainly through the two terminal galactose and sialic acid residues. A smaller contribution comes from the N-acetyl galactosamine unit, while the lactoside seems not to interact with the toxin, thus mainly working as structural motif. This bivalent binding of the branched GM1os holding the B subunit of the protein is often called "two-fingered grip", with the sialic acid acting as thumb and the terminal Galβ(1→3)GalNAc as forefinger. Upon binding, 43% of the buried surface area is due to interactions with sialic acid, 39% is attributable to terminal galactose and 17% is the contribute of the N-acetyl galactosamine.

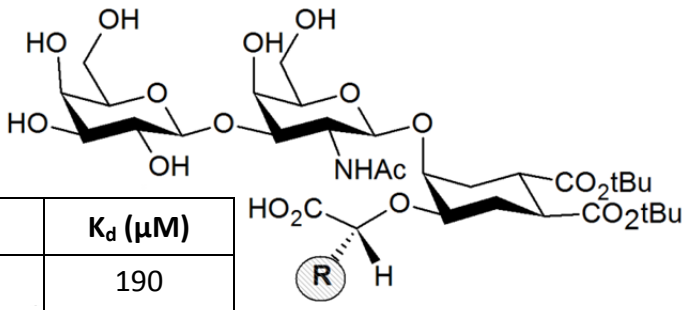
The interactions established between ligand and receptor can be both direct and water-mediated. Each receptor binding site on the toxin was found to be positioned mainly within a single B subunit, with only one water-mediated hydrogen bond from an adjacent B(+1) subunit. Hydroxyl groups of the terminal galactose accept hydrogen bonds directly from nitrogen donors present in the lateral chains of Asn90, Lys91 and Gln61, and are involved in direct solvent mediated bonds with several other amino acid residues. Additionally, the galactose lipophilic face makes extensive contact with Trp88 through hydrophobic and CH-π interactions. The sialic acid gives direct hydrogen bonds with Glu11 and His13 and hydrophobic interactions with Tyr12. Besides, it makes one direct hydrogen bond with the N-acetylgalactosamine and several water-mediated interactions both with the terminal galactose and with other amino acid residues. It is interesting to notice that water molecule 2 mediates the interaction between terminal galactose, sialic acid and the Gly33 residue of an adjacent B(+1) subunit. N-acetylgalactosamine does not exhibit any direct or water-mediated hydrogen bonds with the protein, and neither does the lactose residue. A large number of crystal structures of complexes between galactose derivatives and CTB or LTB (B-pentamer of cholera toxin and heat-labile toxin, respectively) have been reported[17,18,19], including that of Tn antigen[20], which represents the Galβ(1→3)GalNAc forfinger of GM1os. In all these structures, galactose shows the same orientation and essentially identical structures of the binding site.

Considering the importance of the cholera toxin-GM1os interaction in the progression of the cholera disease, it is not surprising that many studies have been reported to better understand and rationalize the complexation mechanism, the kinetics and the thermodynamics of the interaction for GM1 and analogous ligands. To this aim several

techniques have been applied: Surface Plasmon Resonance (SPR) biosensing[21,22], fluorescence spectroscopy[23,24], flow cytometry (FACS)[25], Atomic Force Microscopy (AFM)[26], isothermal titration calorimetry (ITC)[15,27]. Of particular interest is the study made by Turnbull *et al.*[15] which used displacement ITC to measure the key thermodynamics parameters and determine the intrinsic contribution of each monosaccharide residue to the overall CTB-GM1os binding affinity. The GM1 pentasaccharide was dissected into smaller fragments and the standard free energies changes analyzed. It was found that the terminal galactose and the sialic acid residue contribute 54% and 44% of the intrinsic binding energy, confirming the main role played in the complexation by the two terminal GM1 carbohydrates, as also pointed out by crystallographic data.

5.1.4 Synthetic inhibitors for cholera toxin

Since the adhesion of the B subunit of cholera toxin to GM1os on cell surface is the prerequisite for the internalization of the A subunit, many studies have been focused on the synthesis of new multivalent glycosylated ligands that could bind to the CTB pentamer with high affinity. Both natural GM1os-containing derivatives and synthetic GM1os mimics were used as saccharide units. Bernardi *et al.*[28] synthesized a series of glycomimetics that present at the same time a good affinity for CTB subunit and a simpler structure compared to the natural GM1os. This allows an easier synthesis and a higher versatility that make these compounds accessible even in a multigram scale (figure **5.6**).



R	K_d (μM)
methyl	190
cyclohexyl	45
benzyl	10

Figure 5.6 GM1os mimics.

These mimics, called pseudo-GM1 show good affinity for CT (K_d between 190 and 10 µM) especially if compared with simple sugars such as galactose and lactose (K_d of 40,000 and 81,000 µM, respectively). A parallel strategy that can be used combines the use of mimics and their multivalent presentation on macrocyclic scaffolds to increase the overall strength of the protein-carbohydrate interaction. Our group, in collaboration with the group of Prof. Bernardi, synthesized a divalent calixarene ligand bearing two pseudo-GM1 units (R = CH_3) linked through two long and mobile spacers[29] (figure **5.7**).

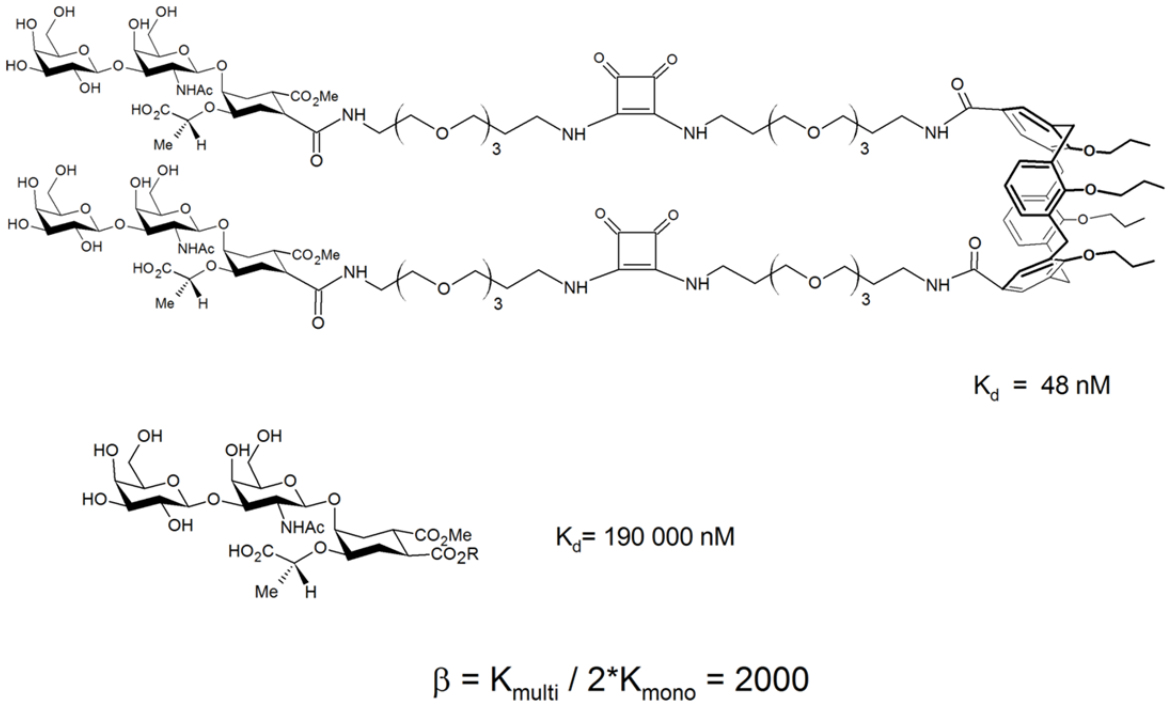


Figure 5.7 Divalent ligand functionalized with two pseudo-GM1 units (up), and pseudo-GM1 mimic (bottom).

Spectrofluorimetric analyses have shown a 2000-fold enhancement in the complexation efficiency for the multivalent ligand compared to the monovalent pseudo-GM1.

Pieters and coworkers have reported the synthesis of a series of highly effective inhibitors of CTB using authentic GM1 oligosaccharides separated with flexible spacers of optimal length to dendritic multivalent scaffolds[30] (figure **5.8**).

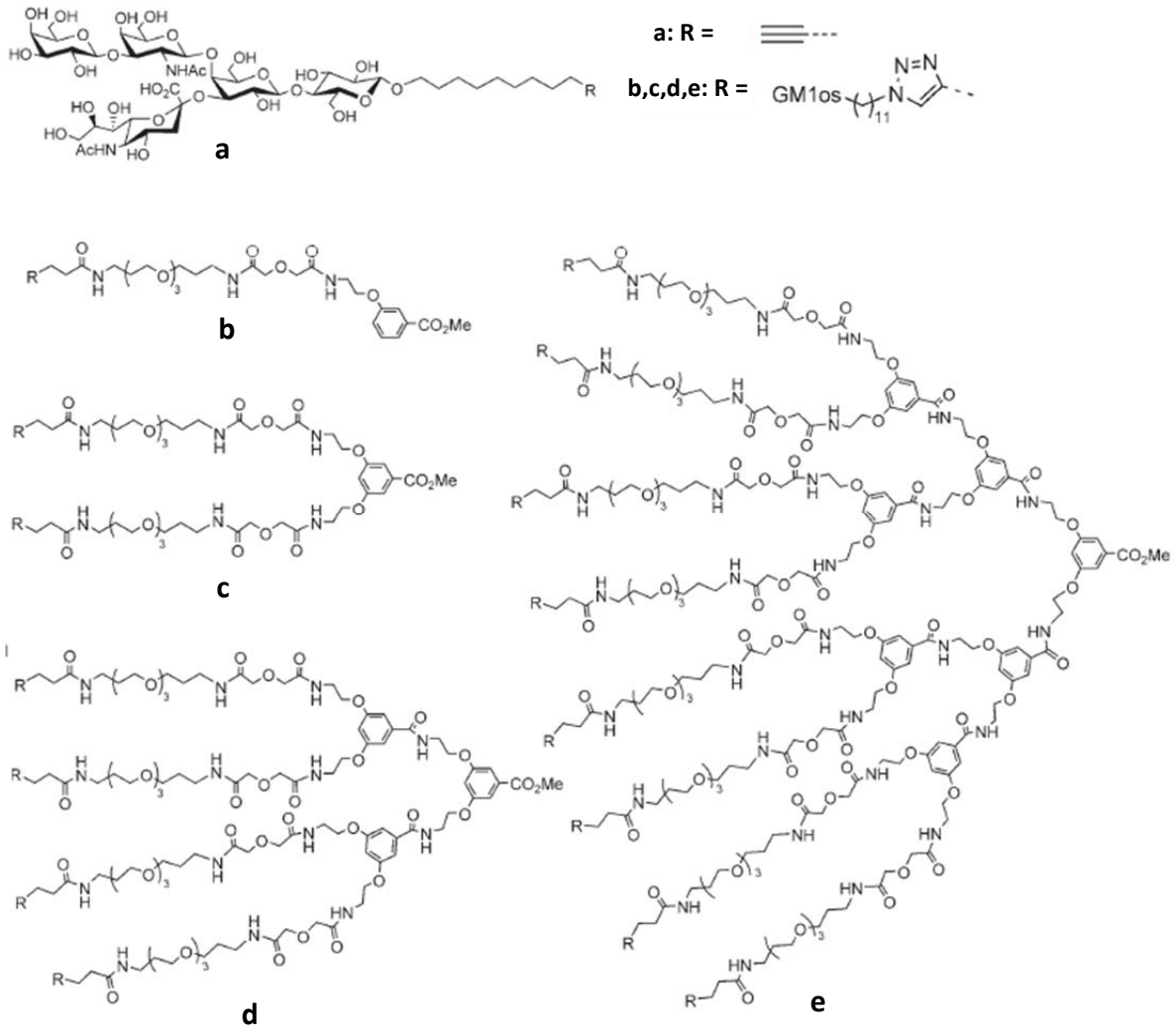


Figure 5.8 GM1os conjugated dendrimeric ligands.

The potency of these dendrimeric compounds was evaluated via ELISA-type assays. The combined effect of high efficiency of the single ligand and multivalency led to inhibitors that showed unprecedented affinity. The octavalent derivative (compound **e** in figure **5.8**) exhibited an IC_{50} of 50 pM, while the monovalent derivative (compound **b** in figure **5.8**) used as reference showed an IC_{50} in the micromolar range (19 μM) (table **5.1**). The octavalent compound **e** showed therefore to be 47,500-fold more potent, per GM1os unit, than the monovalent GM1os conjugate **a**.

Table 5.1 Potency of the dendritic inhibitors **b-e** compared to the monovalent compound **a**, measured by ELISA experiments with CTB-HRP (0.43 nM) and wells coated with GM1 (0.2 μg). Relative potency of compound **a** was taken to be 1 and the potencies of the remaining compounds were calculated with respect to compound **a**.

GM1os derivatives		IC_{50} [M]	Relative potency (per sugar)
Compound	Valency		
a	1	$(1.9 \pm 0.6) \times 10^{-5}$	1 (1)
b	1	$(7 \pm 3) \times 10^{-6}$	2.7 (2.7)
c	2	$(2 \pm 1) \times 10^{-9}$	9 500 (4 750)
d	4	$(2.3 \pm 0.7) \times 10^{-10}$	83 000 (20 750)
e	8	$(5 \pm 1) \times 10^{-11}$	380 000 (47 500)

The analysis of these systems is anyway far from easy. When a multivalent ligand shows enhanced activity compared to its monomeric counterpart, it is difficult to prove if this affinity gain is due to an increase of the intrinsic affinity or if it is caused by the combination of other effects like aggregation or precipitation. Analytical ultracentrifugation (SV-AUC) and dynamic light scattering (DLS) of the dendrimers described above in solution with LTB, demonstrated that ligands not matching the valency of the receptor protein can induce intermolecular aggregation, with formation of space-filling networks and precipitates[31]. Although the subtle differences in structure, LTB and CTB have identical GM1 binding sites[32], therefore are often used interchangeably when testing inhibitors. In particular, it was found that the divalent dendrimer **c** originates 2:5 toxin-ligand aggregates, while the tetravalent derivative **d** presents 2:3 stoichiometry when complexed to the protein (figure **5.9**).

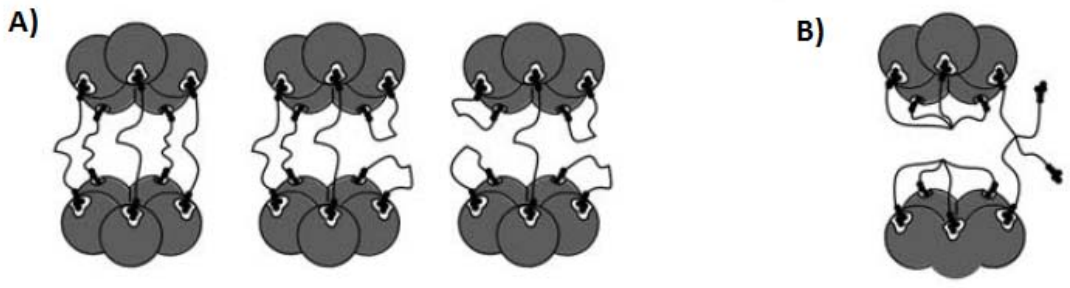


Figure 5.9 Cases of non-matching between the valencies of the protein and of the ligand: A) Three possible structures for the 2:5 (pentavalent toxin:divalent ligand) complex. B) Example of a possible structure for the 2:3 (pentavalent toxin:tetravalent ligand) complex. Image adapted from ref. 31.

It was suggested that, in certain cases, the formation of such aggregates could be responsible for the drastic inhibition enhancement of these multivalent ligands observed by ELISA tests. Affinity measurements made via isothermal titration calorimetry techniques, instead, showed no relevant increase of the intrinsic binding affinity. The formation of these aggregates was demonstrated to take place at μM concentration, but there are no evidences so far that the same aggregation mechanisms are present also at nM concentration, range where the dendritic inhibitors were found effective in the ELISA tests.

Fan and coworkers[33,34] synthesized a series of pentavalent CT inhibitors in which a pentavalent core is conjugated to five linkers, each one ending with a galactose unit (figure **5.10**).

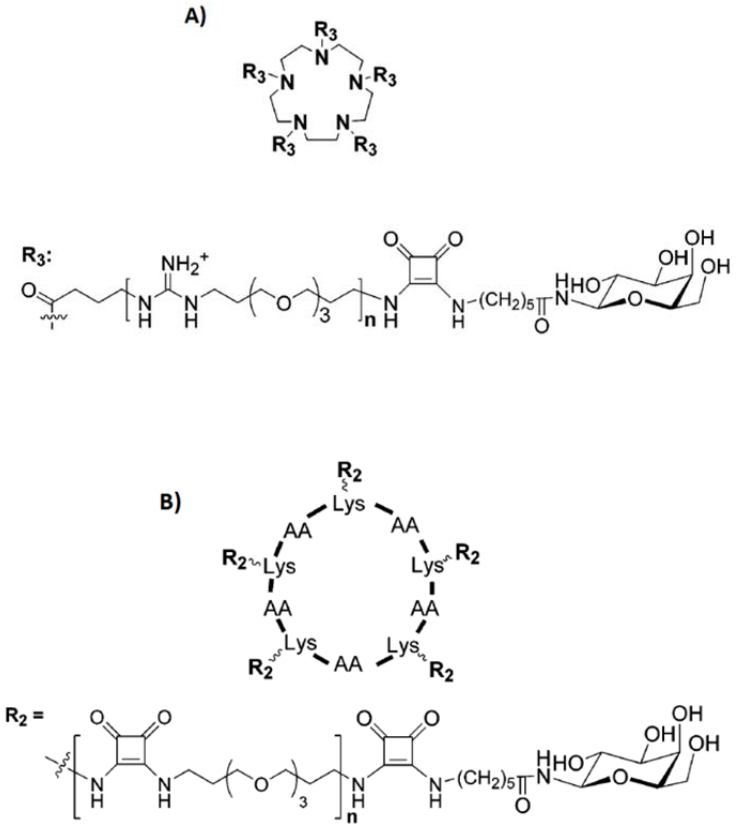


Figure 5.10 Pentavalent galactose inhibitors for CTB based on a pentacyclen core (A) and on a cyclic decapeptide scaffold (B).

Even if a single monomeric galactose unit presents very low CTB affinity, ELISA-type assays showed significant increase in binding affinity for the multivalent system. For instance, the

pentacyclen derivative showed to be 260-fold more potent compared to the monomeric compound. Through DLS studies it was possible to exclude the formation of aggregates also at much higher inhibitor concentrations than that used in the ELISA tests. The formation of a 1:1 protein-ligand complex was also confirmed thanks to X-ray structures. In this case it was possible to establish that the enhancement in the binding affinity was due to this multivalent complex and not to other factors such as aggregation or precipitation as observed in the dendrimer case, when there was a valency mismatch between ligand and receptor.

At this point it is possible to establish what are the characteristics that a good CTB inhibitor should present: a pentavalent core able to connect through linkers of proper length five oligosaccharides having high affinity with the B subunit (possibly GM1os). In this way it should be possible to obtain a very potent and efficient inhibitor presenting the same valency of the cholera toxin. This would hopefully also allow to point out the role of multivalency in the binding process.

5.2 Results and discussion

5.2.1 Synthesis of the pentavalent glycocalix[5]arenes

The aim of the work reported in this chapter is the synthesis of new pentavalent ligands based on a calix[5]arene scaffold able to inhibit with high affinity and efficiency the cholera toxin. Four different glycocalix[5]arenes **1**-**4** (figure **5.11**), functionalized with five units of galactose, lactose, and oligosaccharides of GM2 (GM2os) and GM1 (GM1os) gangliosides, respectively, are described. This work was performed in collaboration with Prof. Han Zuilhof and his research group (Department of Organic Chemistry of the University of Wageningen, NL), which has also provided the GM1os and GM2os derivatives obtained by chemo-enzymatic synthesis[35].

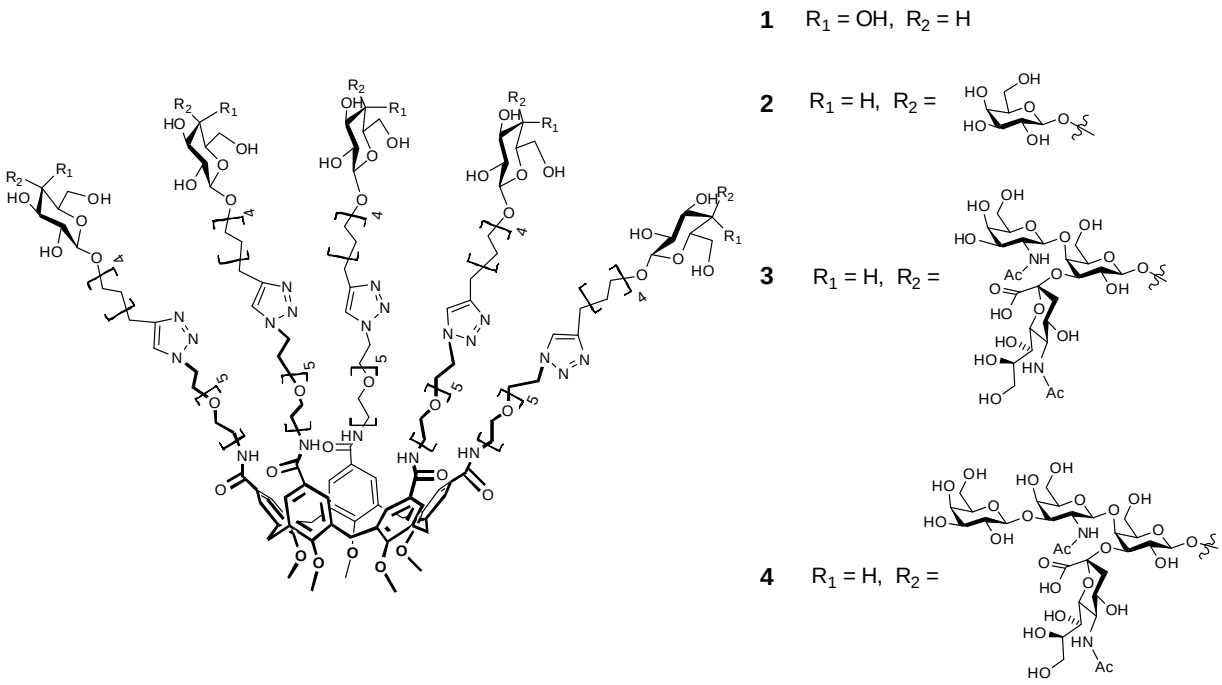


Figure 5.11 Target molecules **1**-**4** containing five units of galactose, lactose and the oligosaccharides of gangliosides GM2 and GM1, respectively.

These compounds all present a calix[5]arene core with methoxy groups at the lower rim. These short chains confer high conformational mobility to the macrocyclic structure. In fact, the methoxy groups are endowed with a reduced steric hindrance and, rapidly passing through the calixarene annulus, give rise to a conformational interconversion process that is very fast even on NMR time scale. In this way the purification and characterization steps are consistently simplified, since the presence of conformer mixtures, very difficult to separate and analyze, is avoided. Besides, the molecule can better adapt itself to the CT structure during the binding process, possibly even correcting mismatched binding. On the other hand, this high mobility leads to a low degree of preorganization of the macrocycle, that is possibly unfavorably contributing to the entropic term and hence to the total free energy ΔG of binding to CT. However, at least for the time being, there is no possibility to synthesize a calix[5]arene based CT inhibitor fixed in a more preorganized cone structure, since no procedures are known in the chemistry of this pentameric metacyclophane to modify the upper rim substituents and simultaneously controlling its conformation[36].

The upper rim of the calixarene inhibitors is functionalized with carbohydrates separated from the macrocyclic core by appropriate linkers (figure **5.12**).

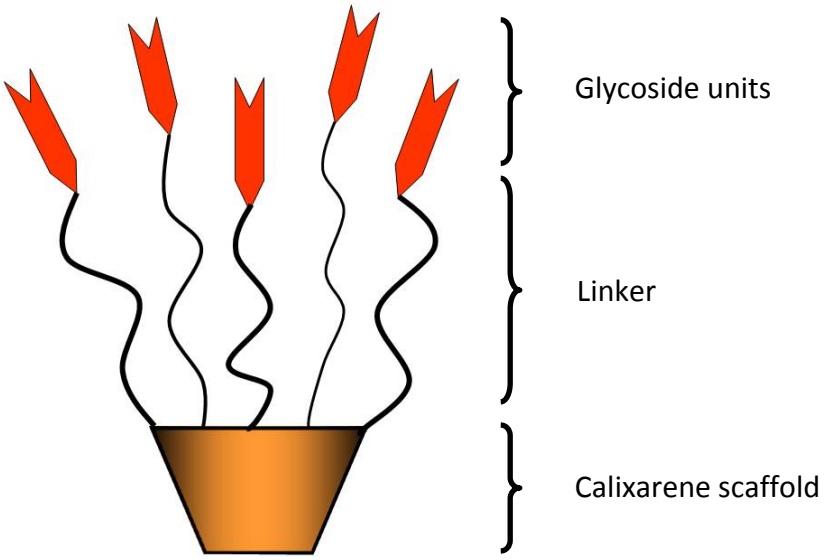


Figure 5.12 Schematic view of the calixarene based CT inhibitors.

Fan and coworkers[34] have demonstrated that it is fundamental to determine the optimal linker length in a synthetic multivalent inhibitor because it is strictly connected to its potency. If the linkers are too short it is impossible for the glycoside groups of the inhibitor to simultaneously reach all the five binding site of the CT B-pentamer. On the other hand, if the linkers are too long, during complexation there would be an unnecessary loss of conformational freedom with a subsequent unfavorable entropic contribution to the binding process. In any case, both situations would lead to a lower binding affinity between inhibitor and CT. For the calixarenes described in this thesis a 30 atom-containing linker was chosen since, according to preliminary modeling data, it should allow the simultaneous interaction of the five carbohydrate units with the B-pentameric subunit of a single toxin.

To obtain these compounds a convergent synthetic approach was used (figure **5.13**). The syntheses of the calixarene scaffold, of the ethylene oxide linker and of the C11-functionalized carbohydrates were carried out separately. The various residues were then connected together. First the calixarene scaffold **10** was reacted with the amino-azide ethylene oxide linker **13** resulting in the azide terminating calix[5]arene **15**, then the alkyne-terminating carbohydrate moieties **16**-**19** were connected to the calixarene via a copper catalyzed “click” cycloaddition.

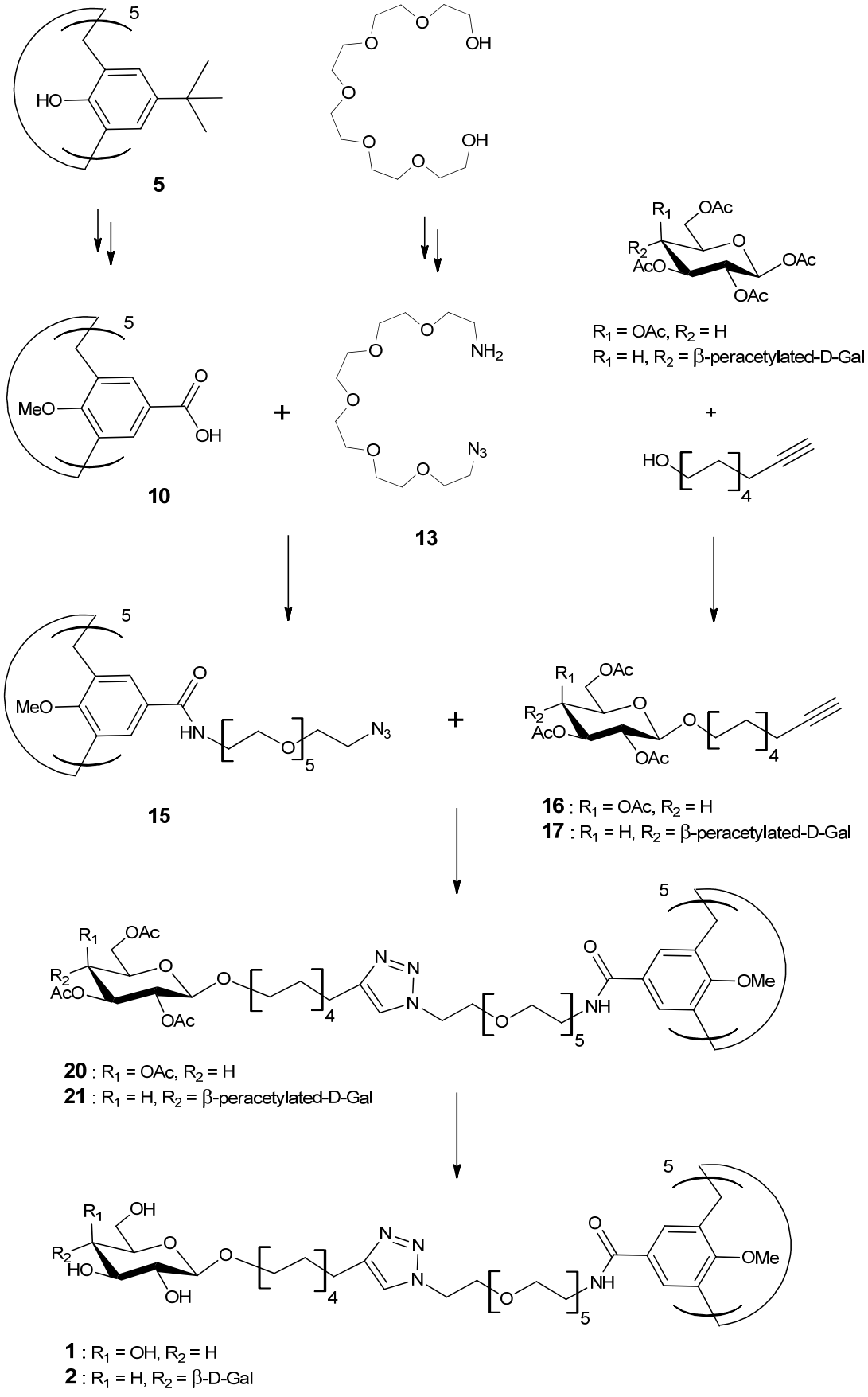


Figure 5.13 Convergent approach leading to galactose and lactose-functionlized calix[5]arene **1** and **2**.

Although galactose and lactose bind to CT much less efficiently compared to the GM2os and GM1os, we planned to first synthesize inhibitors based on these simpler mono- and di-saccharides. These carbohydrates are, in fact, cheaper and available in relatively high amounts, thus allowing their use as simple model compounds in the fine tuning of the coupling reaction conditions. The reaction conditions developed for compounds **1** and **2**, with respect to the copper-catalyzed azide-alkyne cycloaddition, were therefore also applied for the synthesis of compounds **3** and **4**.

5.2.1.1 Synthesis of the calix[5]arene core

The synthesis of the calixarene core started with a one-pot direct condensation between p-*tert*-butyl-phenol and paraformaldehyde in presence of KOH to give the p-*tert*-butyl-calix[5]arene **5** (figure **5.14**). While even numbered calix[n]arenes (n = 4, 6, 8) can be obtained with good yields, especially from phenols bearing bulky substituents like *tert*-butyl[37,38], benzyloxy[39], and adamantyl[40] at para-position, odd numbered calix[n]arenes (n = 5, 7, 9) are obtained in considerably lower yields. The followed procedure[41] to obtain the starting calix[5]arene **5** is quite elaborated and gave only a 9% yield.

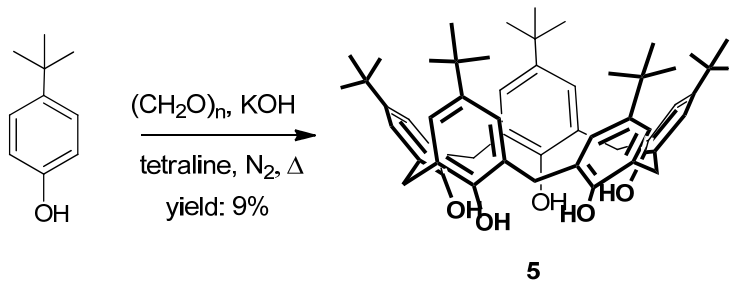


Figure 5.14 Synthesis of para-*tert*-butylcalix[5]arene **5**.

De-*tert*-butylated calix[5]arene **6** was easily obtained according to standard procedures[42] by reaction of *tert*-butyl-calix[5]arene **5** with $AlCl_3$ in toluene (figure **5.15**).

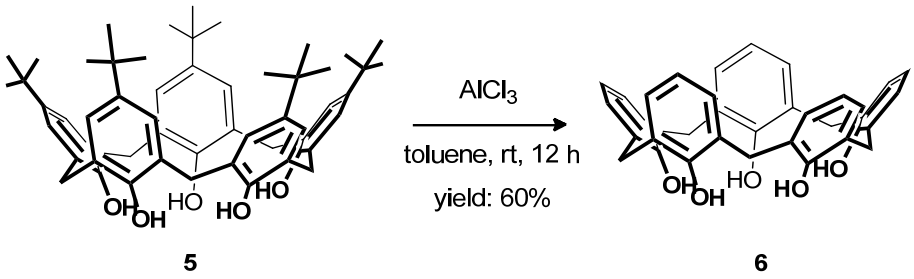


Figure 5.15 Synthesis of calix[5]arene **6**.

At this point there could be two possible ways to prepare the penta-methoxy penta-acid derivative **10** (figure **5.16**).

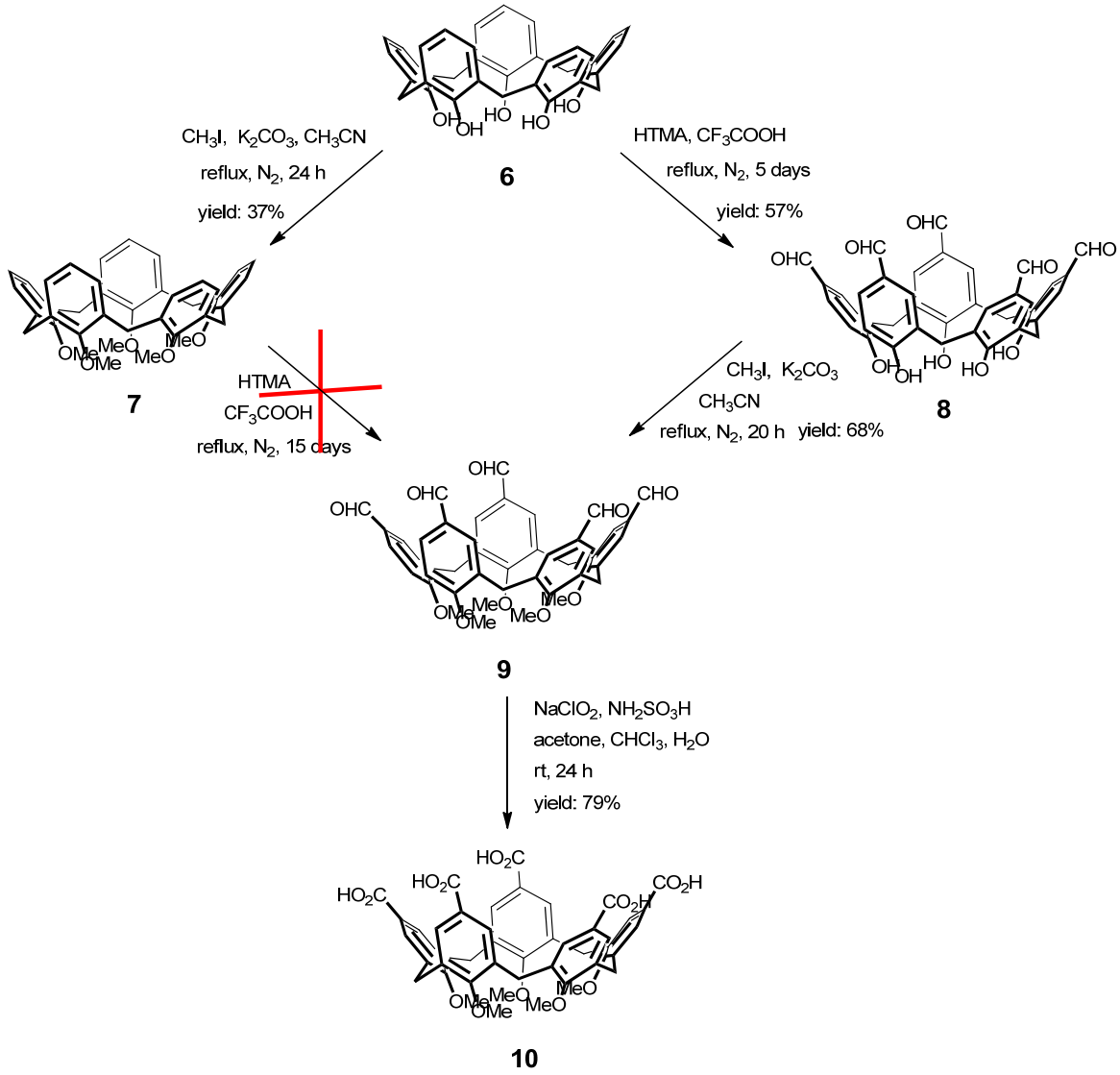


Figure 5.16 Synthesis pathways to obtain calix[5]arene **10**.

One method consists first in functionalizing the lower rim with methyl groups and then the upper rim with formyl groups. The second pathway works the other way around, with formylation at the upper rim followed by methylation of the lower rim. Both ways were investigated to determine most convenient synthetic procedure.

Intermediate **7** was easily obtained by reaction of calix[5]arene **6** with iodomethane in presence of K_2CO_3 as base. The not so high yield is compensated by the relatively easy access to the pure product obtained after two trituration, first in methanol and subsequently in hexane. The following step, the formylation reaction at the upper rim of compound **7**, was carried out with hexamethylentetraamine (HMTA) in trifluoacetic acid (TFA) exploiting the Duff reaction[43]. The reaction showed to be rather slow and only a mixture of products, possibly of partially formylated derivatives, was obtained, together with a small amount of pentaformyl calix[5]arene (the reaction was followed via TLC and ESI-MS analyses). Even forcing the reaction conditions by drastically raising the temperature to 130 °C, the reaction time to 15 days and the number of equivalents of HMTA to 40, did not yield the fully formylated compound.

Due to these synthetic problems, this route was abandoned and we tried to get compound **9** by reversing the order of the reactions. Compound **8** was then obtained completely formylated by reaction of calix[5]arene **6** with only 2 equivalents of HMTA per phenolic unit in TFA after an overall reaction time of 5 days at 100 °C. As most of the lower rim unsubstituted calixarenes, also compound **8** shows a limited solubility in organic solvents that caused some problems in the purification steps. Anyway a sufficiently pure compound could be obtained by trituration in a 1:1 chloroform:hexane mixture. Subsequently compound **8** was methylated at the lower rim by using CH_3I and K_2CO_3 in acetonitrile and the pentamethoxycalix[5]arene **9** was obtained in 68% yield. Oxidation of **9** with $NaClO_2$ and NH_2SO_3H produced the penta-acid **10** that was unambiguously identified by ESI-MS and NMR techniques. In particular, it is interesting to observe in the ^1H NMR spectrum (300 MHz, CD_3OD) the presence of a singlet at 3.93 ppm for the protons of the methylene bridge. This confirms that calixarene **10** undergoes fast conformational interconversion (on NMR time-scale), and results to be conformationally mobile (figure **5.17**).

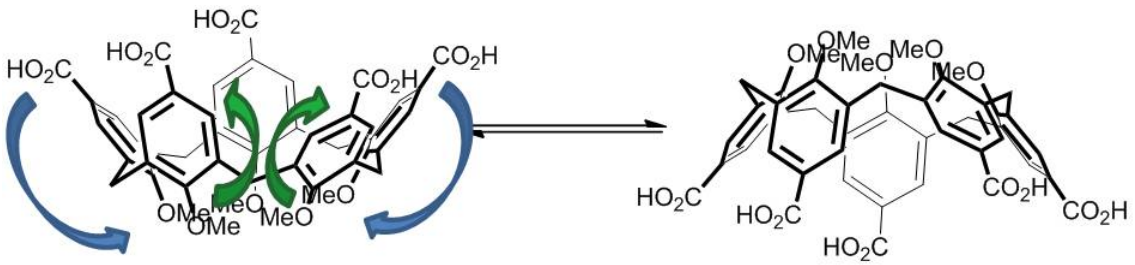


Figure 5.17 The fast passage of the small methoxy groups through the calix[5]arene annulus of compound **10** leads to a rapid interconversion between different conformations.

Carboxylic groups were chosen because they can be used to easily form stable amide bonds by reaction with amine-functionalized compounds in presence of various coupling reagents[44]. Therefore the linkage of the amine spacers to the calixarene scaffold should be carried out in a pretty straightforward manner.

5.2.1.2 Synthesis of the hexaoxyethylene linker

The GM2 and GM1 oligosaccharides provided by prof. Zuilhof's group were already functionalized at the glucose anomeric position with an undec-10-ynyl chain (see below figure **5.20**), therefore it was necessary to synthesize a 20 atom-containing linker to connect the calixarene scaffold to these alkynyl glycosides. In this way an overall 30 atom chain between the core and the sugars could be obtained, thus satisfying the length requirements defined for a proper multivalent ligand-CT binding. To achieve this goal, we planned to synthesize compound **13** (figure **5.18**), an hexa-ethylene-glycol derivative containing an amine function at one end of the chain and an azido group at the other end. In this way, with the aid of the coupling reagents typical of the peptide synthesis[44], the amine group should easily react with the carboxylic acid of the calixarene core through the formation of an amide bond. On the other hand, the azido group is required to carry out the copper(I) catalyzed alkyne-azide cycloaddition (CuAAC). Moreover, the presence of polyoxyethylene chains should enhance the hydrophilic character of the final compounds. This is a matter of primary importance since all the inhibitors synthesized need to be water-soluble, because all the inhibition experiments are performed in aqueous media.

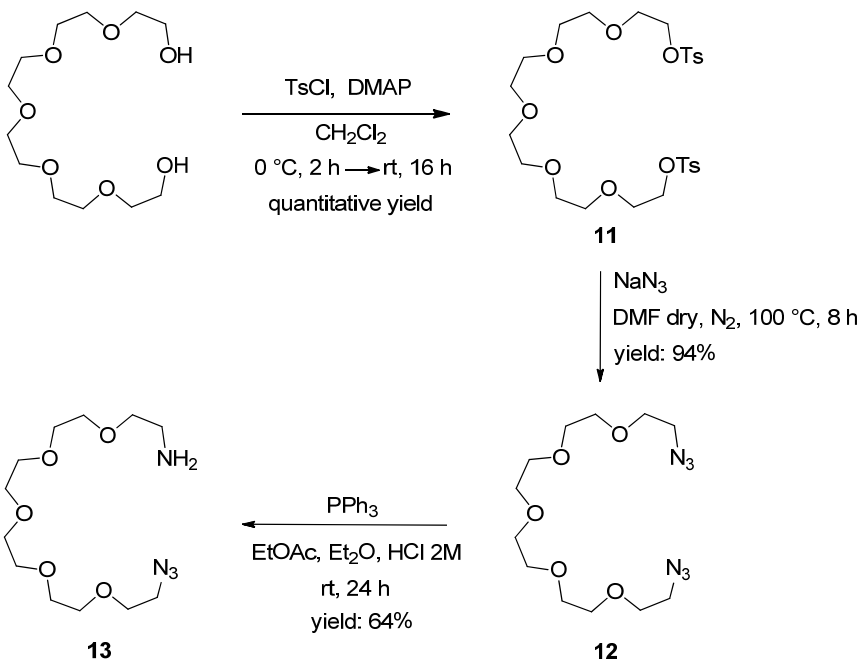


Figure 5.18 Synthetic strategy to obtain the linker **13**.

To obtain the unsymmetrically substituted linker **13**, the first step was the reaction between the hexaethylene glycol and *p*-toluenesulfonyl chloride[45] to give, in nearly quantitative yield, the ditosylated product **11** that was subsequently reacted with NaN_3 to obtain the di-azido compound **12**. To selectively reduce only one of the azido group to amine, it was decided to exploit a desymmetrization reaction where triphenylphosphine is used as reducing agent in a biphasic solvent system (EtOAc-Et_2O/HCl 2M)[46]. As soon as the mono amino product **13** is formed, it becomes protonated and quantitatively diffuses to the acidic aqueous phase. Because of its high polarity, this compound cannot return to the organic phase, avoiding therefore the reduction of the remaining azido group. This reaction is highly selective and no diamino byproduct was formed. The presence of even small amounts of diamino derivatives are highly detrimental since they could lead to oligomeric and/or cross-linked products during the conjugation to the calixarene pentaacid. Besides, the lack of some of the azido group would prevent the "click" reaction to take place, leading to inhibitors with a diminished valency (lower than five). Another interesting advantage of this selective reduction is that product **13** diffuses into the aqueous phase as soon as it is formed, leaving all the excess of reagents in the organic phase. Therefore its purification is extremely simple

(with only extraction steps) allowing to skip the purification also of compounds **11** and **12** and achieving rather good overall yield for **13** (60% starting from the hexaethyleneglycol).

5.2.1.3 Coupling between the hexaethylene glycol linker and the calixarene core

The subsequent step was the coupling between the carboxylic groups of the calix[5]arene **10** and the amine group of the linker **13**, resulting in the formation of amide bonds (figure **5.19**). For this reaction, it was decided to optimize the conditions by trying two different methodologies.

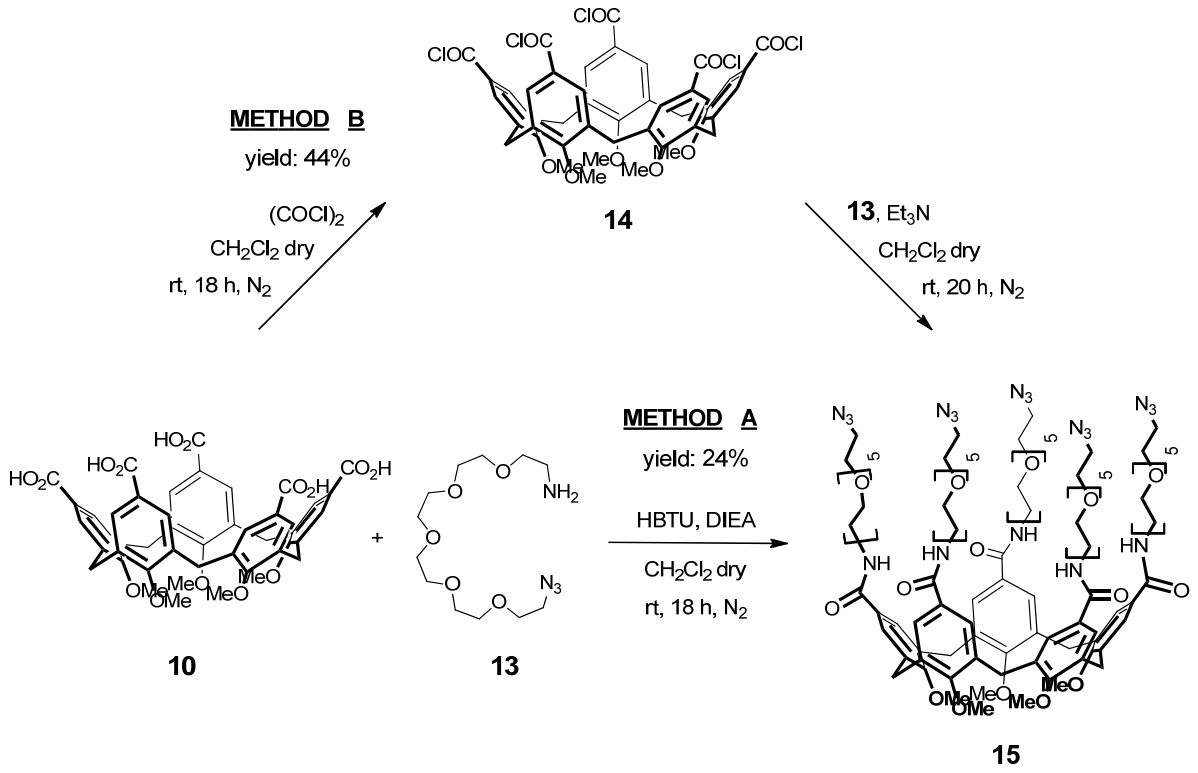


Figure 5.19 Coupling reaction between the pentacid calix[5]arene **10** and linker **13** to give the pentazido compound **15**.

In the first route (method A in figure **5.19**), based on the analogue reaction that was previously successfully exploited on a dicarboxylic acid calix[4]arene[29], HBTU was used as

coupling agent. Even if product **15** was obtained, the purification from HBTU degradation products was quite tricky and time consuming. Multiple separations on flash column were required to obtain the desired pentazido compound **15** pure, with a consequent final low yield of only 24%. The alternative pathway (method B in figure **5.19**) consisted in the activation of the carboxy groups of **10** with oxalyl chloride to form the acyl chloride derivative **14** followed by the reaction with the amino derivative **13** to give the pentazido calix[5]arene **15** substantially in absence of byproducts.

Both routes afforded the penta-functionalization of the starting compound **10**, but method B resulted to be considerably better than method A, not only for the lower amount of impurities formed, but also because the yield resulted to be almost the double (i.e. 44%).

5.2.1.4 Synthesis of 1-β-(undec-10-ynyl)-glycoside derivatives

The synthesis of the 1-β-(undec-10-ynyl)-glycosides **16** and **17** was the following step in the route towards the pentavalent calix[5]arenes **1** and **2**. Compounds **16** and **17** contain a β-galactose and a β-lactose residue, respectively, both present in the parent ganglioside GM1os structure (figure **5.20**). To avoid undesired negative biological behavior in CT binding, arising from using α-anomers not present in the natural GM1 ligands, it is of the utmost importance to obtain anomerically β-pure compounds. β-Galactose and β-lactose were chosen as highly simplified GM1os analogues to optimize the synthetic procedures to be employed in the "click" reactions with compounds **18** and **19** (presenting the GM2 and GM1 oligosaccharide, respectively) which are considerably more challenging and costly in use. Besides, once the galactose and lactose residues are connected to the calixarene backbone, the resulting products **1** and **2** can also be used as inhibitors in the ELISA tests. In fact, also these simple saccharides can take advantage of a multivalent presentation, although characterized by low affinity to CT as monomeric ligands. Anyway, like in GM1, their terminal moiety is a β-galactoside unit, which is the carbohydrate residue that gives the strongest interactions with CT (see paragraph **5.1.3**) and therefore they can be considered as GM1 mimics. In derivatives **16** and **17** the OH groups were protected with acetyl groups, to avoid the free hydroxyl functionalities to interfere in the syntheses and complicate the purification steps. On the contrary compounds **18** and **19** were obtained through chemo-enzymatic

reactions[47] by prof. Zuilhof's group in Wageningen, and therefore they were unprotected. To reduce the number of synthetic steps and avoid loss of valuable material, they were used as such, without protection of the free hydroxyl and carboxyl groups in the following "click" reactions.

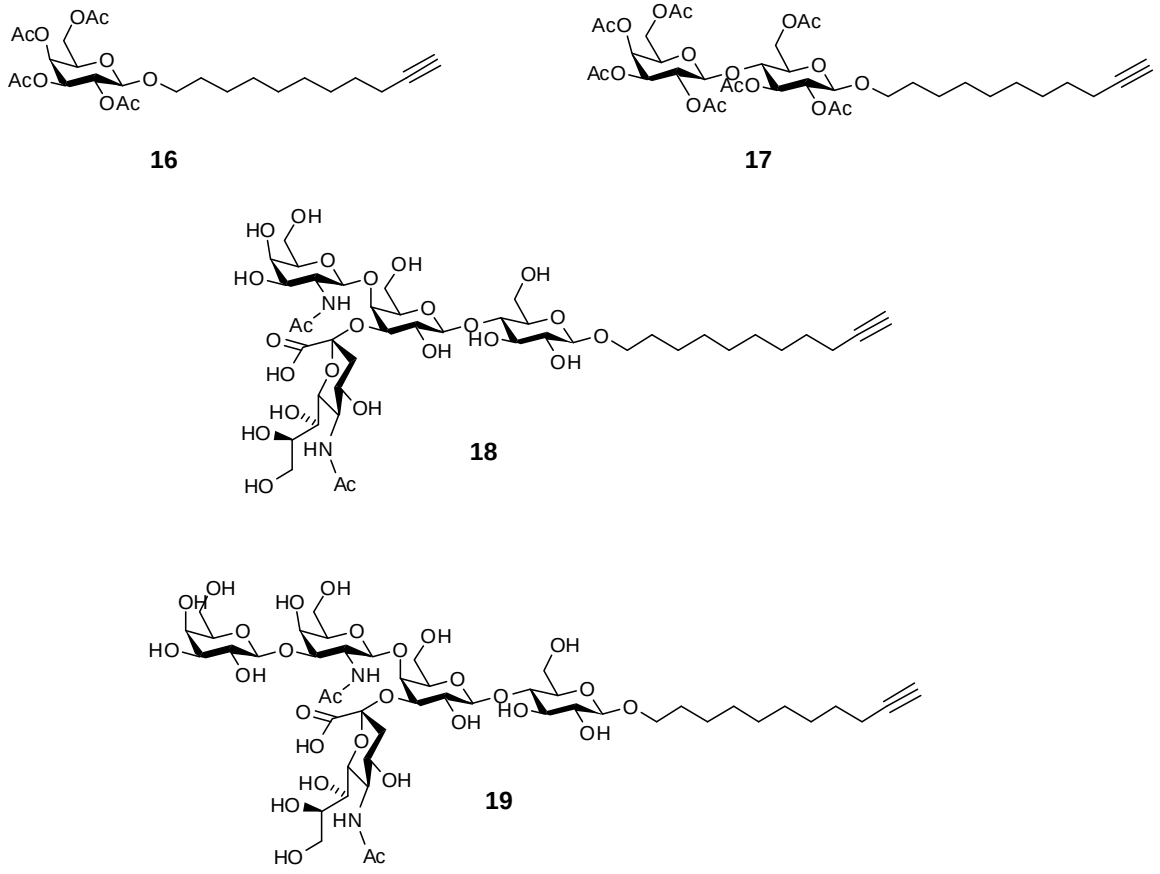


Figure 5.20 1-β-(Undec-10-ynyl)-glycosides derivatives **16**, **17**, **18**, **19** used for the synthesis of the CT inhibitors.

Octa-*O*-acetyl-β-D-lactoside, prepared by acetylation of D-lactose, was functionalized with commercially available 10-undecyn-1-ol (figure **5.21**) according to literature procedure[35,47].

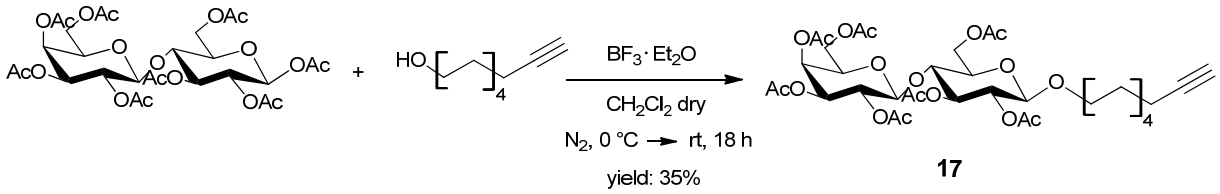


Figure 5.21 Synthesis of the lactose derivative **17**.

The trifluoroborate etherate was used as Lewis acid in this reaction, giving a mixture of β- and α-glycosylated products in a 4:1 ratio. To minimize the formation of the α-anomer, the temperature was initially kept at 0 °C and then slowly allowed to reach room temperature to speed up the reaction. The separation of the two anomers represented one of the main difficulties in these glycosylation reactions. The β-anomer **17** was isolated pure after three consecutive flash column chromatography purifications, with consequent loss of material and a final yield of only 35%.

^{1}H NMR spectrum confirmed the success of the reaction and separation steps. In particular the anomeric proton close to the undecynyl group (identified under the multiplet between 4.48 and 4.40 ppm) is very diagnostic because it presents a coupling constant of about 8 Hz, typical for a β-lactoside and substantially different from that of the α-anomer, whose coupling constant is about 3 Hz and whose position is downfield shifted.

Analogously, compound **16** was synthesized from peracetylated galactose and 10-undecyn-1-ol[35,47]. Also in this case it was possible to obtained the anomerically pure β-compound by flash chromatography. The success of the reaction and purification was confirmed by NMR spectra.

5.2.1.5 "Click" reactions

Conjugation of penta-azido calix[5]arene **15** to the alkyne-terminating glycosides **16**, **17**, **18** and **19** was performed via Copper-catalyzed Azido-Alkyne Cycloaddiction (CuAAC) reaction.

The coupling reactions were performed in a rather similar way for compounds **16** and **17** (galactose and lactose derivatives) on one side, both presenting acetylated carbohydrate units, and for compounds **18** and **19** (GM2os and GM1os derivatives) on the other side, having all free hydroxyl groups. The different solubility of the protected and unprotected alkyne-terminating glycosides suggested in fact the use of different solvents for the CuAAC reactions.

The "click" reaction between the azido groups of calix[5]arene **15** and the C11-alkyne chain of **16** and **17** was performed in DMF and H_2O under microwave irradiation (figure **5.22**).

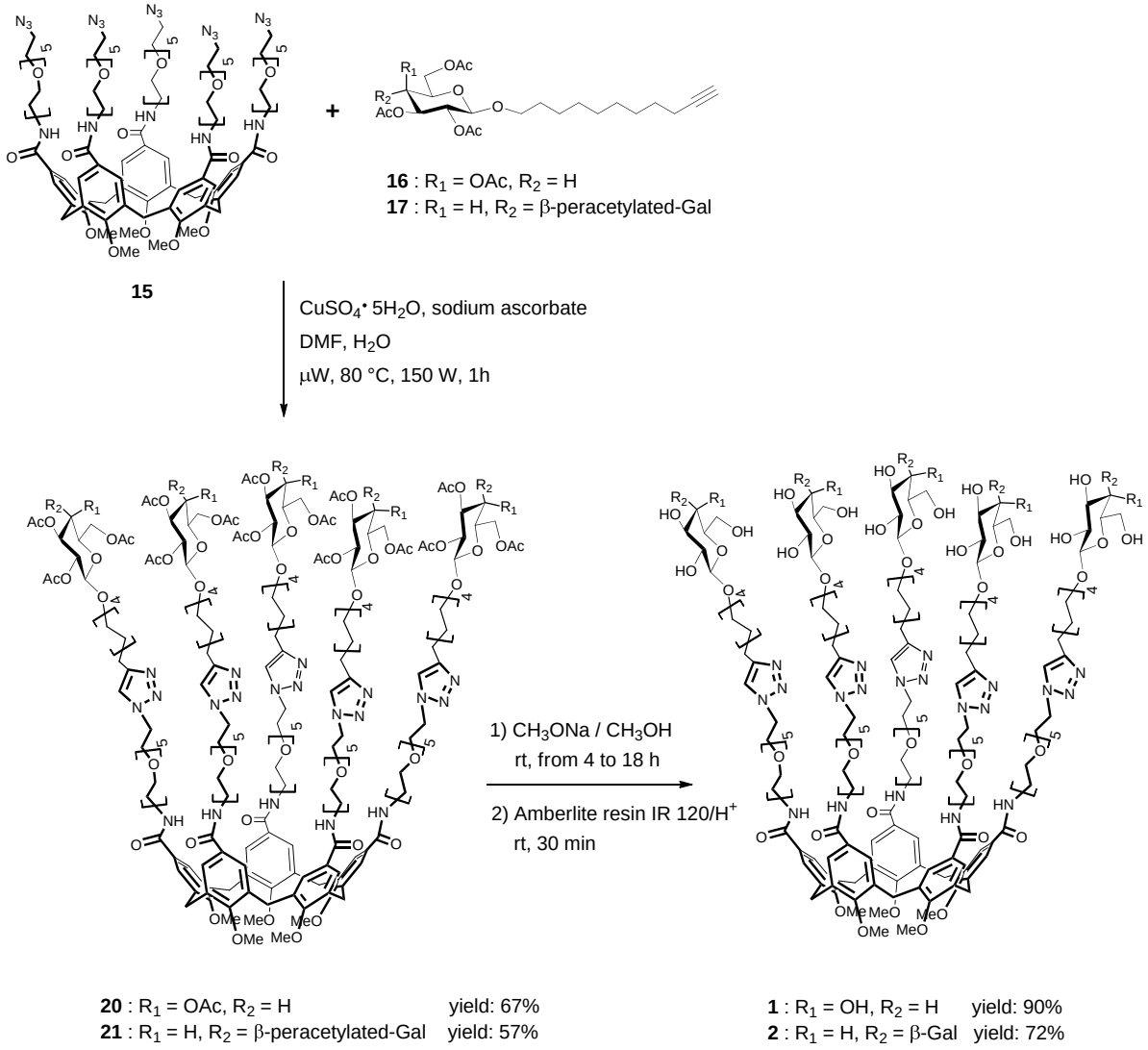


Figure 5.22 Synthesis of pentavalent calix[5]arenes **1** and **2** via CuAAC reaction and subsequent carbohydrates deprotection.

One hour reaction was sufficient to get the pentafunctionalized compounds **20** and **21** in high yields (57% and 67%, respectively) and no evidence of the presence of partially functionalized intermediates was found. Purification was necessary just to remove the excess of glycoside reagents. NMR analyses confirmed the identities of the two products. The ^{1}H NMR spectrum of galactosyl-calix[5]arene **20** (figure **5.23**) presents a signal at 7.77 ppm for the proton of the triazole ring and a peak at 7.61 ppm for the aromatic protons which are particularly diagnostic: the ratio of the corresponding integrals is exactly 1:2, which confirms the complete functionalization of **15** with all its azido functionalites reacted with the galactoside unites **16**. The macrocycle showed to be conformationally mobile even

at room temperature, as the NMR signals of the methylene bridge ($ArCH_2Ar$) suggest. In the ^1H NMR spectrum a singlet at 3.94 ppm can be observed, and in the ^{13}C NMR spectrum the corresponding peak is at 30.6 ppm. Both are typical values for calixarenes in a mobile conformation[48].

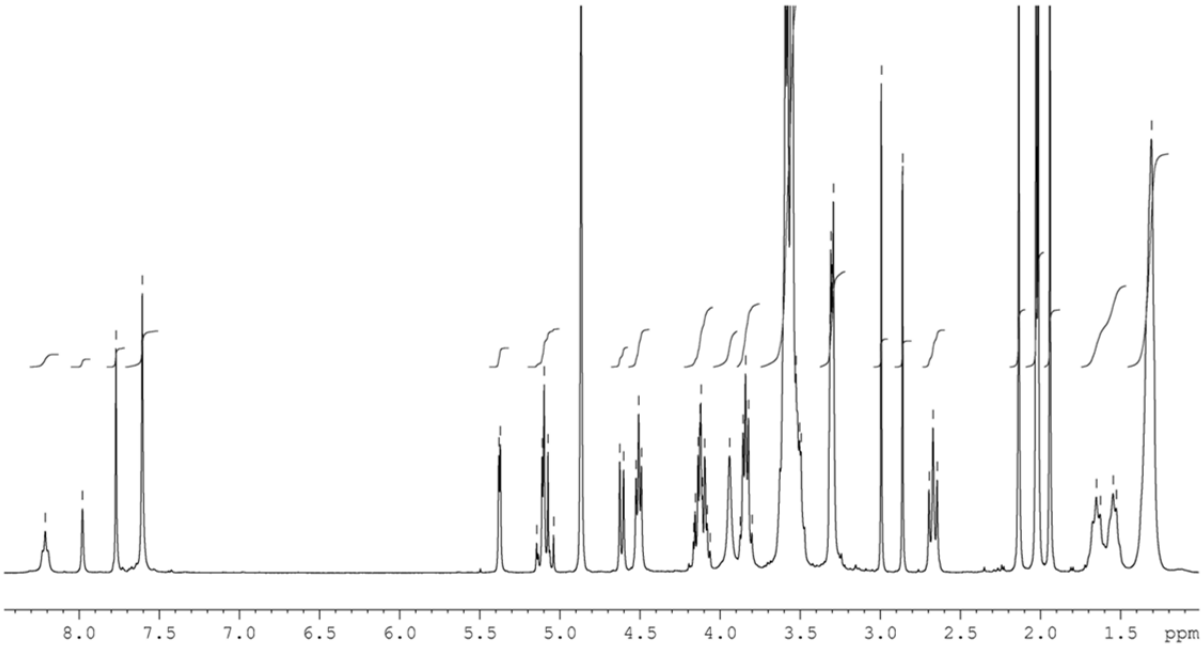


Figure 5.23 ^1H NMR spectrum (300 MHz, CD_3OD) of the peracetylated galactosyl-calix[5]arene **20**.

For compound **21** analogous observations can be made. Integrals of proton signals in ^1H NMR spectrum (figure **5.24**) confirm the penta-functionalization of the calixarene derivative, while chemical shifts of the peak relative to the methylene bridge in ^1H NMR and ^{13}C NMR spectra show that the obtained calixarene presents a mobile conformation.

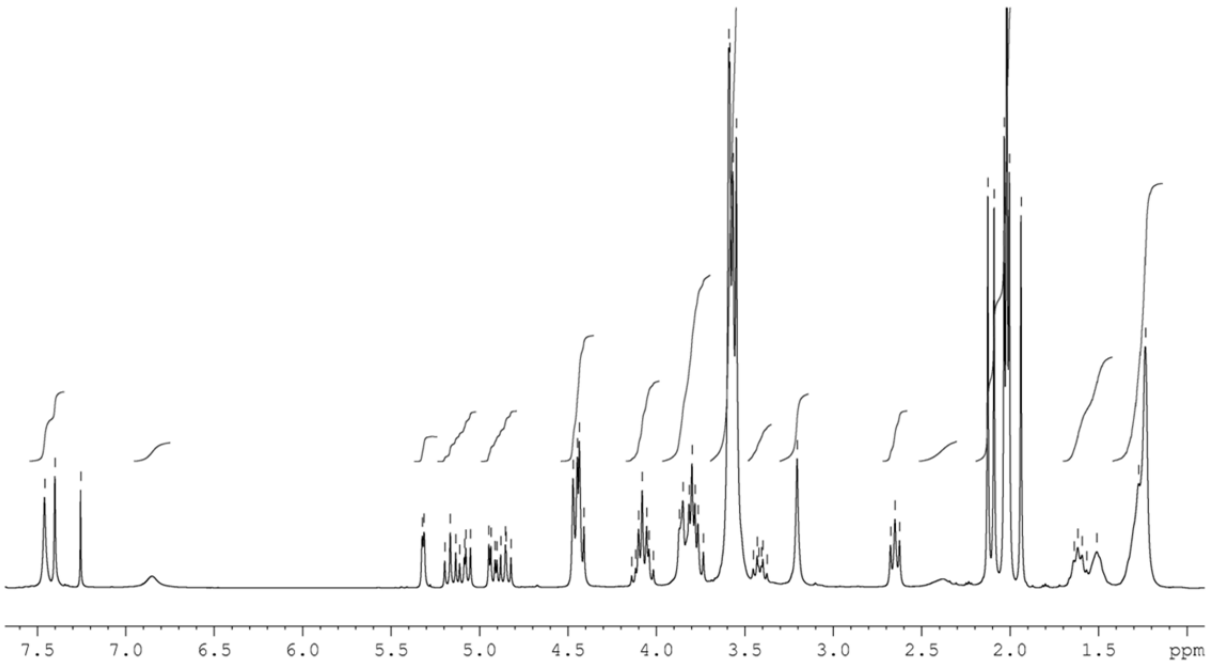


Figure 5.24 ^{1}H NMR spectrum (300 MHz, $CDCl_3$) of the peracetylated lactosyl-calix[5]arene **21**.

The subsequent step was the deprotection of the peracetylated galactosyl and lactosyl calix[5]arenes **20** and **21** to give the final products **1** and **2**, respectively (figure **5.22**). The removal of the acetyl groups was carried out at room temperature under basic conditions, by addition of sodium methoxide to methanol according to the standard Zemplén procedure[49]. Complete deprotection could be achieved in 4–18 hours. Products **1** and **2** were completely characterized by ESI-MS and NMR techniques, with all the peaks assigned also with the aid of 2D NMR experiments. In the ^{1}H NMR spectra (figure **5.25** for **1** and **5.26** for **2**) the complete absence of the acetyl signals between 2.15 and 1.90 ppm shows that the deacetylation was successful. Confirmation of the complete deprotection of the molecules comes also from ^{13}C NMR spectra where the carbonyl and methyl signals of the acetyl groups (in the region between 171 – 169 ppm for the former and 21 – 19 for the latter) are also absent.

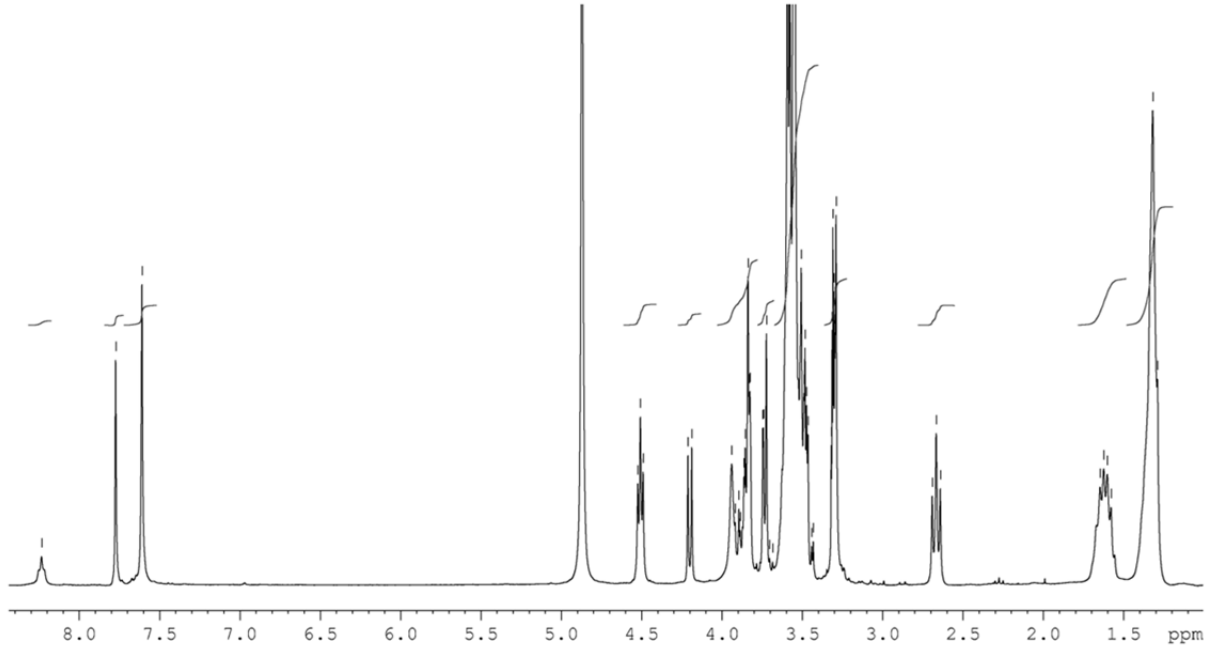


Figure 5.25 ^{1}H NMR spectrum (300 MHz, CD_3OD) of galactosyl-calix[5]arene **1**.

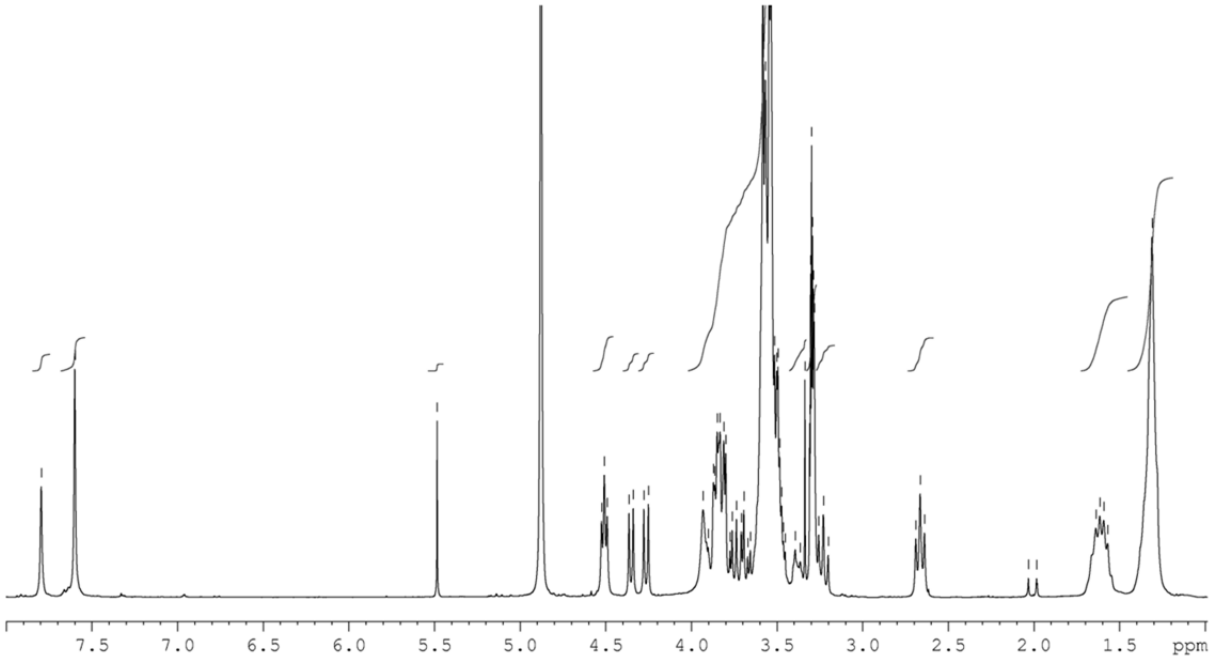


Figure 5.26 ^{1}H NMR spectrum (300 MHz, CD_3OD) of lactosyl-calix[5]arene **2**.

The synthesis of calix[5]arenes **3** and **4** was obtained in a single step from the penta-azido calix[5]arene **15** and undec-10-ynyl GM2 and GM1 oligosaccharides **18** and **19**, respectively

(figure **5.27**). CuAAC reactions and procedures analogues to those used for the synthesis of calix[5]arenes **20** and **21** were exploited.

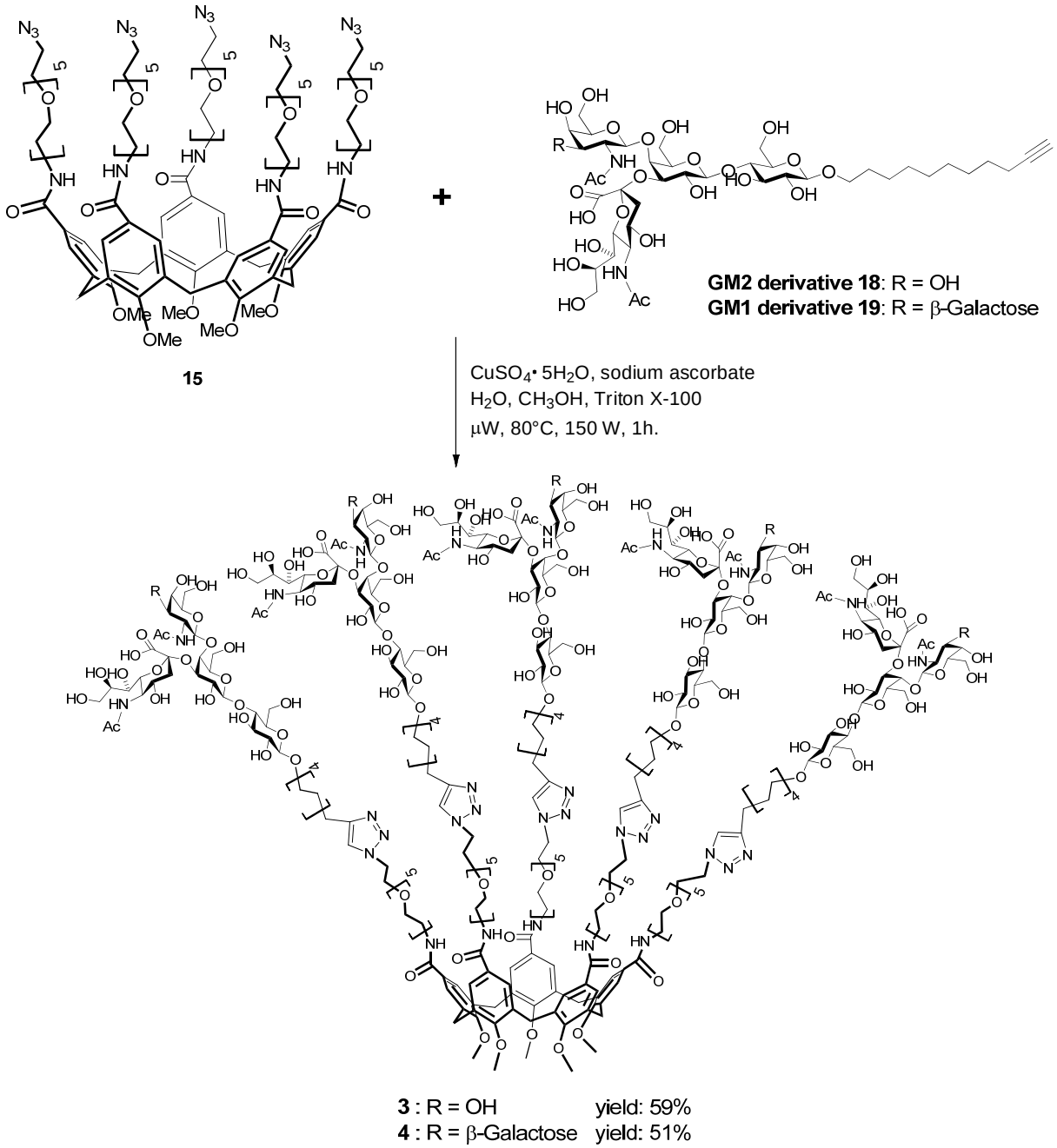


Figure 5.27 Synthesis of the pentavalent calix[5]arene **3** and **4** via CuAAC reaction.

GM1os and GM2os derivatives **18** and **19** are well soluble in water, show limited solubility in methanol and are insoluble in almost all less polar organic solvents. Thus, the CuAAC

reaction was carried out in water to allow most of the reactants to be well dissolved and therefore readily available for the reaction. To allow the solubilization also of the azide-terminating calix[5]arene **15**, it was necessary to solubilize first this compound in a small amount of methanol, and then add the other reagents. A drop of detergent (Triton X-100) was added to limit the formation of calixarene aggregates which would limit the accessibility of the azido reaction sites to the alkyne. The reaction was performed under microwave irradiation, giving the penta-glycosyl calixarenes **3** and **4** in 1 hour. The high polarity of the final compounds, due to the presence of the unprotected carbohydrate moieties, did not allow the purification from the excess of saccharides via standard direct phase flash chromatography. For this reason, size exclusion chromatography was performed (Sephadex G-15, eluent: 100% water) obtaining products **3** and **4** in 59% and 51% yields, respectively, pure enough to be used in the preliminary biological tests.

Products **3** and **4** were characterized by NMR techniques (figure **5.28** for GM2-calix[5]arene **3** and figure **5.29** for GM1-calix[5]arene **4**).

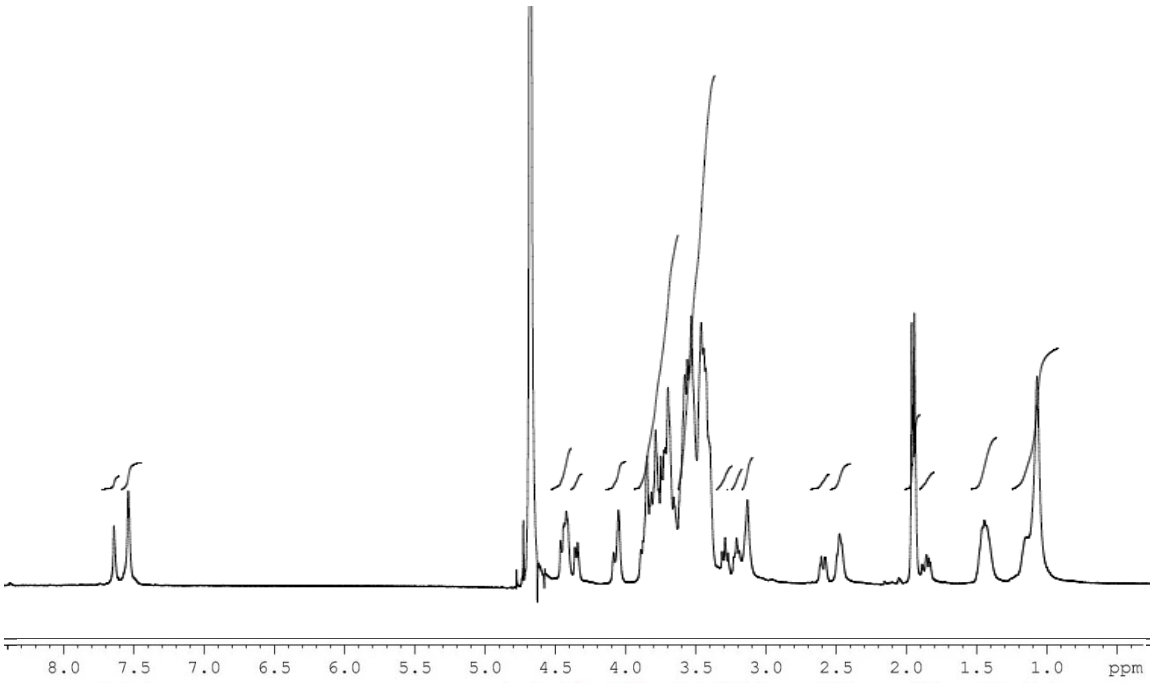


Figure 5.28 ^{1}H NMR spectrum (400 MHz, D_2O) of the GM2-calix[5]arene **3**.

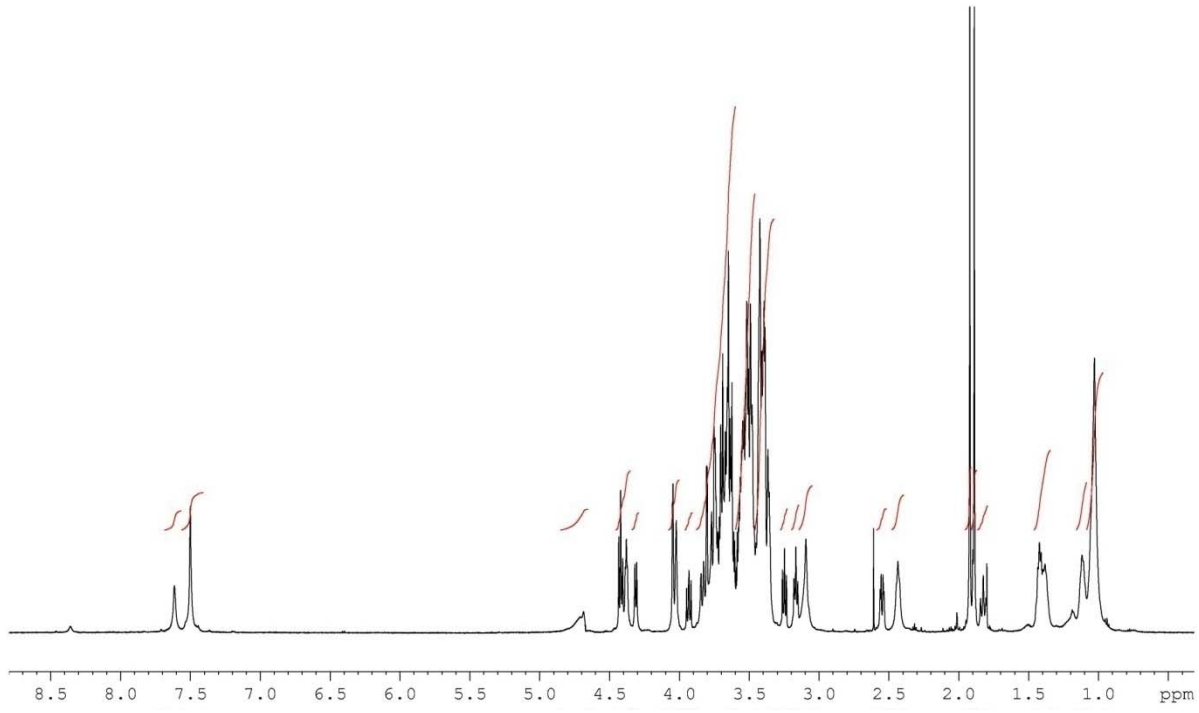


Figure 5.29 ^{1}H NMR spectrum (600 MHz, D_2O) of the GM1-calix[5]arene **4**.

The resulting spectra show to be rather complex, especially in the region between 3.0 and 4.5 ppm due to the overlapping resonances of the different carbohydrate moieties. Nevertheless, it is possible to observe that complete penta-functionalization has occurred for both molecules **3** and **4**. Very diagnostic also in this case are the integral ratios of the singlet for the triazole proton around 7.6 ppm and the singlet for the aromatic protons around 7.5 ppm, which are exactly 1:2. Also the singlets around 1.9 ppm for the acetyl protons of the sialic acid and of the N-acetyl-galactosamine, whose integrals correspond to 30 protons in total, confirm that all the azido groups reacted with the undec-10-ynyl GM oligosaccharides. HR-ESI-MS also confirmed that the CuAAC reactions were successful and that the desired products **3** and **4** were obtained.

5.2.2 Synthesis of monomeric inhibitors

Using the same procedures for the synthesis of the multivalent calixarene-based inhibitors, the synthesis of monomeric inhibitors based on a monomeric 4-methoxybenzoic acid structure was also designed. Starting from this acid it is possible to obtain the complete

series of monomeric analogues by simply changing the glycosylated reagent during the "click" reaction. So far only the lactose-based monomer **24** has been synthesized (figure **5.30**), but in the ongoing project it is planned to obtain also the galactose, GM2os and GM1os containing monomeric compounds by coupling the azido group of **22** with the alkyne group of the already synthesized compound **16**, **18**, **19**. These monomers were designed to be used as control compounds during cholera toxin inhibition tests to quantify the contribution of the glycoside cluster effect in the overall binding by comparing the IC_{50} values of a pentavalent inhibitor with the corresponding data obtained for its analogous monovalent molecule.

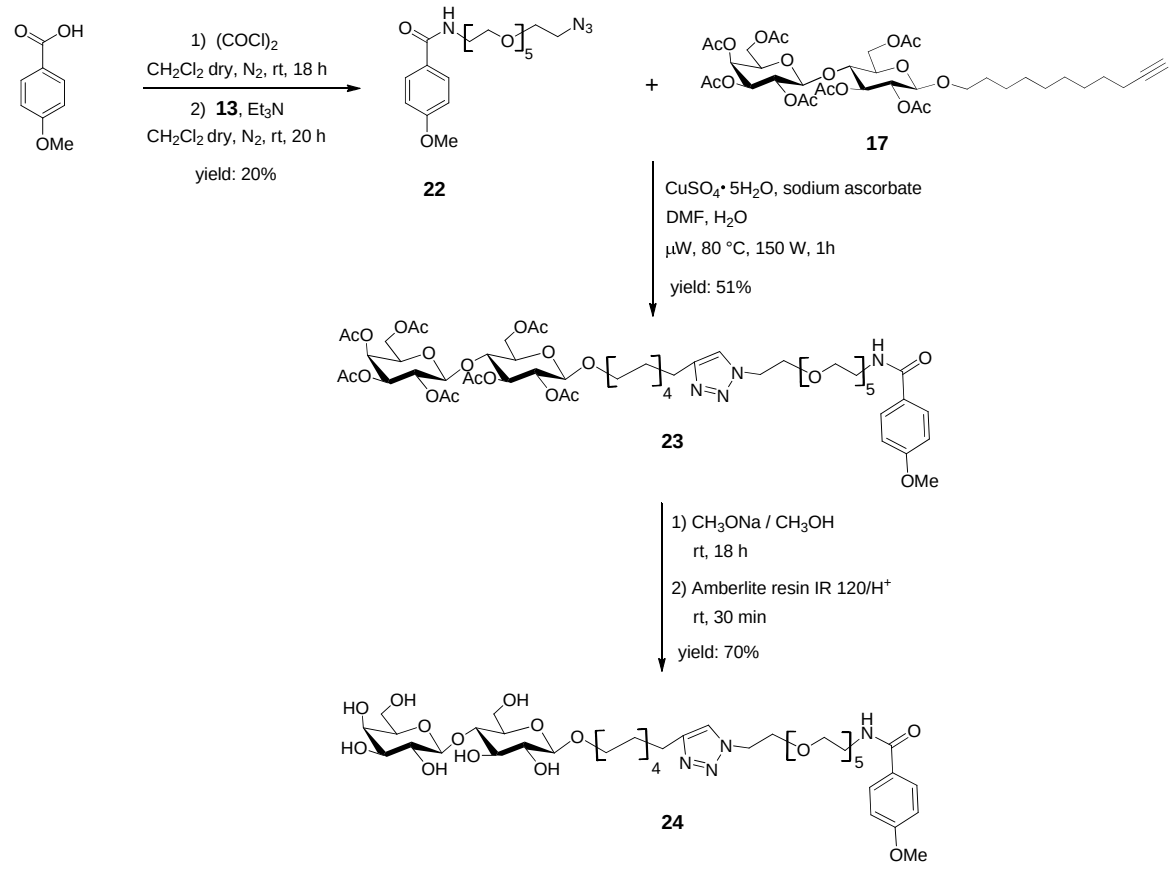


Figure 5.30 Synthesis of the monomeric lactosylated compound **24**.

As for the corresponding calix[5]arene **15**, compound **22** was synthesized by coupling 4-methoxybenzoic acid (commercially available) with the linker **13** through activation of the

carboxylic group as acyl chloride by using oxalyl chloride. The subsequent conjugation between **22** and the lactose derivative **17** via "click" reaction performed under microwave irradiation led to product **23** that was then deprotected from the acetyl groups with the standard Zemplén method. The obtained final product **24** was completely characterized via NMR and ESI-MS techniques.

5.2.3 Preliminary ELISA tests

Preliminary ELISA (*Enzyme-Linked ImmunoSorbent Assay*) experiments were performed with GM1-calix[5]arene **4**. In these ELISA experiments, native ganglioside GM1 was adsorbed on the well surface of the ELISA plates via hydrophobic interactions. Remaining unreacted sites were blocked by a 1% BSA (Bovine Serum Albumin) solution in PBS (*Phosphate Buffered Saline,* pH = 7.4). Subsequently, in separate vials, phosphate buffered solutions of fixed concentrations of CTB-HRP (*Cholera Toxin B subunit HorseRadish Peroxidase conjugate)* (50 ng/mL) were mixed with variable concentrations of GM1-calix[5]arene **4** resulting in twelve vials with concentration for the calixarenes ranging from 10^{-4} M to10^{-15} M. These mixtures were allowed to incubate for two hours at room temperature and then transferred to the GM1-coated wells. These solutions were incubated for 30 minutes at room temperature to allow the cholera toxin to equilibrate between the coated natural-GM1 and the GM1-calix[5]arene **4** present in solution. Then the solution was removed by flicking and the wells washed with 0.1% BSA, 0.05% Tween 20 in PBS solution. Subsequently an OPD (*Ortho-Phenylene Diamine*) solution was added to the wells, and allowed to react with HRP for 15 minutes at room temperature shielded from light. Via enzymatic oxidative coupling OPD is transformed by the CT-linked HRP protein to 2,3-diaminophenazine, which gives yellow-orange to brown solutions, according to the concentration. The optical density (OD) of these solutions is linearly correlated to the concentration of CTB-HRP that had adsorbed on the wells by interaction with the ganglioside GM1. In figure **5.31** this technique is schematically presented. From these experiments, rough estimations of the IC_{50} values for the tested inhibitor could be deduced. Control experiments were performed putting CT-HRP in BSA-coated wells where native GM1 was not present. The low OD acquired from this experiment was subtracted from the value acquired from the calixarene experiments to correct for non-

specific binding of the toxin to the wells by hydrophobic interactions. Also, OPD was allowed to stand for 15 minutes in native GM1-coated, BSA-blocked wells to make sure that the mixture did not color in due time.

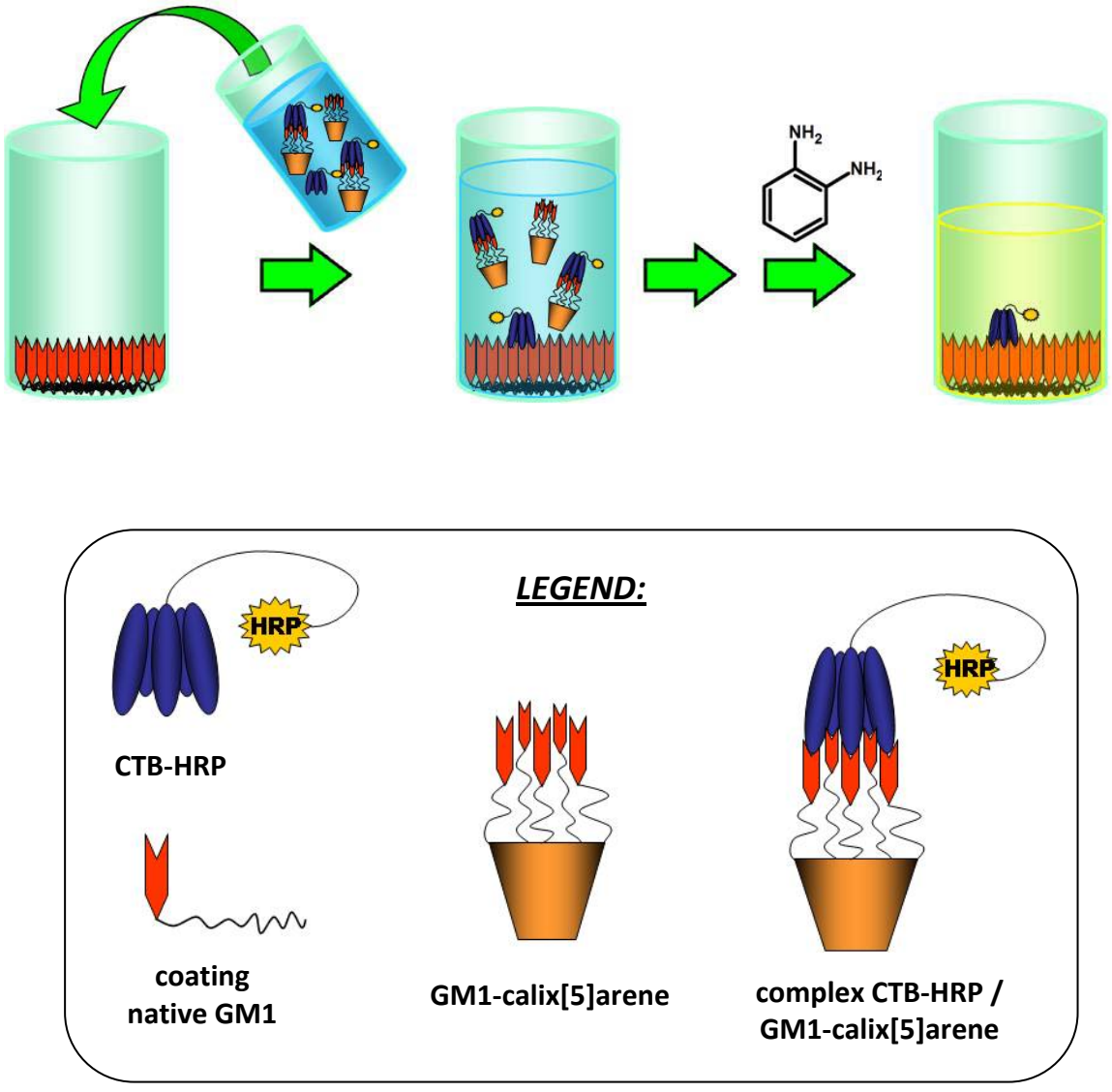


Figure 5.31 Schematic presentation of an ELISA experiment methodology for inhibitor GM1-calix[5]arene**4**.

All the tests were performed in duplicate. The more intense the observed yellow color, the more the CT has linked to the coated native GM1 in the wells and the less efficient the binding of the inhibitor is (figure **5.32**).

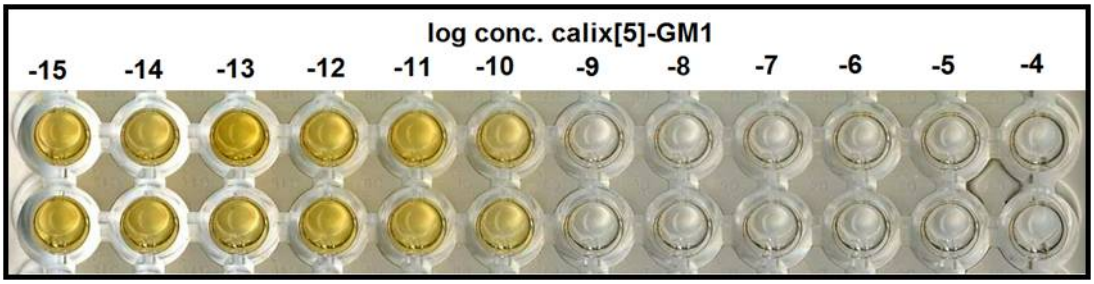


Figure 5.32 ELISA experiment in duplo for inhibition of CT with GM1-calix[5]arene **4**.

The data analysis was performed by first measuring the optical density of the blank, i.e. the native GM1-coated well with all components except the inhibitor and toxin, this gave the lowest response (OD = 0.05). Then, in presence of only the CT, without inhibitor, and this gave the highest response (OD = 1.74). These two values represent the minimum and the maximum values of optical density (0% and 100% binding of the CT to the GM1-coating of the wells) for the experiment. Subsequently, the concentration of the inhibitor GM1-calix[5]arene **4** is varied exponentially and the optical density response measured. All values obtained are combined in a dose-response graph, where the percentage of inhibition is reported in function of the inhibitor concentration. From the curve, it is possible to obtain the IC_{50} value (*half maximal inhibitory concentration*), which is a measure of the effectiveness of a compound in inhibiting a biological function. In other words, it is a quantitative measure that indicates how much of a particular drug, in our case GM1-calix[5]arene **4**, is needed to inhibit 50% of cholera toxin present in solution. The lower the IC_{50} values is, the higher the effectiveness of the analyzed inhibitor.

The ELISA experiments performed so far gave only a preliminary indication on the efficiency of the ligand tested. Very interestingly, the pentavalent calix[5]arene-based glycoconjugate **4** presented extremely high affinity for cholera toxin, with an IC_{50} value in the low nanomolar range.

Considering the preliminary character of this assay, an accurate quantitative treatment of the data is not possible at the moment. Results obtained so far need to be confirmed and to be determined with more precision by performing ELISA tests in a smaller concentration range for the inhibitor. Besides, also experiments with the monomeric compounds need to be performed in order to quantify the multivalent effect.

Other measurements are currently in progress in the laboratory of Wageningen in order to confirm these results and possibly obtain precise and accurate IC_{50} values for this and the remaining inhibitors (**1** – **3**) synthesized during this work.

5.3 Conclusions

In this chapter, the syntheses of four different pentavalent glycocalix[5]arenes and one monovalent compound for cholera toxin inhibition have been reported. The pentavalent compounds have the same valency as the B_5 pentamer of the target protein, therefore they are potentially able to inhibit the CT via formation of 1:1 complexes rather than giving origin to intermolecular aggregates of variable stoichiometry. The synthesis of these compounds was obtained by using a convergent approach. The calix[5]arene core **10**, the linker **13** and the glycosylated units **16**, **17**, **18** and **19** (with galactose, lactose, GM2os and GM1os residues, respectively) were synthesized separately and then conjugated together. The calix[5]arene lower rim was functionalized with methyl groups in order to give a mobile conformation to the molecule and therefore simplify the synthetic steps and the purification procedures. By using the typical peptide synthesis methodologies, it was possible to link the carboxyl groups of the calix[5]arene scaffold to the terminal amine functionality of the hydrophilic linker **13** with the formation of an amide bond. Via Cu-catalyzed 1,3-dipolar cycloaddiction reaction (also known as one of the most representative reaction of "click" chemistry) it was possible to link together the alkyne group of the undec-10-ynyl glycosides and the azido group at the end of the aglycone part by the formation of a triazole ring. By using this strategy, it was possible to synthesize the pentavalent glycocalix[5]arenes **1** - **4** functionalized with five residues of galactose, lactose, GM2os and GM1os, respectively. GM2 and GM1 oligosaccharides were prepared via chemoenzymatic synthesis[47] in the Department of Organic Chemistry at the Wageningen University, NL. The synthesis of the monovalent compound **24** was also achieved. This compound, which presents a residue of lactose, is the exact analogue of a monomeric unit of calix[5]arene **2**. The most interesting inhibitor, the GM1-calix[5]arene **4**, was also preliminary tested in some ELISA experiments. The IC_{50} value obtained in low nanomolar range suggested that this compound could be one

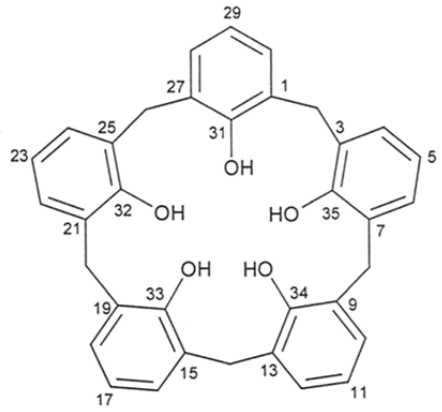
of the most efficient cholera toxin inhibitors known so far. More accurate experiments are ongoing in the group of Prof. Zuilhof in the Department of Organic Chemistry of the Wageningen University (NL) at the time of the writing of this thesis.

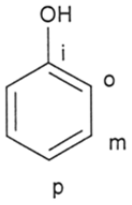
5.4 Experimental Part

General Information. All moisture sensitive reactions were carried out under nitrogen atmosphere, using previously oven-dried glassware. All dry solvents were prepared according to standard procedures, distilled before use and stored over 3 Å or 4 Å molecular sieves. Most of the solvents and reagents were obtained from commercial sources and used without further purification. Analytical TLC was performed using prepared plates of silica gel (Merck 60 F-254 on aluminum) and then, according to the functional groups present on the molecules, revealed with UV light or using staining reagents: $FeCl_3$ (1% in H_2O/CH_3OH 1:1), H_2SO_4 (5% in EtOH), ninhydrin (5% in EtOH), basic solution of $KMnO_4$ (0.75% in H_2O). Merck silica gel 60 (70-230 mesh) was used for flash chromatography and for preparative TLC plates. ^1H NMR e ^{13}C NMR spectra were recorded on Bruker AV300, Bruker AV400, Bruker DPX400 and Bruker AV600 equipped with cryoprobe spectrometers (observation of ^1H nucleus at 300 MHz, 400 MHz, 600 MHz respectively, and of ^{13}C nucleus at 75 MHz, 100 MHz and 151 MHz respectively). All chemical shifts are reported in part per million (ppm) using the residual peak of the deuterated solvent, whose values are referred to tetramethylsilane (TMS, $\delta_{TMS} = 0$), as internal standard. All ^{13}C NMR spectra were performed with proton decoupling. Electrospray ionization (ESI) mass analyses were performed with a Waters spectrometer, while high resolution ESI mass analyses were recorded on a Thermo Scientific Q Exactive spectrometer. Melting points were determined on an Electrothermal apparatus in closed capillaries. Microwave reactions were performed using a CEM Discovery System reactor.

Nomenclature of calix[5]arene compounds.

In this thesis the simplified nomenclature proposed by Gutsche is used to name the calix[5]arene compounds. The positions on the macrocycle are numbered as indicated in the following figure.





5,11,17,23,29-Penta-*tert*-butylcalix[5]arene (5)

p-*tert*-Butylphenol (75.0 g, 0.50 mol), paraformaldehyde (50.0 g, 1.67 mol) and KOH 86% pure (9.0 g, 0.16 mol) were suspended in 30 mL water and 1 L tetralin in a three neck round bottomed flask equipped with mechanical stirrer, thermometer and Dean-Stark trap with condenser. The mixture was stirred for 1 h and 30 min at 80-85 °C under slow nitrogen flow. The reaction was then rapidly heated to 180-185 °C and the N_2 flow was increased to help water removal. The mixture changes color from yellow to dark red. The reaction was kept at this temperature for 10 minutes, then heated at 160 °C for 3 hours and then allowed to cool to room temperature. The formed solid was filtered off and washed with toluene. The tetralin and toluene solution were combined and evaporated to dryness under reduced pressure. The residue (a brownish gum) was vigorously stirred in $CHCl_3$ (250 mL) and HCl 1N (125 mL) at room temperature for 15 minutes. The precipitate was filtered off, the organic layer separated and washed with water (2 x 100 mL), dried over anhydrous Na_2SO_4 and the solvent removed *in vacuo*. The residue was refluxed in 250 mL acetone for 30 minutes. The hot solution was filtered to remove the solid impurities. The obtained solution was concentrated to 150 mL and cooled first to room temperature and then to -20 °C to allow the precipitation of product **5** as white crystals. The purity of the final product and the mixture composition during the different phases of the reaction and work-up could be monitored via TLC (eluent: petroleum ether/acetone 9:1). Yield: 8.6%. **^{1}H NMR** (300 MHz, $CDCl_3$): δ (ppm) 8.64 (s, 5H, ArOH); 7.21 (s, 10H, Ar meta); 3.80 (bs, 10H, $ArCH_2Ar$); 1.26 (s, 45H, CH_3). **M.p.** : 310-312 °C. The product shows the same spectroscopic characteristics reported in literature[41].

Calix[5]arene (6)

$AlCl_3$ in powder (1.76 g, 13.2 mmol) was added into a round bottomed flask containing p-tert-butylcalix[5]arene **5** (1.03 g, 1.26 mmol) in 10 mL toluene. The mixture was vigorously stirred at room temperature for 12 hours and monitored via TLC (eluent: petroleum ether/acetone 20:1). The reaction was then quenched by dropwise addition of 10 mL of 1N HCl. The precipitate was filtered off over a Celite pad and washed with CH_2Cl_2 (30mL). The organic layer was washed with HCl 1N (20 mL) and the aqueous phases extracted with CH_2Cl_2 (2 x 20 mL). The combined organic layers were subsequently washed with brine, water, dried over anhydrous $MgSO_4$, filtered and concentrated under reduced pressure. The crude was purified first by precipitation from CH_3OH and then by recrystallization from toluene, giving compound **6** as a white solid. Yield: 60%. **^{1}H NMR** (300 MHz, $CDCl_3$): δ (ppm) 8.91 (s, 5H, ArOH); 7.21 (d, 10H, J = 7.5 Hz, Ar meta); 6.84 (t, 5H, J = 7.5 Hz, Ar para); 3.80 (bs, 10H, $ArCH_2Ar$). **M.p.** > 275 °C. The product shows the same spectroscopic characteristics reported in literature[42].

31,32,33,34,35-Pentamethoxycalix[5]arene (7)

In a two neck round bottomed flask the calix[5]arene **6** (1.35 g, 2.55 mmol) was dissolved in 100 mL of CH_3CN. K_2CO_3 (10.6 g, 76.5 mmol) and CH_3I (4.70 mL, 76.5 mmol) were added and the mixture was refluxed under nitrogen atmosphere for 24 hours. The progress of the reaction was checked via TLC (eluent: hexane/AcOEt 9:1). The solvent was removed under reduced pressure and the residue redissolved in CH_2Cl_2 and washed first with HCl 1N and then with water till neutral pH. The organic layer was recovered and the solvent removed under reduced pressure, to give a yellow solid that was purified by trituration first with CH_3OH, then with hexane. The product **7** was obtained as a white solid. Yield: 37%. **^{1}H NMR** (300 MHz, $CDCl_3$): δ (ppm) 6.96 (d, 10H, J = 7.4 Hz, ArH meta); 6.79 (t, 5H, J = 7.4 Hz, ArH para); 3.86 (s, 10H, $ArCH_2Ar$); 3.20 (s, 15H, OCH_3). **^{13}C NMR** (75 MHz, $CDCl_3$): δ (ppm) 156.6 (Ar-ipso); 134.5 (Ar-ortho); 128.9 (Ar-meta); 123.1 (Ar-para); 60.4 (-OCH_3); 30.9 ($ArCH_2Ar$). **ESI-MS(+)**: m/z 601.1 [100% $(M+H)^+$]. **M.p.** : 240.1-244.3 °C.

5,11,17,23,29-Pentaformylcalix[5]arene (8)

Calix[5]arene **6** (0.41 g, 0.78 mmol) was added to a solution of HMTA (2.50 g, 17.8 mmol) in 50 mL TFA and the mixture was refluxed for 5 days under N_2. The reaction was regularly checked via TLC (eluent: AcOEt/CH_3OH 9:1). The solvent was then removed under reduced

pressure and the residue dissolved in 12 mL of a 1:1 CH_2Cl_2/HCl 1M solution. The mixture was stirred at room temperature for 24 hours. The aqueous phase was extracted with CH_2Cl_2 (5 x 5mL). The combined organic phases were washed with water (2 x 10mL), dried over anhydrous Na_2SO_4, filtered and the solvent removed under vacuum. The residue was purified by trituration in $CHCl_3$/hexane 1:1 to give the product **8** as a brownish solid. Yield: 57%. **^{1}H NMR** (300 MHz, $CDCl_3$/CD_3OD 9:1): δ (ppm) 9.75 (s, 5H, CHO); 7.70 (s, 5H, ArOH); 7.22 (s, 10H, ArH); 3.45 (s, 10H, $ArCH_2Ar$). **^{13}C NMR** (75 MHz, $CDCl_3$/CD_3OD 9:1): δ (ppm) 192.1 (CHO); 132.5 (Ar-ipso); 129.3 (Ar-ortho); 127.1 (Ar-meta); 126.1 (Ar-para); 31.4 ($ArCH_2Ar$). **ESI-MS(+)**: m/z 693.6 [100% $(M+Na)^+$]. **M.p.** > 300 °C.

5,11,17,23,29-Pentaformyl-31,32,33,34,35-Pentamethoxycalix[5]arene (9)

In two neck round bottomed flask pentaformylcalix[5]arene **8** (0.7 g, 1.05 mmol) was dissolved in 150 mL of dry CH_3CN, then K_2CO_3 (4.5 g, 32 mmol) and CH_3I (2 mL, 32 mmol) were added and the mixture was refluxed for 20 h under nitrogen atmosphere. The reaction was monitored via TLC (eluent: AcOEt/CH_3OH 9:1). The solvent was removed under reduced pressure and the residue dissolved in 150 mL of a 1:1 solution CH_2Cl_2/HCl 1M. The mixture was stirred for 2 h at room temperature. The organic layer was recovered, and the aqueous phase extracted with CH_2Cl_2 (2 x 50 mL). The combined organic phases were washed with water (2 x 100 mL), dried over anhydrous Na_2SO_4, filtered and the solvent removed under reduced pressure. The product **9** was obtained as a brown solid. Yield: 68%. **^{1}H NMR** (300 MHz, $CDCl_3$): δ (ppm) 9.72 (s, 5H, CHO); 7.50 (s, 10H, ArH); 3.92 (s, 10H, $ArCH_2Ar$); 3.25 (s, 15H, OCH_3). **^{13}C NMR** (75 MHz, $CDCl_3$): δ (ppm) 192.4 (CHO); 161.8 (Ar-ipso); 134.8 (Ar-ortho); 132.0 (Ar-meta); 130.8 (Ar-para); 60.6 (-OCH_3); 30.7 ($ArCH_2Ar$). **ESI-MS(+)**: m/z 763.80 [100% $(M+Na)^+$]. **M.p.** : 217-219 °C.

5,11,17,23,29-Pentacarboxy-31,32,33,34,35-Pentamethoxycalix[5]arene (10)

Pentaformyl-pentamethoxycalix[5]arene **9** (0.25 g, 0.34 mmol) was dissolved in a two neck round bottomed flask in 100 mL of a mixture acetone/$CHCl_3$ 1:1, and cooled to 0 °C with an ice-water bath. In another flask a solution of $NaClO_2$ 80% pure (0.47 g, 4.20 mmol) was dissolved in the minimum amount of water and then sulfamic acid (0.49 mg, 5.04 mmol) was added. This solution was slowly poured into the reaction flask. The mixture was stirred at 0 °C for 15 minutes and then slowly warmed up to room temperature and left for 24 hours.

The reaction was monitored via TLC (eluent: AcOEt/CH_3OH 9:1). The solvent was then removed under reduced pressure and the residue triturated with 1M HCl. After filtration on a Buchner funnel, the product **10** was obtained as an orange solid. Yield: 79%. **^{1}H NMR** (300 MHz, CD_3OD): δ (ppm) 7.75 (s, 10H, ArH); 3.93 (s, 10H, $ArCH_2Ar$); 3.30 (s, 15H, OCH_3). **^{13}C NMR** (75 MHz, CD_3OD): δ (ppm) 194.0 (CO); 162.0 (Ar-ipso); 135.8 (Ar-ortho); 132.0 (Ar-meta); 126.8 (Ar-para); 61.2 (-OCH_3); 31.9 ($ArCH_2Ar$). **ESI-MS(+)**: m/z 843.8 [100% $(M+Na)^+$].

Hexaethyleneglycol-di-*p*-toluenesulfonate (11)

A solution of triethylamine (4.9 mL, 35.0 mmol) in CH_2Cl_2 (10 mL) was added dropwise to a solution of *p*-toluenesolfonyl chloride (8.17 g, 43.0 mmol) and hexaethylene glycol (5.00 g, 17.2 mmol) in CH_2Cl_2 (30 mL) cooled to 0 °C with an ice-water bath. A catalytic amount of DMAP was added and the suspension was stirred for 2 hours at 0 °C and then at room temperature for other additional 16 hours. The reaction was monitored via TLC (eluent: petroleum ether/AcOEt 8:2). The mixture was then washed with 0.5M HCl (2 x 50 mL) and with brine (50 mL), then dried over anhydrous Na_2SO_4, filtered and the solvent removed under reduced pressure to give the product **11** as a yellow oil that was used without further purifications in the following reaction. Yield: quantitative. **^{1}H NMR** (300 MHz, $CDCl_3$): δ (ppm) 7.77 (d, 4H, J = 8.1 Hz, ArH ortho); 7.32 (d, 4H, J = 8.1 Hz, ArH meta); 4.13 (t, 4H, J = 4.8 Hz, CH_2OTs); 3.67-3.55 (m, 20H, OCH_2); 2.42 (s, 6H, $ArCH_3$). The product shows the same spectroscopic characteristic reported in literature[45].

1,17-diazide-3,6,9,12,15-pentaoxaheptadecane (12)

In a round bottomed flask, the hexaethylene glycol-di-p-toluenesulfonate **11** (10.1 g, 17.2 mmol) was dissolved in 100 mL of dry DMF. NaN_3 (11.2 g, 172.0 mmol) was then added to the solution and the mixture was stirred at 100 °C for 8 h in nitrogen atmosphere. The reaction was poured into cold water and the product extracted with ethyl ether (5 x 15 mL). The combined organic phases were washed with water (2 x 30 mL) and the solvent evaporated under reduced pressure. Product **12** was obtained as a yellow oil that was not further purified and used directly in the following reaction. Yield: 94%. **^{1}H NMR** (300 MHz, $CDCl_3$): δ (ppm) 3.55-3.45 (m, 20H, OCH_2); 3.20 (t, 4H, J = 5.0 Hz, CH_2N_3). The product presents the same spectroscopic characteristics reported in literature[50].

17-azide-3,6,9,12,15-pentaoxaheptadecane-1-amine (13)

To a solution of diazide-pentaoxaheptadecane **12** (1.45 g, 4.36 mmol) in AcOEt (7mL) and diethyl ether (7 mL) were added HCl 2M (12 mL) and triphenylphosphine (1.07 mg, 4.10 mmol). The biphasic mixture was stirred at room temperature for 24 h under nitrogen atmosphere. The reaction was constantly checked via TLC (eluent: CH_2Cl_2/CH_3OH 9:1). The aqueous phase was washed with CH_2Cl_2 (3 x 15mL), basified with an aqueous solution of KOH and extracted with CH_2Cl_2 (5 x 20 mL). The organic phase was dried over anhydrous $MgSO_4$, filtered and concentrated under reduced pressure to give product **13** as a yellow oil. Yield: 64%. **^{1}H NMR** (400 MHz, $CDCl_3$): δ (ppm) 3.60-3.40 (m, 18H); 3.30 (t, 2H, J = 5.2 Hz, $\underline{CH_2}CH_2NH_2$); 3.18 (t, 2H, J = 5.2 Hz, CH_2N_3); 2.65 (t, 2H, J = 5.2 Hz, $C\underline{H_2}NH_2$); 1.38 (sb, 2H, NH_2). The product presents the same spectroscopic characteristics reported in literature[50].

5,11,17,23,29–Pentakis(17-azide-3,6,9,12,15-pentaoxaheptadecane-1-aminocarbonyl)–31, 32,33,34,35–pentamethoxycalix[5]arene (15)

Method A) In a Schlenk tube, kept under N_2, calix[5]arene **10** (0.10 g, 0.12 mmol) was suspended in 15 mL of dry CH_2Cl_2. Then the amine chain **13** (0.28 g, 0.91 mmol) was added and, in the order, also DIEA (0.2 mL, 1.22 mmol) and HBTU (0.35 g, 0.91 mmol). The mixture was vigorously stirred for 18 h at room temperature under nitrogen atmosphere. The progress of the reaction was monitored via TLC (eluent: CH_2Cl_2/CH_3OH 9:1). The reaction was quenched by addition of a saturated solution of $NaHCO_3$ (20 mL) and extracted with CH_2Cl_2 (2 x 20mL). The combined organic phases were washed with HCl 0.5M and water till neutral pH, dried over anhydrous Na_2SO_4 and the solvent evaporated under vacuum. The crude was purified by flash chromatography (eluent: CH_2Cl_2/CH_3OH 9:1) giving calix[5]arene **15** as a yellow oil. Yield: 24%.

Method B) In a round bottomed flask 0.12 g of calix[5]arene **10** (0.14 mmol) and 0.51 mL of oxalyl chloride (5.82 mmol) were solubilized in 15 mL of dry CH_2Cl_2 under nitrogen atmosphere. The solution was stirred for 18 h at room temperature and then the solvent evaporated to dryness. The residue was dissolved again in 5 mL of dry CH_2Cl_2 and then added dropwise to a solution of amine compound **13** (0.29 g, 0.87 mmol) and NEt_3 (0.12 mL, 0.87 mmol) in 5 mL of dry CH_2Cl_2. The mixture was stirred for 20 h at room temperature under nitrogen atmosphere. TLC (eluent: $CHCl_3$/CH_3OH 94:6) were performed to check the progress of the reaction. The mixture was then washed with 1M HCl, an aqueous solution of Na_2CO_3

and water till neutral pH. The solvent was removed under vacuum and the crude purified by flash chromatography (eluent: $CHCl_3/CH_3OH$ 95:5) to give the product **15** as a yellow oil. Yield: 44%. **^{1}H NMR** (400 MHz, $CDCl_3$): δ (ppm) 7.50 (s, 10H, ArH); 7.07 (bs, 5H, CONH); 3.89 (s, 10H, $ArCH_2Ar$); 3.65-3.50 (m, 110H, OCH_2, $C\underline{H}_2NHCO$); 3.34 (t, 10H, J = 4.8 Hz, CH_2N_3); 3.28 (s, 15H, OCH_3). **^{13}C NMR** (100 MHz, $CDCl_3$): δ (ppm) 167.5 (CO); 159.3 (Ar-ipso); 134.3 (Ar-ortho); 129.8 (Ar-para); 128.2 (Ar-meta); 70.6, 70.4, 70.2, 70.0, 69.7 ($OC\underline{H}_2$); 60.9 ($OC\underline{H}_3$); 50.6 ($C\underline{H}_2N_3$); 39.7 ($C\underline{H}_2NHCO$); 30.7 ($ArC\underline{H}_2Ar$). **ESI-MS(+)**: m/z 1153.9 [100% $(M+2Na)^{2+}$]; 1084.1 [40% $(M-N_2+2Na)^{2+}$].

Undec-10-ynyl-2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranoside (16)

Compound **16** was already present in our laboratory, synthesized according to literature procedures[35,47]. Purification via flash chromatography (elution gradient: hexane/AcOEt 8:1 → 8:2) gave the product **16** as a colorless oil. **^{1}H NMR** (300 MHz, $CDCl_3$): δ (ppm) 5.38 (dd, 1H, J_{4-5} = 0.9 Hz, J_{4-3} = 3.4, H_4); 5.20 (dd, 1H, J_{1-2} = 7.9 Hz , J_{2-3} = 10.5 Hz, H_2); 5.00 (dd, 1H, J_{2-3} = 10.5 Hz, J_{3-4} = 3.4 Hz , H_3); 4.44 (d, 1H, J_{1-2} = 7.9 Hz, H_1); 4.22-4.07 (m, 2H, $H_{6a,}$ H_{6b}); 3.92-3.83 (m, 2H, H_5, β-OCH_a); 3.46 (m, 1H, β-OCH_b); 2.21-2.12 (m, 5H, $C\underline{H}_2C{\equiv}CH$, Ac); 2.04 (s, 6H, Ac); 1.98 (s, 3H, Ac); 1.93 (t, 1H, J = 2.6 Hz , $CH_2C{\equiv}C\underline{H}$); 1.62-1.45 (m, 4H, β-$OCH_2C\underline{H}_2$, $C\underline{H}_2CH_2C{\equiv}CH$); 1.45-1.22 (m, 10H, CH_2 aliphatic chain). **^{13}C NMR** (75 MHz, $CDCl_3$): δ ppm 170.3, 170.2, 170.1, 169.3 (Ac); 101.3 (C_1); 84.6 ($C\underline{{\equiv}}CH$); 70.9 (C_3); 70.5 (C_5); 70.2 (β-$OC\underline{H}_2$); 68.8 (C_2); 68.2 ($C{\equiv}\underline{C}H$); 67.0 (C_4); 61.2 (C_6); 29.3, 29.2, 29.0, 28.6, 28.4, 28.0, 25.7 ($\underline{C}H_2$ aliphatic chain); 20.7, 20.6, 20.5 ($\underline{C}H_3C{=}O$); 18.3 ($\underline{C}H_2C{\equiv}CH$). **ESI-MS(+)** m/z: 521.1 [100% $(M+Na)^+$].

Peracetylated-galactosylcalix[5]arene (20)

Calix[5]arene **15** (32.0 mg, 14.1 µmol) and the β-galactoside derivative **16** (52.9 mg, 106 µmol) were dissolved in 2.5 mL of DMF in a microwave tube. $CuSO_4{\cdot}5H_2O$ (2.0 mg, 8.5 µmol), sodium ascorbate (3.3 mg, 16.9 µmol) and 0.5 mL H_2O were then added. The mixture was heated at 80 °C by microwave irradiation (150 W) for 60 minutes. When the reaction was completed (checked via TLC, eluent: CH_2Cl_2/CH_3OH 10:0.5), it was quenched by addition of water (15 mL) and extracted with AcOEt (5 x 15 mL). The combined organic layers were dried over anhydrous Na_2SO_4, filtered and the solvent removed *in vacuo*. The crude was purified by preparative TLC on plates (eluent: CH_2Cl_2/CH_3OH 9:1) giving product **20** as a yellow oil. Yield: 67%. **^{1}H NMR** (300 MHz, CD_3OD): δ (ppm) 8.21 (bs, 5H, CONH); 7.77 (s, 5H, CH

triazole); 7.61 (bs, 10H, ArH); 5.38 (d, 5H, J = 2.7 Hz, H_4); 5.16-5.02 (m, 10H, H_3, H_2); 4.61 (d, 5H, J_{1-2} = 7.3 Hz, H_1); 4.51 (t, 10H, J = 5.0 Hz, $OCH_2C\underline{H}_2$-triazole); 4.20-4.05 (m, 15H, H_5, H_{6a}, H_{6b}); 3.94 (bs, 10H, $ArCH_2Ar$); 3.88-3.77 (m, 15H, $OC\underline{H}_2CH_2$-triazole, β-$OC\underline{H}_a$); 3.70-3.44 (m, 105H, $OC\underline{H}_2$, β-$OC\underline{H}_b$, $CONHC\underline{H}_2$); 3.29 (s, 15H, $ArOC\underline{H}_3$); 2.67 (t, 10H, J = 7.6 Hz, triazole-$C\underline{H}_2CH_2CH_2$); 2.13 (s, 15H, Ac); 2.02 (s, 15H, Ac); 2.01 (s, 15H, Ac); 1.94 (s, 15H, Ac); 1.71-1.59 (m, 10H, triazole-$CH_2C\underline{H}_2CH_2$), 1.59-1.47 (m, 10H, β-$OCH_2C\underline{H}_2$), 1.40- 1.23 (m, 50H, CH_2 aliphatic chain). **^{13}C NMR** (75 MHz, CD_3OD) δ ppm : 170.6, 170.1, 169.8 (Ac); 168.2 (C=ONH); 159.5 (Ar ipso); 147.6 (C_q triazole); 134.4 (Ar ortho); 129.3 (Ar para); 128.3 (Ar meta); 122.6 ($\underline{C}H$ triazole); 100.8 (C_1); 71.0 (C_3); 70.3 (C_5); 70.2, 70.0, 69.9, 69.5 ($O\underline{C}H_2$); 69.2 ($O\underline{C}H_2CH_2$-triazole); 69.1 (C_2); 67.5 (C_4); 61.2 (C_6); 60.1 ($ArO\underline{C}H_3$); 49.9 ($OCH_2\underline{C}H_2$-triazole); 39.6 ($CONH\underline{C}H_2$); 30.6 ($Ar\underline{C}H_2Ar$); 29.2, 29.0, 28.9 ($\underline{C}H_2$ aliphatic chain); 25.6 (triazole-$\underline{C}H_2CH_2CH_2$); 24.9 ($\underline{C}H_2$ aliphatic chain); 19.4, 19.2, 19.1 ($\underline{C}H_3CO$). **ESI-MS(+)** m/z: 1608.7 [100% $(M+3Na)^{3+}$]; 1211.8 [35% $(M+4Na)^{4+}$].

Galactosylcalix[5]arene (1)

The peracetylated-galactosylcalix[5]arene **20** (45.0 mg, 9.46 μmol) was dissolved in 5 mL of CH_3OH, and drops of a freshly prepared methanol solution of MeONa were added till pH 8-9. The mixture was stirred at room temperature for 4 hours. The progrss of the reaction was monitored via ESI-MS analysis. Amberlite resin IR 120/H^+ was subsequently added for quenching, and the mixture was gently stirred for 30 min till neutral pH. The resin was then filtered off and the solvent removed under vacuum to give pure product **1** as a yellow oil. Yield. 90%. **^{1}H NMR** (300 MHz, CD_3OD): δ (ppm) 8.23 (bs, 5H, CONH); 7.77 (s, 5H, CH triazole); 7.61 (s, 10H, ArH); 4.51 (t, 10H, J = 5.0 Hz , $OCH_2C\underline{H}_2$-triazole); 4.20 (d, 5H, J = 7.1 Hz, H_1); 3.94 (bs, 10H, $ArCH_2Ar$); 3.92-3.80 (m, 20H, H_4, β-OCH_a, $OC\underline{H}_2CH_2$-triazole); 7.77-7.69 (m, 10H, H_{6a}, H_{6b}); 3.66-3.40 (m, 120H, β-OCH_b, H_2, H_3, H_5, OCH_2, $CONHC\underline{H}_2$); 3.29 (s, 15H, OCH_3); 2.67 (t, 10H, J = 7.6 Hz, triazole-$C\underline{H}_2CH_2CH_2$); 1.72-1.53 (m, 20H, triazole-$CH_2C\underline{H}_2CH_2$, β-$OCH_2C\underline{H}_2$); 1.44-1.24 (m, 50H, CH_2 aliphatic chain). **^{13}C NMR** (75 MHz, CD_3OD): δ ppm 168.2 (C=ONH); 159.5 (Ar ipso); 147.6 (C_q triazole); 134.4 (Ar ortho); 129.3 (Ar para); 128.3 (Ar meta); 122.6 ($\underline{C}H$ triazole); 103.6 (C_1); 75.2, 73.7, 71.2 (C_2, C_3, C_5); 70.1, 70.0, 69.9, 69.4, 69.2, 69.1, 68.9 ($O\underline{C}H_2$, β-$O\underline{C}H_2$, C_4); 61.1 (C_6); 60.1 ($ArO\underline{C}H_3$); 49.9 ($OCH_2\underline{C}H_2$-triazole); 39.6 ($CONH\underline{C}H_2$); 30.6 ($Ar\underline{C}H_2Ar$); 29.4, 29.2, 29.1, 29.0, 28.9, 25.7 ($\underline{C}H_2$ aliphatic chain); 24.9 (triazole-$\underline{C}H_2CH_2CH_2$). **ESI-MS(+)** m/z: 1328.1 [100% $(M+3Na)^{3+}$]; 1001.7 [90% $(M+4Na)^{4+}$].

Undec-10-ynyl-2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl-(1→4)-2,3,6-tri-*O*-acetyl-β-D-glucopyranoside (17)

A solution of octa-*O*-acetyl-β-lactose (400 mg, 0.59 mmol) and 10-undecyn-1-ol (198 mg, 1.18 mmol) in dry CH_2Cl_2 (4 mL) was stirred at room temperature for 30 minutes with molecular sieves 4 Å under nitrogen atmosphere. The solution was then cooled to 0 °C with an ice-water bath and $BF_3{\cdot}Et_2O$ (670 µL, 5.43 mmol) was added dropwise. The mixture, kept under N_2, was allowed to slowly reach room temperature and left react for 18 hours. The reaction was monitored via TLC (eluent: petroleum ether/AcOEt 1:1). The mixture was then poured into cold water (4 °C, 15 mL) and extracted with CH_2Cl_2 (3 x 20 mL). The combined organic phases were washed with an aqueous solution of $KHCO_3$ (20mL) and brine (20 mL). then the organic phase was dried over anhydrous Na_2SO_4, filtered and the solvent removed under reduced pressure. The crude was purified by flash chromatography (elution gradient: petroleum ether/AcOEt 2:1 → 1:1). The product **17** was obtained pure as a colorless oil. Yield: 35%. **^{1}H NMR** (300 MHz, $CDCl_3$): δ (ppm) 5.33 (d, 1H, J = 3.3 Hz, H_4'); 5.17 (t, 1H, J = 9.3 Hz, H_3); 5.09 (dd, 1H, $J_{1'-2'}$ = 7.9 Hz, $J_{2'-3'}$ = 10.4 Hz, H_2'); 4.93 (dd, 1H, $J_{2'-3'}$ = 10.4 Hz , $J_{3'-4'}$ = 3.3 Hz, H_3'); 4.86 (dd, 1H, J_{2-3} = 9.3 Hz, J_{1-2} = 8.1 Hz, H_2); 4.48-4.40 (m, 3H, H_1', H_1, H_{6a}); 4.15-4.00 (m, 3H, H_{6b}, $H_6'_a$, $H_6'_b$); 3.88-3.72 (m, 3H, β-OCH_a, H_5', H_4); 3.57 (m, 1H, H_5); 3.42(m, 1H, β-OCH_b); 2.18-2.11 (m, 5H, C<u>H</u>$_2$C≡CH, Ac); 2.10 (s, 3H, Ac); 2.05-2.00 (m, 12H, Ac); 1.95 (s, 3H, Ac); 1.92 (t, 1H, J = 2.0 Hz , C≡C<u>H</u>); 1.56-1.43 (m, 4H, β-OCH_2C<u>H</u>$_2$, C<u>H</u>$_2CH_2$C≡CH); 1.40-1.20 (m, 10H, CH_2 aliphatic chain). The product presents the same spectroscopic characteristics reported in literature[47].

Peracetylated-lactosylcalix[5]arene (21)

Calix[5]arene **15** (27.5 mg, 12.0 µmol) and the β-lactoside compound **17** (70.0 mg, 89.0 µmol) were dissolved in 2.5 mL DMF in a microwave tube. $CuSO_4{\cdot}5H_2O$ (2.6 mg, 10.4 µmol), sodium ascorbate (4.4 mg, 22.2 µmol) and 0.5 mL H_2O were then added. The mixture was heated at 80 °C by microwave irradiation (150 W) for 60 minutes. When the reaction was completed (checked via TLC, eluent: CH_2Cl_2/CH_3OH 94:6), it was quenched by addition of water (15 mL) and extracted with AcOEt (5 x 15 mL). The combined organic layers were dried over anhydrous Na_2SO_4, filtered and the solvent removed *in vacuo*. The crude was purified by flash chromatography (elution in gradient: CH_2Cl_2/CH_3OH 96:4 → 95:5) giving product **21** as a yellow oil. Yield: 57%. **^{1}H NMR** (300 MHz, $CDCl_3$): δ (ppm) 7.46 (bs, 10H, ArH); 7.40 (s,

5H, CH triazole); 6.83 (bs, 5H, CONH); 5.33 (d, 5H, J = 3.3 Hz, H_4'); 5.17 (t, 5H, J = 9.3 Hz, H_3); 5.09 (dd, 5H, $J_{1'-2'}$ = 7.9 Hz, $J_{2'-3'}$ = 10.4 Hz, H_2'); 4.93 (dd, 5H, $J_{2'-3'}$ = 10.4 Hz , $J_{3'-4'}$ = 3.3 Hz, H_3'); 4.86 (dd, 5H, J_{2-3} = 9.3 Hz, J_{1-2} = 8.1 Hz, H_2); 4.50-4.38 (m, 25H, H_1', H_1, H_{6a}, $OCH_2C\underline{H}_2$-triazole); 4.15-4.00 (m, 15H, H_{6b}, $H_6'_a$, $H_6'_b$); 3.90-3.71 (m, 35H, H_5', H_4, β-$OC\underline{H}_a$, $ArC\underline{H}_2Ar$, $OC\underline{H}_2CH_2$-triazole); 3.65-3.47 (m, 105H, $OC\underline{H}_2$, H_5, $CONHC\underline{H}_2$); 3.42 (m, 5H, β-$OC\underline{H}_b$); 3.21 (s, 15H, $ArOC\underline{H}_3$); 2.65 (t, 10H, J = 7.7 Hz, triazole-$C\underline{H}_2CH_2CH_2$); 2.12 (s, 15H, Ac); 2.09 (s, 15H, Ac); 2.04-1.98 (m, 60H, Ac); 1.94 (s, 15H, Ac); 1.65-1.55 (m, 10H, triazole-$CH_2C\underline{H}_2CH_2$); 1.55-1.44 (m, 10H, β-$OCH_2C\underline{H}_2$); 1.39-1.15 (m, 50H, CH_2 aliphatic chain). **^{13}C NMR** (100MHz, $CDCl_3$): δ ppm 170.4, 170.3, 170.2, 170.1, 169.8, 169.6, 169.1 (Ac); 167.3 (C=ONH); 159.2 (Ar ipso); 148.2 (C_q triazole); 134.2 (Ar ortho); 129.9 (Ar para); 128.2 (Ar meta); 121.7 ($\underline{C}$H triazole); 101.1 (C_1'); 100.6 (C_1); 76.3 (C_4); 72.8 (C_3); 72.5 (C_5); 71.7 (C_2); 71.0 (C_3'); 70.6 (C_5'); 70.5 ($O\underline{C}H_2$); 70.2 (β-$O\underline{C}H_2$); 69.8 ($O\underline{C}H_2$); 69.6 ($O\underline{C}H_2CH_2$-triazole); 69.1 (C_2'); 66.6 (C_4'); 62.1 (C_6); 60.8 (C_6'); 60.7 ($ArO\underline{C}H_3$); 50.0 ($OCH_2\underline{C}H_2$-triazole); 39.7 ($CONH\underline{C}H_2$); 31.0 ($Ar\underline{C}H_2Ar$); 29.5, 29.3, 25.8, 25.7 ($\underline{C}H_2$ aliphatic chain); 20.9, 20.8, 20.6, 20.5 ($\underline{C}H_3C$=O). **ESI-MS(+)** m/z: 1571.8, 100% $[M+4Na]^{4+}$.

Lactosylcalix[5]arene (2)

The peracetylated-lactosylcalix[5]arene **21** (42.0 mg, 6.8 μmol) was dissolved in 5 mL of CH_3OH, and drops of a freshly prepared methanol solution of MeONa were added till pH 8-9. The mixture was stirred at room temperature for 18 hours. The progress of the reaction was monitored via ESI-MS analysis. Amberlite resin IR 120/H^+ was subsequently added for quenching, and the mixture was gently stirred for 30 min till neutral pH. The resin was then filtered off and the solvent removed under vacuum to give pure product **2** as a yellow oil. Yield: 72%. **^{1}H NMR** (300 MHz, CD_3OD): δ (ppm) 7.79 (s, 5H, CH triazole); 7.60 (s, 10H, ArH); 4.51 (t, 10H, J = 5.0 Hz, $OCH_2C\underline{H}_2$-triazole); 4.35 (d, 5H, J = 7.3 Hz, H_1'); 4.26 (d, 5H, J = 7.8 Hz, H_1); 3.93 (bs, 10H, $ArCH_2Ar$); 3.90-3.65 (m, 40H, H_4', H_{6ab}', H_{6ab}, β-OCH_a, $OC\underline{H}_2CH_2$-triazole); 3.65-3.35 (m, 135H, H_3, H_4, H_5, H_2', H_3', H_5', β-OCH_b, OCH_2, $CONHC\underline{H}_2$); 3.28 (s, 15H, OCH_3); 3.23 (t, 5H, J = 8.4 Hz , H_2); 2.66 (t, 10H, J = 7.6 Hz, triazole-$C\underline{H}_2CH_2CH_2$); 1.69-1.53 (m, 20H, CH_2 aliphatic chain); 1.42-1.23 (m, 50H, CH_2 aliphatic chain). **^{13}C-NMR** (75 MHz, CD_3OD): δ ppm 168.2 (C=ONH); 159.5 (Ar ipso); 147.5 (C_q triazole); 134.4 (Ar ortho); 129.3 (Ar para); 128.3 (Ar meta); 122.8 ($\underline{C}$H triazole); 103.7 (C_1'); 102.8 (C_1); 79.3 (C_4); 75.7, 75.1, 75.0, 73.4, 73.3, 71.1 (C_2, C_3, C_5, C_2', C_3', C_5'); 70.1, 70.0, 69.9, 69.5, 69.2, 69.0 ($O\underline{C}H_2$, β-$O\underline{C}H_2$,); 68.9

(C_4'); 61.1, 60.6 (C_6, C_6'); 60.1 (ArO<u>C</u>H_3); 50.0 (OCH_2<u>C</u>H_2-triazole); 39.6 (CONH<u>C</u>H_2); 30.6 (Ar<u>C</u>H_2Ar); 29.4, 29.2, 29.1, 29.0, 28.8, 25.7 (<u>C</u>H_2 aliphatic chain); 24.8 (triazole-<u>C</u>H_2CH_2CH_2). **ESI-MS(+)** m/z: 1597.8 [100% $(M+3Na)^{3+}$]; 1204.1 [60% $(M+4Na)^{4+}$].

GM2-calix[5]arene (3)

Calix[5]arene **15** (15.7 mg, 6.93 µmol) was dissolved in 0.5 mL of CH_3OH in a microwave tube. Then the GM2os derivative **18** (50.5 mg, 52.1 µmol), previously synthesized by chemo-enzymatic procedures[35,47], was added together with samples of $CuSO_4 \cdot 5H_2O$ (0.52 mg, 2.1 µmol), sodium ascorbate (0.82 mg, 4.2 µmol), 4 mL of H_2O and a drop of Triton X-100. The mixture was heated at 80 °C by microwave irradiation (150 W) for 60 minutes. The reaction progress was monitored via TLC (eluent: AcOEt\$CH_3OH$\$H_2O$\AcOH 4:2:1:0.1) and ESI-MS analyses. The crude mixture was purified via size exclusion column chromatography (Sephadex G-15, eluent: H_2O 100%) giving product **3** as a white solid sufficiently pure to be used in the preliminary ELISA experiments. Yield: 59%. **^{1}H NMR** (D_2O, 400 MHz): δ (ppm) 7.65 (s, 5H, CH triazole); 7.54 (s, 10H, ArH); 4.70 (m, 5H, $H_{1\text{-Gal}}$ (determined by HSQC)); 4.47-4.42 (m, 15H, $H_{1\text{-GalNAc}}$); 4.36-4.35 (d, 5H, J = 7.6 Hz, $H_{1\text{-Glc}}$); 4.08-4.06 (m, 10H); 3.89-3.66 (m, 95H, Ar<u>C</u>H_2Ar); 3.58-3.43 (m, 150H); 3.32-3.30 (m, 5H, $H_{2\text{-Gal}}$); 3.23-3.19 (m, 5H, $H_{2\text{-Glc}}$); 3.14 (bs, 10H); 2.61-2.58 (m, 5H, $H_{3a\text{-Neu5Ac}}$); 2.48 (bs, 10H); 1.97 (s, 15H, $Ac_{\text{-Neu5Ac}}$); 1.95 (s, 15H, Ac_{GalNAc}); 1.89-1.84 (m, 5H, $H_{3b\text{-Neu5Ac}}$); 1.51-1.36 (bs, 20H, CH_2 aliphatic chain); 1.23-0.95 (50H, m, CH_2 aliphatic chain). **HR-ESI-MS(-)**: m/z 1440.46696 [100% $(M-5H)^{5-}$] calcd: 1438.65884.

GM1-calix[5]arene (4)

Calix[5]arene **15** (15.7 mg, 6.93 µmol) was dissolved in 0.5 mL of CH_3OH in a microwave tube. Then the GM1os derivative **19** (59.8 mg, 52.1 µmol), previously synthesized by chemo-enzymatic procedures[35 47], was added together with samples of $CuSO_4 \cdot 5H_2O$ (0.52 mg, 2.1 µmol), sodium ascorbate (0.82 mg, 4.2 µmol), 4 mL of H_2O and a drop of Triton X-100. The mixture was heated at 80 °C by microwave irradiation (150 W) for 60 minutes. The reaction progression was monitored via TLC (eluent: AcOEt\$CH_3OH$\$H_2O$\AcOH 4:2:1:0.1) and ESI-MS analyses. The crude mixture was purified via size exclusion column chromatography (Sephadex G-15, eluent: H_2O 100%) giving product **4** as a white solid**,** sufficiently pure to be used in the preliminary ELISA tests. Yield: 51%. **^{1}H NMR** (D_2O, 600 MHz): δ (ppm) 7.61 (s, 5H, CH triazole); 7.50 (s, 10H, Ar); 4.69 (bs, 5H, $H_{1\text{-GalNAc}}$); 4.43 (d, 5H, J = 8.0, H1-$_{Gal}$'); 4.44-4.41

(m, 5H, $H_{1\text{-}Gal}$); 4.38-4.37 (m, 5H); 4.32-4.31 (5H, d, J = 7.9 Hz, $H_{1\text{-}Glc}$); 4.05-4.02 (m, 15H); 3.94-3.92 (m, 5H, $H_{2\text{-}GalNAc}$); 3.86-3.60 (m, 105H, ArC$\underline{H}_2$Ar); 3.60-3.45 (m, 80H); 3.46-3.33 (m, 75H); 3.26-3.24 (m, 5H, $H_{2\text{-}Gal}$); 3.18-3.15 (m, 5H, $H_{2\text{-}Glc}$); 3.10 (bs, 10H); 2.56-2.54 (m, 5H, $H_{3a\text{-}Neu5Ac}$); 2.45-2.42 (m, 10H, triazole-C$\underline{H}_2$CH$_2$CH$_2$); 1.92 (s, 15H, Ac_{Neu5Ac}); 1.89 (s, 15H, Ac_{GalNAc}); 1.85-1.80 (m, 5H, $H_{3b\text{-}Neu5Ac}$); 1.44-1.35 (m, 20H, CH_2 aliphatic chain); 1.14-1.08 (m, 10H, C$\underline{H}_2$CH$_2$OC$_{1\text{-}Glc}$), 1.03 (m, 40H, CH_2 aliphatic chain). **^{13}C NMR** (D_2O, 151 MHz): δ (ppm) 175.4, 175.0, 174.5 (C=O); 169.0 (Ar$\underline{C}$=ONH); 159.7 (Ar ipso); 134.7 (Ar ortho); 129.2 (Ar para); 128.9 (Ar meta); 123.7 ($\underline{C}$H triazole); 105.1 ($C_{1\text{-}Gal}$'); 103.0 ($C_{1\text{-}Gal}$); 102.8 ($C_{1\text{-}GalNAc}$); 102.5 ($C_{1\text{-}Glc}$); 102.0 ($C_{2\text{-}Neu5Ac}$); 80.7 ($C_{3\text{-}GalNAc}$); 79.0 ($C_{4\text{-}Glc}$); 77.6 ($C_{4\text{-}Gal}$); 75.2 ($C_{5\text{-}Gal'}$); 75.1 ($C_{5\text{-}Glc}$); 74.9 ($C_{3\text{-}Gal}$); 74.7; 74.4 ($C_{5\text{-}Gal}$); 73.4 ($C_{6\text{-}Neu5Ac}$); 73.1 ($C_{2\text{-}Glc}$); 72.8 ($C_{3\text{-}Gal'}$); 72.6 ($C_{7\text{-}Neu5Ac}$); 71.0 ($C_{2\text{-}Gal}$'); 70.9 (β-O$\underline{C}$H$_2$); 70.4 ($C_{2\text{-}Gal}$); 70.1-69.8 (-O$\underline{C}$H$_2$-); 69.4 (-O$\underline{C}$H$_2$-); 69.2; 69.0; 68.9; 68.4; 68.3 ($C_{4\text{-}GalNAc}$); 63.1; 61.4; 61.3; 60.9; 60.5; 52.0; 51.5 ($C_{2\text{-}GalNAc}$); 50.3; 40.0; 37.2; 31.1 (Ar$\underline{C}$H$_2$Ar); 29.3, 29.1, 29.0, 28.9, 28.8 (CH_2 aliphatic chain); 25.6 ($\underline{C}$H$_2$CH$_2$OC$_{1\text{-}Glc}$); 25.0 (triazole-$\underline{C}$H$_2$CH$_2$CH$_2$); 23.0 (NHC(O)$\underline{C}$H$_{3\text{-}GalNAc}$); 22.4 (NHC(O)$\underline{C}$H$_{3\text{-}Neu5Ac}$). **HR-ESI-MS(-)**: m/z 1600.50134 [100% $(M\text{-}5H)^{5-}$] calcd: 1600.71123.

17-Azide-3,6,9,12,15-pentaoxaheptadecane-1-aminocarbonyl-p-methoxybenzene (22)

Oxalyl chloride (1.5 mL, 16.0 mmol) was added to a solution of 4-methoxybenzoic acid (0.30 g, 2.0 mmol) in 15 mL of dry CH_2Cl_2 and the mixture was stirred at room temperature under N_2 for 18 h. The solvent was then removed under vacuum and the residue dissolved again in 5 mL of dry CH_2Cl_2. This solution was added dropwise to a round bottomed flask containing the amine compound **13** (0.91 g, 3.0 mmol) and NEt_3 (0.5 mL, 3.0 mmol) in 10 mL of dry CH_2Cl_2. The mixture was let react for 20 h at room temperature under nitrogen atmosphere. The reaction was monitored via TLC (eluent: AcOEt). A 1M HCl solution (20 mL) was then added to quench the reaction, and the product extracted with CH_2Cl_2 (2 x 20 mL). The combined organic phases were washed with $NaHCO_3$ saturated aqueous solution (15 mL), brine (15 mL), water (15 mL), dried over anhydrous Na_2SO_4, filtered and the solvent evaporated under reduced pressure. The crude was purified by flash chromatography (elent: AcOEt/acetone 9:1). Product **22** was obtained pure as a yellow oil. Yield: 20%. **^{1}H NMR** (300 MHz, $CDCl_3$): δ (ppm) 7.74 (d, 2H, J = 8.9 Hz, ArH meta); 6.87 (d, 2H, J = 8.9 Hz, ArH ortho); 6.78 (bs, 1H, CONH); 3.80 (s, 3H, OCH_3); 3.65-3.54 (m, 22H, OCH_2, CONHC$\underline{H}_2$); 3.32 (t, 2H, J = 5.0 Hz, CH_2N_3). **^{13}C NMR** (75 MHz, $CDCl_3$): δ ppm 167.0 (C=O); 162.0 (Ar ipso); 128.8 (Ar

meta); 126.9 (Ar para); 113.6 (Ar ortho); 70.6, 70.5, 70.2, 70.0, 69.9 (O<u>C</u>H_2); 55.4 (O<u>C</u>H_3); 50.6 (<u>C</u>H_2N_3); 39.7 (CONH<u>C</u>H_2). **ESI-MS(+)** m/z: 463.0 [100% $(M+Na)^+$]; 435.0 [60% $(M-N_2+Na)^+$].

Peracetylated-lactosyl monomer (23)

Compound **22** (17.0 mg, 38.6 µmol) and the β-lactoside compound **17** (45.5 mg, 57.9 µmol) were dissolved in 2.5 mL DMF in a microwave tube. Samples of $CuSO_4 \cdot 5H_2O$ (9.6 mg, 38.6 µmol), sodium ascorbate (15.3 mg, 77.2 µmol) and 0.5 mL H_2O were then added. The mixture was heated at 80 °C by microwave irradiation (150 W) for 60 minutes. The progress of the reaction was monitored via TLC (eluent: AcOEt/acetone 9:1). The reaction was subsequently quenched by addition of water (15 mL) and the product extracted with AcOEt (5 x 15 mL). The combined organic layers were dried over anhydrous Na_2SO_4, filtered and the solvent removed under reduced pressure. The crude was purified by flash chromatography (eluent: AcOEt/acetone 85:15) giving product **23** as a yellow oil. Yield: 51%. **^{1}H NMR** (400 MHz, $CDCl_3$): δ (ppm) 7.76 (d, 2H, J = 8.5 Hz, ArH meta); 7.52 (bs, 1H, C=ONH); 6.88 (d, 2H, J = 8.5 Hz, ArH ortho); 6.88 (s, 1H, CH triazole); 5.31 (d, 1H, J = 3.3 Hz, H_4'); 5.17 (t, 1H, J = 9.3 Hz, H_3); 5.09 (dd, 1H, $J_{1'-2'}$ = 7.9 Hz, $J_{2'-3'}$ = 10.4 Hz, H_2'); 4.93 (dd, 1H, $J_{2'-3'}$ = 10.4 Hz , $J_{3'-4'}$ = 3.3 Hz, H_3'); 4.85 (dd, 1H, J_{2-3} = 9.3 Hz, J_{1-2} = 8.1 Hz, H_2); 4.55-4.37 (m, 5H, H_1', H_1, H_{6a}, $OCH_2C\underline{H}_2$-triazole); 4.14-4.00 (m, 3H, H_{6b}, $H_6'_a$, $H_6'_b$); 3.90-3.71 (m, 8H, H_5', H_4, , OCH_3, β-O<u>C</u>H_a, O<u>C</u>H_2CH_2-triazole); 3.68-3.51 (m, 21H, O<u>C</u>H_2 , H_5, CONH<u>C</u>H_2); 3.42 (m, 1H, β-O<u>C</u>H_b); 2.66 (bs, 2H, triazole-C<u>H</u>$_2CH_2CH_2$); 2.12 (s, 3H, Ac); 2.08 (s, 3H, Ac); 2.04-1.98 (m, 12H, Ac); 1.93 (s, 3H, Ac); 1.65 (bs, 2H, triazole-CH_2C<u>H</u>$_2CH_2$); 1.51 (bs, 2H, β-OCH_2C<u>H</u>$_2$); 1.39- 1.15 (m, 10H, CH_2 aliphatic chain). **^{13}C NMR** (100 MHz, $CDCl_3$): δ ppm 170.4, 170.3, 170.1, 170.0, 169.8, 169.6, 169.1 (Ac); 167.0 (C=ONH); 162.1 (Ar ipso); 128.9 (Ar meta); 126.9 (Ar para); 113.6 (Ar ortho); 101.1 (C_1'); 100.6 (C_1); 76.3 (C_4); 72.8 (C_3); 72.6 (C_5); 71.7 (C_2); 71.0 (C_3'); 70.6 (C_5'); 70.5 (O<u>C</u>H_2); 70.2 (β-O<u>C</u>H_2); 69.9 (O<u>C</u>H_2); 69.5 (O<u>C</u>H_2CH_2-triazole); 69.1 (C_2'); 66.6 (C_4'); 62.1 (C_6); 60.8 (C_6'), 55.4 (ArO<u>C</u>H_3), 50.3 (OCH_2C<u>H</u>$_2$-triazole); 39.7 (CONH<u>C</u>H_2); 29.5, 29.3, 25.8 (<u>C</u>H_2 aliphatic chain); 20.9, 20.8, 20.7, 20.6, 20.5 (<u>C</u>H_3C=O). **ESI-MS(+)** m/z: 1249.4 [100% $(M+Na)^+$].

Lactosyl monomer (24)

The peracetylated compound **23** (24.4 mg, 19.9 µmol) was dissolved in 5 mL of CH_3OH, and drops of a freshly prepared methanol solution of MeONa were added till pH 8-9. The mixture

was stirred at room temperature for 18 hours. The progress of the reaction was monitored via ESI-MS analysis. Amberlite resin IR 120/H^+ was subsequently added for quenching, and the mixture was gently stirred for 30 min till neutral pH. The resin was then filtered off and the solvent removed under vacuum to give pure product **24** as a yellow oil. Yield: 70%. **^{1}H NMR** (400 MHz, CD_3OD): δ (ppm); 7.83 (m, 2H, ArH meta); 7.79 (s, 1H, CH triazole); 7.00 (m, 2H, ArH ortho); 4,54 (t, 2H, J = 5.1 Hz , $OCH_2C\underline{H}_2$-triazole); 4.38 (d, 1H, J = 7.5 Hz, H_1'); 4.29 (d, 1H, J = 7.8 Hz, H_1); 3.95-3.81 (m, 9H, H_4', H_{6a}, H_{6b}, OCH_3, β-OCH_a, $OC\underline{H}_2CH_2$-triazole); 3.81-3.70 (m, 2H, $H_6'_a$, $H_6'_b$); 3.70-3.45 (m, 26H, H_3, H_4, H_2', H_3', H_5', β-OCH_b, OCH_2, $CONHC\underline{H}_2$); 3.41 (m, 1H, H_5); 3.26 (t, 1H, J = 8.4 Hz , H_2); 2.69 (t, 2H, J = 7.4 Hz, triazole-$C\underline{H}_2CH_2CH_2$); 1.72-1.57 (m, 4H, CH_2 aliphatic chain); 1.43-1.28 (m, 10H, CH_2 aliphatic chain). **^{13}C-NMR** (100 MHz, CD_3OD): δ ppm 168.4 (C=ONH); 162.5 (Ar ipso); 147.7 (C_q triazole); 126.3 (Ar para); 128.8 (Ar meta); 122.7 ($\underline{C}$H triazole); 113.3 (Ar ortho); 103.7 (C_1'); 102.8 (C_1); 79.3 (C_4); 75.7, 75.1, 75.0, 73.4, 73.3, 71.2 (C_2, C_3, C_5, C_2', C_3', C_5'); 70.2, 70.1, 69.9, 69.2, 69.0 ($O\underline{C}H_2$, β-$O\underline{C}H_2$); 68.9 (C_4'); 61.1 (C_6'); 60.6 (C_6); 54.5 ($ArO\underline{C}H_3$); 49.9 ($OCH_2\underline{C}H_2$-triazole); 39.5 ($CONH\underline{C}H_2$); 29.4, 29.2, 29.0, 28.8, 25.7 ($\underline{C}H_2$ aliphatic chain); 24.8 (triazole-$\underline{C}H_2CH_2CH_2$). **ESI-MS(+)** m/z: 955.9 [100% $(M+Na)^+$]; 489.7 [45% $(M+2Na)^{2+}$].

CTB_5 inhibition assay

Each well of a 96-well microtiter plate (F96 Maxisorp™, Nunc) was coated with a 100 µL native GM1 solution (2 µg/mL, 1.3 µM) (Sigma-Aldrich) in PBS (*Phosphate Buffered Saline*) and incubated at 37 °C for 16 hours. Unattached GM1 was removed by washing with PBS (2 x 450 µL), and then the remaining free binding sites were blocked by incubation with 100 µL of a 1% (w/v) BSA (*Bovine Serum Albumin*) solution in PBS for 30 min at 37 °C. Subsequently the wells were washed with PBS (3 x 450 µL). Samples of CTB-HRP (*Cholera Toxin B subunit HorseRadish Peroxidase conjugate*) (Sigma-Aldrich) at constant concentration (50 ng/mL) and GM1-calix[5]arene **4** at variable concentration (between 10^{-4} M and 10^{-15} M) in 0.1% BSA, 0.05% Tween-20 in PBS were incubated at room temperature for 2 hours and then transferred to the coated plate. After 30 min of incubation at room temperature the solution was removed by flicking followed by washing steps with 0.1% BSA, 0.05% Tween-20 in PBS (3 x 450 µL). 100 µL of a freshly prepared solution of OPD (*Ortho Phenylene Diamine*) (25 mg OPD·2HCl, 7.5 mL 0.1M citric acid, 7.5 mL 0.1M sodium citrate and 6 µL of a 30% H_2O_2 solution, pH was adjusted to 6.0 with NaOH) were added to each well and allowed to react in

absence of light for 15 minutes. The color-forming reaction was quenched by addition of 50 µL of 1M H_2SO_4. The adsorbance at 490 nm was measured immediatly after by a µQuant plate reader.

5.5 References

1 Marrero, K.; Sánchez, A.; Rodríguez-Ulloa, A.; González, L. J.; Castellanos-Serra, L.; Paz-Lago, D.; Campos, J.; Rodríguez, B. L. et al. *Research in Microbiology* **2009**, 160 (1), 48-56.

2 Ryan K. J.; Ray C. G. ed. **2004**. *Sherris Medical Microbiology* (4th ed.). McGraw Hill.

3 Faruque, S. M.; Nair, G. B. ed. **2008**. *Vibrio cholerae: Genomics and Molecular Biology*. Caister Academic Press.

4 World Health Organization (WHO), *Weekly Epidemiological Record*, **2010**, *85*, 117-128.

5 Kirn, T. J.; Lafferty, M. J.; Sandoe, C. M. P.; Taylor, R. K. *Molecular Microbiology,* **2000**, 35(4), 896-910.

6 Falnes, P. O.; Sandvig, K. *Curr. Opin. Cell Biol.* **2000**, *12*, 407-413.

7 Wernick, N. L. B.; Chinnapen, D. J. F.; Cho, J. A.; Lencer, W. I. *Toxins*, **2010**, *2*, 310-325.

8 Sixma, T. K.; Pronk, S. E.; Kalk, K. H.; Wartna, E. S.; van Zanten B. A.; Wittholdt, B.; Hol, W. G. J. *Nature*, **1991**, *351*, 371-377.

9 Zhang, R. G.; Scott, D. L.; Westbrook, M. L.; Nance, S.; Spangler, B. D.; Shipley, G. G.; Westbrook, E. M. *J. Mol. Biol.* **1995**, *251*, 563-573.

10 Merritt, E. A.; Sarfaty, S.; van der Akker, F.; L'Hoir, C.; Martial J. A.; Hol, W. G. J. *Protein Science*, **1994**, *3*, 166-175.

11 Ahn-Yoon, S.; DeCory, T. R.; Baeumner, A. J.; Durst, R. A. *Anal. Chem.* **2003**, *75*, 2256-2261.

12 a) Song, X.; Nolan, J.; Swanson, B. I. *J. Am. Chem. Soc*. **1998**, *120*, 4873-4874; b) Singh, A. K.; Harrison, S. H.; Schoeniger, J. S. *Anal. Chem*. **2000**, *72*, 6019-6024; c) Clarke, S. C. *Br. J. Biomed. Sci.* **2005**, *62*, 40-46.

13 Svennerholm, L. J. *Neurochem*, **1963**, *10* (9), 613-623.

14 Svennerholm, L. J. *Lipid Res*., **1964**, *5* (2), 145-157.

15 Turnbull, W. B.; Precious B. L.; Homans, S. W. *J. Am. Chem. Soc*., **2004**, *126*, 1047-1054.

16 Merritt, E. A.; Kuhn, P.; Sarfaty, S.; Erbe, J. L.; Holmes, R. K.; Hol, W. G. *J. Mol. Biol.* **1998**, *282*, 1043-1059.

17 Pickens, J. C.; Merritt, E. A.; Ahn, M.; Verlinde, C. L. M. J.; Hol, W. G. J.; Fan, E. *Chem. Biol.* **2002**, *9*, 215-224.

18 Merritt, E. A.; Sarfaty, S.; Feil, I. K.; Hol, W. G. J. *Structure* **1997**, *5*, 1485-1499.

19 Fan, E.; Merritt, E. A.; Zhang, Z.; Pickens, J. C.; Roach, C.; Ahn, M.; Hol, W. G. J. *Acta Crystallogr. D*, **2001**, *D57*, 201-212.

20 van den Akker, F.; Steensma, E.; Hol, W. G. J. *Protein Sci.* **1996**, *5*, 1184-1188.

21 Kuziemko, G. M.; Stroh, M.; Stevens, R. C. *Biochemistry,* **1996**, *35*, 6375-6384.

22 MacKenzie, C. R.; Hirama, T.; Lee, K. K.; Altman, E.; Young, N. M. *J. Biol. Chem.* **1997**, *272*, 5533-5538.

23 Picking, W. L.; Moon, H.; Wu, H.; Picking, W. D. *Biochim. Biophys. Acta,* **1995**, *1247*, 65-73.

24 Mertz, J. A.; McCann, J. A.; Picking, W. D. *Biochem. Biophys. Res. Commun.* **1996**, *226*, 140-144.

25 Lauer, S.; Goldstein, B.; Nolan, R. L.; Nolan, J. P. *Biochemistry,* **2002**, *41*, 1742-1751.

26 Cai, X.-E.; Yang, J. *Biochemistry,* **2003**, *42*, 4028-4034.

27 Masserini, M.; Freire, E.; Palestini, P.; Calappi, E.; Tettamanti, G. *Biochemistry*, **1992**, *31*, 2422-2426.

28 Arosio, D.; Baretti, S.; Cattaldo, S.; Potenza, D.; Bernardi, A. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 3831.

29 Arosio, D.; Fontanella, M.; Baldini, L.; Mauri, L.; Bernardi, A.; Casnati, A.; Sansone, F.; Ungaro, R. *J. Am. Chem. Soc.*, **2005**, *127*, 3660-3661.

30 Pukin, A. V.; Branderhorst, H. M.; Sisu, C.; Weijers, C. A. G. M.; Gilbert, M.; Liskamp, R. M. J.; Visser, G. M.; Zuilhof, H.; Pieters, R. J. *ChemBioChem*, **2007**, *8*, 1500-1503.

31 Pieters, J. R.; Turnbull, B.W.; Zuilhof, H.; Sisu, C.; Baron, A. J.; Branderhorst, H. M.; Connell, S. D.; Weijers, C. A. G. M., Renko de Vries, Hayes, E. D.; Pukin, A. V.; Gilbert, M.; Visser, G. M. *ChemBioChem*, **2009**, *10*, 329-337.

32 A. Holmner, A.; Askarieh, G.; Okvist, M.; Krengel, U. *J. Mol. Biol.,* **2007**, *371*,754-764.

33 Merritt, A. M.; Zhang, Z.; Pickens, J. C.; Ahn, M.; Hol, W. G. J.; Fan, E. *J. Am. Chem. Soc.*, **2002**, *124*, 8818-8824.

34 Zhang, Z.; Liu, J.; Verlinde, C. L. M. J.; Hol, W. G. J.; Fan, E. *J. Org. Chem.,* **2004**, *69*, 7737-7740.

35 Weijers, C. A. G. M.; Franssen, M. C. R.; Visser, G. M. *Biotechnology Advances*, **2008**, *26* 436-456.

36 Notti, A.; Parisi, M. F.; Pappalardo, S. *Selectively modified calix[5]arenes,* **2001,** Ed. Asfari, Zouhair.

37 Gutsche, D. C.; Iqbal, M. *Org. Synth*., **1993**, *Vol. VIII*, 75.

38 Gutsche, D. C.; Dhawan, B.; Leonis, M., Stewart, D. *Org. Synth*. *Coll*., **1993**, *Vol. VIII*, 77.

39 Casnati, A.; Ferdani, R.; Pochini, A.; Ungaro, R. *J. Org. Chem.*, **1997***, 62*, 6236.

40 Lubitov, I. E.; Shokova, E. A.; Kovalev, V. V. *Synlett.*, **1993**, 647.

41 Gutsche, D. C. *Oppi Briefs*, **1993,** *25*, 137.

42 Bell, S. E. J.; Browne, J. K.; McKee, V.; M. McKervey, A.; Malone, J.F.; O'Leary, M.; Walker, A. *J. Org. Chem.,* **1998,** *63,* 489-501.

43 Smith, W. E. *J. Org. Chem.* **1972**, *37 (24)*, 3972-3973.

44 Valeur, E.; Bradley, M. *Chem. Soc. Rev*., **2009**, *38*, 606–631.

45 Chuanju, Z.; Shijun, L.; Jinqiang, Z.; Kelong, Z.; Ning, L.; Feihe, H. *Organic Letters,* **2007,** *9*, 5553-5556.

46 Schwabacher, A. W.; Lane, J. W; Scheisher, M. W.; Leigh K. M.; Johnson, C. W. *J. Org. Chem.*, **1998**, *63*, 1727-1729.

47 Pukin, A. V.; Weijers, C. A. G. M.; van Lagen, B.; Sun, R. W. B.; Gilbert, M.; Karwaski, M.; Florack, D. E. A.; Jacobs, B. C.; Tio-Gillen, A. P.; van Belkum, A.; Endtz, H. P.; Vissera, G. M.; Zuilhof, H. *ChemBioChem*, **2008**, *343*, 636-650.

48 Jaime, C.; de Mendoza, J.; Prados, P.; Nieto, P. M.; Sanchez, C. *J. Org. Chem.,* **1991**, *5*6, 3372-3376.

49 Zemplén, G.; Pascu, E. *Ber. Dtsch. Chem. Ges.*, **1929**, *62*, 1613-1618.

50 Iyer, S. S. et al. *Tetrahedron Letters*, **2004,** *45* , 4285-4288.

Acknowledgements

At the end of these amazing three years of Ph.D., I would like to thank all the people, colleagues and friends, who shared with me this part of my life and made this Ph.D. period so interesting and enjoyable.

Firstly, I would like to thank Prof. Alessandro Casnati and Dr. Francesco Sansone, my tutors, who gave me the possibility to join their research group. Thanks for supervising me, for leading me in the project during these three years, for giving me useful advices and ideas and the for freedom to work in an independent way within the Ph.D. project. Thanks also for the efforts you put in the correction of this thesis.

Thanks to Prof. Jesús Jimenez-Barbero, who hosted me in his research group in Madrid (E) for three months. Thanks to Javier Cañada, for all the interesting and useful work discussions. Thanks to Luis Calle and to Filipa Marcelo for teaching me with so much patience and passion many NMR techniques and modeling methods that were completely new for me.

Thanks to Prof. Han Zuilhof and to Dr. Carel Weijers, who provided the GM1os and GM2os derivatives and hosted me in their research group in Wageningen (NL) to perform the conjugation of these compounds with calix[5]arene scaffolds. Thanks to Jaime Garcia Hartjes who shared with me this project and gave a substantial contribution to it. Thanks for helping me with the chemical issues, for teaching me the how to perform ELISA tests, and for making my stay there so enjoyable.

Thanks to Prof. Wolf-Dieter Fessner, to Dong Yi and to Ning He for disclosing to me the secrets of the enzymatic synthesis during my stay at the Darmstadt University (D) aimed at chemoenzymatic modification of glycocalixarenes.

Thanks to Prof. Hans-Joachim Gabius and to Dr. Sabine André for the biological assays and for the helpful and stimulating discussions concerning the design of the calixarene clusters for galectin inhibition.

Thanks to Dr. Cyrille Grandjean for providing modified LacNAc isothiocyanates.

Now I would like to thank all the wonderful people who worked with me in Parma in the "lab. 192". Ragazzi, ma quanto è stato bello lavorare con voi!!! Le ore di lavoro non sembravano neanche tali tanto ci siamo divertiti insieme. Vale, grazie! Come avrei fatto senza di te? Sei stata una colonna portante del laboratorio. Unire tanta pazienza a tanta professionalità e simpatia è veramente una cosa rara. Gabri, i tuoi consigli sono stati sempre preziosissimi! Ma quand'è che ci fai conoscere le tue figlie? Fanti, Guerrini, quanti bei momenti insieme! Tra ristate, canzoni e buon umore avete contribuito in modo sostanziale a rendere il 192 davvero un "laboratorio perfetto"! Grazie Maik, sei un amico prezioso nonché un confidente fidato. Gio, grazie per tutti i momenti di divertimento insieme, per le tue perle, per le serate da te e per tanto altro ancora. Grazie ai miei tesisti che mi hanno aiutato col progetto della tesi, nonché hanno allietato le mie giornate: grazie Della, Belletz, Alice e Paola (ora dottoranda). Grazie anche alle altre dottorande, Miky e Ire, che avete condiviso con me gioie e dolori della chimica. E' stato bello e divertente lavorare con voi. Dona, Fulvio, ci siamo fatti grandi ristate insieme, grazie per le belle giornate in lab. che mi avete regalato. Grazie Piermi per la tua allegria e per le belle serate in compagnia! Mara, Ylenia, Aurora, Serena è stato un onore e un piacere lavorare con voi! Grazie a tutte le persone che hanno trascorso più o meno tempo al 192 contribuendo a rendere indimenticabili questi anni di lavoro/divertimento. Ho imparato qualcosa da tutti voi e nel corso del tempo da colleghi siete diventati amici!

Grazie anche ai ragazzi dei vicini lab. 187 e 213. Grazie a Rocco e Alex, che avete condiviso con me i tre anni di dottorato, ancora mitiche sono le gesta di Gargnano! Grazie agli amici (chimici-)fisici! Grisu, Davo, Cri, Scion come non ricordare tutte le allegre pause pranzo insieme in mensa! Grazie in generale a tutti i ragazzi di Chimica che hanno trascorso con me questi tre anni.

Un grazie speciale va a te, Carlo, per aver condiviso così tanto con me e per essermi stato vicino in tante occasioni. Sei una persona speciale.

Grazie Aimi e Meli per la vostra amicizia, per avermi sempre supportata e sopportata, per aver condiviso con me tanti momenti felici e non. Ragazzi, siete degli amici preziosi e unici. Davvero non so come farei senza di voi!

Thanks to all the people met in Madrid. When I arrived there I felt immediately at home, thanks to you! A special thanks goes to my flat mates. Muchisimas gracias Maite, I really enjoyed the time we spent together! Grazie Frances, se i miei tre mesi a Madrid sono stati

così divertenti e speciali é soprattutto grazie a te, ma quanto ci siamo divertite insieme! Thanks to Dorota, my Erasmus friend, we were a great party couple! Luis and Filipa, thanks, gracias, grazie! I want to say that to you in all the languages! Thanks for all the nice, happy and fun moments together! Thanks to João, Khouzaima, Ana, Lidi! All of you contributed in a substantial way to make my stay in Madrid unforgettable. Thanks also to Angeles, Alvaro, Virginia and all the other people of the CIB-CSIC. Thanks to Pablo, Patricia, Gema, Rafa, Carlos, Fernando…! It was really nice to meet you.

Thank also to all the people met at the Darmstadt University. In particolar, thanks to Steffen, for the company and fun in our shared office!

A special thank goes to a very special person. Jaime, to meet you in my life was one of the best things ever happened to me. You are simply incredible! When we are together my heart is filled with pure joy. Thanks for all the emotions you always give me, for the time you want to spend with me and for all the wonderful moments together.

Infine vorrei ringraziare tantissimo la mia famiglia. Grazie a Francesca per i piccoli e grandi momenti di ogni giorno che abbiamo da sempre condiviso. Un enorme grazie va a mamma e papà che mi hanno sempre sostenuta e incoraggiata, dandomi fiducia e permettendomi di arrivare fino a qui.

Silvia