UNIVERSITA' DEGLI STUDI DI PARMA

Dottorato di ricerca in Scienze Chimiche

Ciclo XXVIII (2013-2015)

New multivalent calixarene-based ligands for protein recognition and inhibition

Tutors:

Prof. Francesco Sansone

Prof. Alessandro Casnati

PhD candidate: Marta Giuliani

Coordinator:

Prof. Roberto Cammi

2016

Contents

List of abbreviations……I

Abstract……II

Chapter 1 Calixarene-based ligands for protein targeting 3

1.1 Targeting protein-protein and protein-ligand interactions 4

1.2 Multivalency 6

1.3 Calixarenes 10

1.4 Calixarene-based multivalent ligands for proteins 13

1.5 References 22

Chapter 2 Mixed glycocalix[4]arene-DOPC liposomes for targeted drug-delivery

2.1 Introduction 26

2.1.1 Carbohydrates and their interactions with proteins 26

2.1.2 Carbohydrate Binding Proteins: the Lectins 27

2.1.3 Sugar transporter proteins: GLUT's family 30

2.1.4 Drug delivery and liposomes 31

2.1.5 1,3-dipolar cycloaddition: CuAAC 37

2.2 Results and discussions 38

2.2.1 Synthesis of the linker and conjugation to the calixarene scaffold 39

2.2.2 Functionalization of the saccharide units 40

2.2.3 Conjugation to the calixarene scaffold and deacetylation step 42

2.2.4 Synthesis of the glucosylated monomer 3 45

2.2.5 Preparation of mixed liposomes 48

2.3 Conclusions 53

2.4 Experimental part 54

2.5 References 65

Chapter 3 Glycocalix[n]arenes for immunostimulation .....

3.1 Introduction 70

3.1.1 Carbohydrates: eliciting the immune response 70

3.1.2 Vaccines based on carbohydrates 72

3.1.3 Semi-synthetic and synthetic carbohydrate-based vaccines 73

3.1.4 Encapsulated bacteria: *Streptococcus pneumoniae* and its 19F serotype 76

3.1.5 Vaccines against Streptococcus pneumoniae 78

3.1.6 Glycocalixarenes as immunostimulants against SP serotype 19F 79
3.2 Results and discussion 83
3.2.1 Synthesis of the calix[n]arene scaffolds 83
3.2.2 Synthesis of the saccharide unit 87
3.2.3 Synthesis of glycocalixarenes 1-4 97
3.2.4 Computational studies and NOESY experiments for protected glycocalix[4]arenes 36 and 38 105
3.2.5 Synthesis of the monomeric glycosylated system 108
3.2.6 Biological studies 109
3.3 Conclusions 111
3.4 Experimental part 112
3.5 References 137
3.6 APPENDIX: NMR spectra 143
Chapter 4 Structural studies of protein-ligand interactions .....
4.1 Introduction 150
4.1.1 The ^{1}H-^{15}N HSQC experiment and the chemical shift perturbation to study protein-ligand interactions 150
4.1.2 Protein-ligand crystallization 152
4.1.3 Target proteins 156
4.1.4 Studied ligands 161
4.2 Results and discussion 163
4.2.1 Synthesis of guanidinocalix[4]arenes 163
4.2.2 NMR binding studies 165
4.2.3 Crystallization trials 177
4.2.4 Synthesis of glycocalix[6]arenes 181
4.2.5 RSL-calixarene binding: NMR studies 190
4.3 Conclusions 191
4.3.1 Experimental part 193
4.4 References 207

LIST OF ABBREVIATIONS

The following list reports the significance of various abbreviations and acronyms present in this thesis. In the case of non standard acronyms, used in some part or abbreviate names of particular cases are not in the list, but reported in the chapters.

Ar: Aromatic

Boc: t-Butyloxycarbonyl

DCM: Dichloromethane (CH_2Cl_2)

DIPEA: diisopropylethylamine

DMF: Dimethylformamide

DMSO: Dymethyl sulfoxide

ESI-MS: Electron Spray Ionization Mass Spectroscopy

EtOAc: Ethyl Acetate

NMR: Nuclear Magnetic Resonance

NOESY: NUclear Overhauser Effect/Enhancement Spectrocopy

Pht: Phthalimido

TBAB: Tetrabutylammonium bromide

TFA Trifluoro acetic acid

TLC: Thin Layer Chromatography

UV-VIS: Ultraviolet-visible spectroscopy

Abstract

One of the challenges that concerns chemistry is the design of molecules able to modulate protein-protein and protein-ligand interactions, since these are involved in many physiological and pathological processes. The interactions occurring between proteins and their natural counterparts can take place through reciprocal recognition of rather large surface areas, through recognition of single contact points and single residues, through inclusion of the substrates in specific, more or less deep binding sites. In many cases, the design of synthetic molecules able to interfere with the processes involving proteins can benefit from the possibility of exploiting the multivalent effect. Multivalency, widely spread in Nature, consists in the simultaneous formation between two entities (cell-cell, cell-protein, protein-protein) of multiple equivalent ligand-recognition site complexes. In this way the whole interaction results particularly strong and specific. Calixarenes furnish a very interesting scaffold for the preparation of multivalent ligands and in the last years calixarene-based ligands demonstrated their remarkable capability to recognize and inhibit or restore the activity of different proteins, with a high efficiency and selectivity in several recognition phenomena. The relevance and versatility of these ligands is due to the different exposition geometries of the binding units that can be explored exploiting the conformational properties of these macrocycles, the wide variety of functionalities that can be linked to their structure at different distances from the aromatic units and to their intrinsic multivalent nature.

With the aim of creating new multivalent systems for protein targeting, the work reported in this thesis regards the synthesis and properties of glycocalix[n]arenes and guanidino calix[4]arenes for different purposes. Firstly, a new bolaamphiphile glycocalix[4]arene in 1,3-alternate geometry, bearing cellobiose, was synthesized for the preparation of targeted drug delivery systems based on liposomes. The formed stable mixed liposomes obtained by mixing the macrocycle with DOPC were shown to be able of exploiting the sugar units emerging from the lipid bilayer to agglutinate Concanavalin A, a lectin specific for glucose. Moreover, always thanks to the presence of the glycocalixarene in the layer, the same liposomes demonstrated through preliminary experiments to be uptaken by cancer cells overexpressing glucose receptors on their exterior surface more efficiently respect to simple DOPC liposomes lacking glucose units in their structure. Then a small library of

glycocalix[n]arenes having different valency and geometry was prepared, for the creation of potentially active immunostimulants against *Streptococcus pneumoniae*, particularly the 19F serotype, one of the most virulent. These synthesized glycocalixarenes bearing β-N-acetylmannosamine as antigenic unit were compared with the natural polysaccharide on the binding to the specific anti-19F human polyclonal antibody, to verify their inhibition potency. Among all, the glycocalixarene based on the conformationally mobile calix[4]arene resulted the more efficient ligand, probably due its major possibility to explore the antibody surface and dispose the antigenic units in a proper arrangement for the interaction process. These results pointed out the importance of how the different multivalent presentation in space of the glycosyl units can influence the recognition phenomena.

At last, NMR studies, using particularly ^{1}H-^{15}N HSQC experiments, were performed on selected glycocalix[6]arenes and guanidino calix[4]arenes blocked in the cone geometry, in order to better understand protein-ligand interactions. The glycosylated compounds were studied with *Ralstonia solanacearum* lectin, in order to better understand the nature of the carbohydrate-lectin interactions in solution. The series of cationic calixarene was employed with three different acidic proteins: GB1, Fld and alpha synuclein. Particularly GB1 and Fld were observed to interact with all five cationic calix[4]arenes but showing different behaviours and affinities.

Chapter 1

Calixarene-based ligands for protein targeting

1.1 Targeting protein-protein and protein-ligand interactions

Proteins are sophisticated macromolecules whose amino acid sequence, the primary structure, determines unique 3D folding and functions. They supervise and are involved in countless processes that have as common motif the occurrence of a recognition phenomenon towards specific counterparts. These can be small molecules or other proteins, nucleic acid filaments, polysaccharides.

There are proteins, the antibodies, involved in the activation of the immune response,[1] others play a role in signal transduction[2] and transmission working for instance as selective receptors exposed on the cell surface, in cell proliferation, growth, and differentiation.[3] They can be also structural components providing strength and support to cells,[4] or be involved in the transport/storage of different types of molecules. Many act as enzymes able to catalyze biological processes and molecular transformations.[5]

Proteins are also the receptors on the outer membrane of pathogens they use to adhere to cell membranes and start infections,[6] or the targets that pathogens select on the cell membrane of the hosts to invade an organism, or the tools used by cancer cells to proliferate and spread metastases.[7,8]

Protein-ligand and Protein-Protein interactions (PPIs) have then a crucial role both in physiological and pathological processes. The comprehension of the nature of these interactions and then their inhibition or, on the contrary, the restore of (self)recognition functions lost in consequence of alterations, for example, in the genetic code can then help in developing efficient therapies.

In these recognition processes, proteins can exploit more or less deep binding sites and/or more or less wide regions, called hot points, of their surface.[9,10] Each protein has a unique exterior surface composed of hydrophobic, charged and hydrophilic domains, and unique recognition sites. The knowledge of mechanism and structural features, put in the game by the protein in the recognition process identified as object of study is crucial to design molecules potentially able to interfere with it. Moreover, also residues, regions, contact points far from the real recognition site/area can constitute interesting targets for developing bioactive ligands if the interaction with them can induce allosteric effects useful to modulate the protein activity.

Ligands can then be designed following different approaches: for the recognition and coverage of a relative extended area of the protein surface, that could also be around an active site of the macromolecule, for direct and specific interaction with the protein binding/active sites, for the recognition of single contact points.

The design of synthetic molecules acting as modulators of all these interactions has been widely regarded as a formidable goal for the discovery of new biologically active derivatives.[11] Many drugs currently on the market perform their therapeutic action going to interact with protein targets and, beyond these, a plethora of compounds have been designed, prepared and tested along the years with the aim of interfering with processes involving proteins.

Overlooking, because farer from the type of research conducted in this PhD work, the case of small molecules designed to be used as inhibitors of enzymes and then prepared in the form of mimics of their natural substrates, ligands have been and are created with the aim of inhibiting protein aggregation[12,13] or specific protein interactions.[12,14] In this field there are, for instance, the attempts done against neurodegenerative diseases such as Alzheimer. The aggregation of amyloid beta (Aβ) peptide monomers in vivo, forming oligomers, characterizes this illness and causes several problems. In order to avoid the negative process of self-assembly of amyloid beta into fibrillar plaques, synthetic ligands can be prepared to bind to the native state of these mutant proteins blocking its self-assembly. In the Alzheimer's disease (AD), Kim and Lee[15] described how fullerene is able to inhibit strongly the amyloid peptide aggregation at the early stage, specifically binds to the central hydrophobic motif, KLVFF, of Aβ peptides. In this way fullerene could be a good candidate in the research of new drugs against AD.

On the other hand, other approaches can be explored to prepare bio-active compounds that, on the contrary, induce or reinforce PPIs determining and stabilizing homo- or heterooligomerization[16] processes that generate a beneficial/physiological effect.

Qureshy and co-workers[17] demonstrated for example that a nonpeptide molecule was capable of inducing EPOR dimerization and mimicking the biological activities of erythropoietin EPO, that controls the proliferation and differentiation of erythroid progenitor cells into red blood cells.

In order to enhance the binding efficiency and selectivity of the designed ligands, chemists try to optimise the complementary of their molecules with the selected counterparts in the

protein structures,. identifying hydrophobic domains, charged groups, polarized residues, hydrogen bonding donors and acceptors, and their reciprocal positions and distances.
However, for the mentioned targets like relatively extended surfaces, multiple binding sites, clusters of repeated units, another powerful tool can be taken in account and give a helpful contribution that is the possibility to exploit the phenomenon of multivalency.

1.2 Multivalency

Multivalency[18] can be defined as the ability of an entity, particle or molecule to bind another entity, particle or molecule simultaneously establishing multiple identical binding site-ligand complexes or contacts, each determined by non covalent interactions. (Fig. **1**).
The type of interactions that can occur are hydrogen bonding, cation- π interactions as well as π-π stacking, hydrophobic, ionic, and Van der Waals forces.
Multivalency is a powerful concept that Nature employs to achieve specific and strong, yet reversible interactions. This occurs in particular when the single complex between substrate and corresponding binding site would be too weak to ensure selective recognition and stability to the assembled species.
One of the most important features of a multivalent interaction is that it raises the overall binding constant respect to the sum of the binding constants of the corresponding single monovalent events.

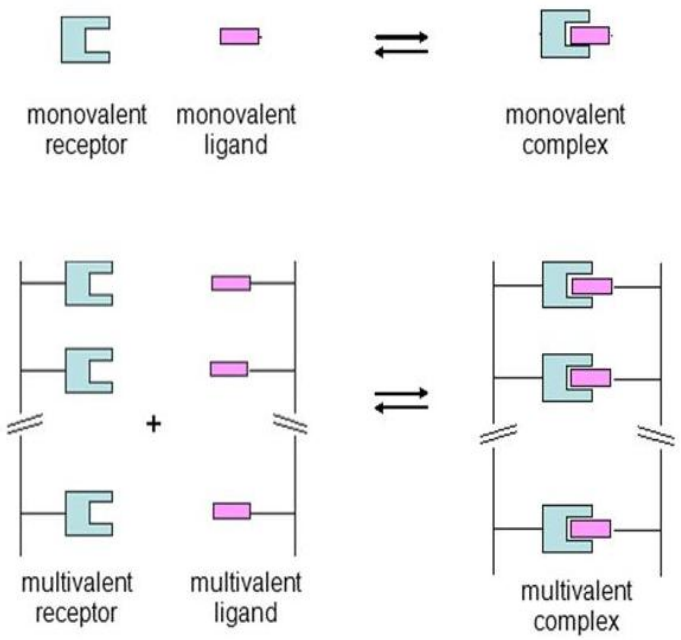


Fig. 1: monovalent and multivalent complex formation comparison

There are different ways for a multivalent ligand to bind a multivalent receptor. Obviously the first interaction is intermolecular. After that, when a complexed form is already established, both inter- and intra- binding events are possible. 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.

Often another concept, the cooperativity,[19] generally intended as the influence of the binding of one ligand on the receptor's affinity toward further binding, is associated with multivalency[20].

Multivalency is also frequently correlated with a favourable *avidity entropy*, a statistical term that increases rapidly with the valency of the complex, always favouring binding and counter balancing the loss of conformational entropy that takes place during multivalent complex formation. On a thermodynamic point of view, the next formula, given by Jenks[21] can be used for a multivalent binding:

eq 1) $\Delta G^\circ_{multi} = n\,\Delta G^\circ_{mono} + \Delta G^\circ_{interaction}$

where ΔG°_{mono} is the standard free energy given by the monovalent binding of a ligand to a receptor and ΔG°_{multi} is the standard free energy for a multivalent ligand. ΔG°_{multi} considers

also the avidity ($\Delta G^\circ_{interaction}$), that is an additional free energy component taking into account and balancing the favourable and unfavourable effects of tethering.

To facilitate the thermodynamic analysis a new parameter was given by Kitov an Bundle[22], rearranging the terms in **eq 1**, based on the concept that multivalent interaction is the result of a first intermolecular binding, followed by intramolecular subsequent interactions. Hence, $\Delta G^\circ_{mono} = \Delta G^\circ_{inter}$ and it is separated from the free energies of all subsequent interactions $(n-1)\Delta G^\circ_{intra} = (n-1)\,\Delta G^\circ_{mono} + \Delta G^\circ_{interaction}$. Therefore the two authors created the following **eq.2**:

eq 2) $$\Delta G^\circ_{multi} = n\,\Delta G^\circ_{inter} + (n-1)\Delta G^\circ_{intra} - RT \ln \Omega_n$$

governed by the Boltzmann-like distribution law. They called in a new way the ΔG°_{multi} using $\Delta G^\circ_{avidity}$ and a third term was added into **eq 2** ($- RT \ln \Omega_n$), that agrees with the generalized Boltzmann-Gibbs definition of entropy, is expressed in entropy units. It was named this last term as "avidity entropy", $\Delta S^\circ_{avidity}$, since it represents the probability of the interaction rather than its strength. Moreover this type of entropy is unique to multivalent interactions, measuring the disorder in the distribution of microscopically distinct complexes. It is a positive factor that allows the multivalent interactions between multivalent ligands and receptors. When the number of binding sites per protein receptor and the number of branches of the ligand increase, $\Delta S^\circ_{avidity}$ can be fundamental.

Actually, it is important to specify that Kitov an Bundle[22] consider an interaction between an oligomeric protein receptor and a multivalent ligand under the two limiting conditions. The first one is that all the n binding sites of the receptor and all the branches of multivalent ligand are independent but have identical binding properties. The second one is that only one multivalent ligand can bind to the oligomeric receptor at a time; if there is a steric impediment, there will be any further interaction even with unoccupied binding sites. Therefore no aggregates were considered by the two researchers.

To characterize a multivalent effect, a parameter was introduced by Whitesides and co-workers,[23] the enhancement factor β. This represents the ratio between the binding constant for the multivalent binding (K_{multi}) of a multivalent ligand to a multivalent receptor and the binding constant for the monovalent binding (K_{mono}) of a monovalent ligand to the same multivalent receptor:

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

An advantage of this enhancement factor is that it can be used even if the multiplicity of the effective binding interactions is unknown. A disadvantage is that it simultaneously also includes the influence of the cooperativity and the symmetry effect.

Another useful factor is the relative potency (rp), defined as the ratio between the IC_{50} value of the monovalent ligand ($IC_{50\ mono}$) and the multivalent ligand ($IC_{50\ multi}$):

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

IC_{50} value is the concentration of a compound needed to inhibit the studied process by 50% and gives a measure of its effectiveness.

When the valency *n* of the cluster is known, the enhancement factor β and the relative potency rp can be both normalised by dividing their values by *n* giving rise to the β/*n* and rp/*n* parameters which measure the gain of each ligating unit obtained by its multivalent presentation respect to the same ligating unit used singularly in a monovalent interaction with the multivalent receptor. These parameters are very good tools to compare the efficiency of clusters having different valency, topology and linker.

In this way clusters exhibiting high β/*n* or rp/*n* values (>>1) are more efficient ligands/inhibitors than their monovalent counterpart (the single ligating unit) and show a positive multivalent effect, while low β/*n* or rp/*n* values (<<1) indicate a lower affinity/inhibition than the monovalent ligand and are therefore characterized by a negative multivalent effect.

On the basis of these considerations and knowledges, in the last years many research groups involved in the development of ligands for proteins exploited multivalency to design and synthesize multivalent systems as potentially highly efficient protein modulators. The presentation in fact of multiple copies of multivalent ligands have unique properties since they can probe multivalent cell/protein-surface with several binding events. They have also many advantages since the scaffold structure, identity of binding elements, number of binding units, and the linker connecting ligating units to the scaffold can be varied systematically, based on structural and chemical properties of the target proteins. [24]

Furthermore in some cases PPIs or protein-ligand interactions require an high level of preorganization. Multivalent ligands have the capability to display the multiple copies of the ligating units in a pre-organized way, giving peculiar properties to these designed compounds.

There are several families of scaffolds of diverse structure, flexibility, and valency with great importance in the development of synthetic multivalent ligands for targeting PPIs and protein-ligand interactions[25]. They have different chemical features but all generally consist of the main core, the scaffold, that bears several covalent connections determining the valency of the cluster, the linkers or spacers, with the peripheral functionalities or ligating units.

1.3 Calixarenes

Calixarenes[26] are ideal scaffolds for the creation of multivalent ligands with biological application, as protein recognition and inhibition.[10,27] These macrocycles have gained the attention in the field of Supramolecular Chemistry[28], the chemistry beyond the molecules, in the last decades and actually present an intrinsically multivalent structure

Calix[n]arenes are compounds obtained from the condensation of para-substituted phenols and formaldehyde in basic conditions. The name of these synthetic macrocycles derives from their similarity to the shape of a Greek vase called *calyx* crater (Figure **2a**). The term *arene* is correlated to the presence of the aryl residues as main components of the macrocycle (Figure **2b,c**).

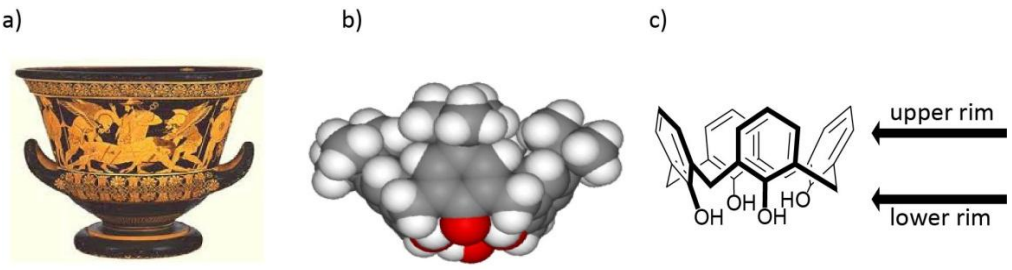


Fig. 2: a) "calyx krater", Greek vase, from which the name of this class of macrocycles derives; b) CPK-model and c) molecular formula of *p-tert*-butyl-calix[4]arene in cone geometry.

As shown in figure **2c** it is possible to distinguish the lower (or narrow) rim, where are present the hydroxyl groups, and the upper (or wide) rim, identified by the para-positions of the phenol units.

It is possible to selectively synthesize calixarenes with different valency, that is with a different number of phenolic units, easily varied at least from 4 to 8.[29] For example, changing the base in the reaction to form these macrocycles, the main product is calix[4]arene or calix[6]arene if NaOH or KOH are used, respectively.

The even-numbered calix[n]arenes (n=4, 6 and 8), reported in figure **3,** are the most common macrocycles of this family but it is possible to obtain also odd-numbered calix[n]arenes (n=5, 7 and 9) although in lower yields.

This means that it is possible to create multivalent systems based on a core of the same nature, the calixarene structure, but displaying a variable valency in term of binding units.

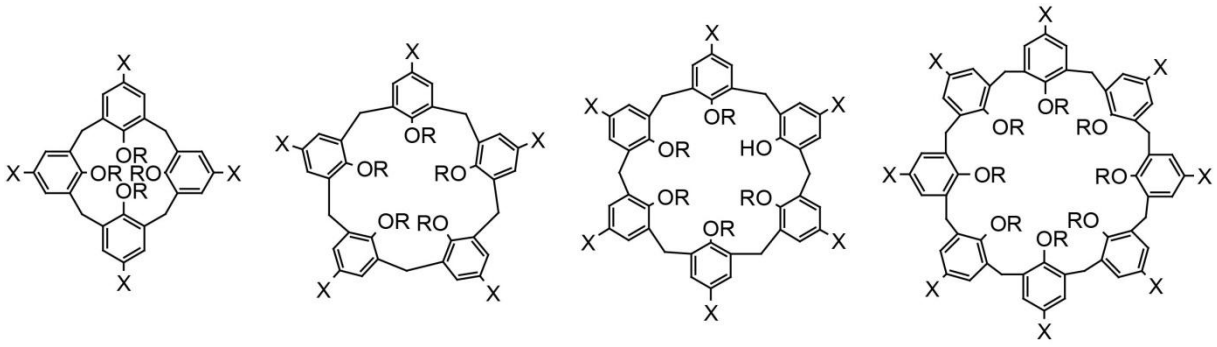


Fig. 3: different valencies of calix[n]arenes (n = number of X units = 4, 5, 6 and 8)

Another advantage coming from the use of calix[n]arenes to build multivalent ligands is the possibility to explore the effects of different orientations of the ligating units in the space taking constant the nature of scaffold and linkers. Calix[n]arenes with valency major than 4 (n>4) are for example conformationally mobile in solution. They are usually present as a mixture of conformers that possess similar energies and rapidly interconvert from one to the other and so the conformational control is difficult. Extensive bridging at the upper or lower rim can fix these larger derivatives in one particular structure. Indeed, also for calix[4]arenes the tetramethoxy-and tetraethoxy derivatives are conformationally mobile, but when at the lower rim are introduced bulky substituents, typically larger than the ethyl group, it is possible lock the macrocycle into four possible following efficient stereoselective procedures based on well-established experimental conditions aimed at reaching this result.[26]

Cone, partial cone, 1,3-alternate and 1,2-alternate are the names of these conformations reported in figure **4**. It is clear as these four different geometries and, in addition, the conformationally mobile analogues allow the design of small families of multivalent ligands

characterized by a subtly modulated orientation of the ligating units. This increases the chance of finding out a proper ligand for a protein and facilitates the identification of a relationship between efficacy of the binding and orientation of the ligating units in the space.

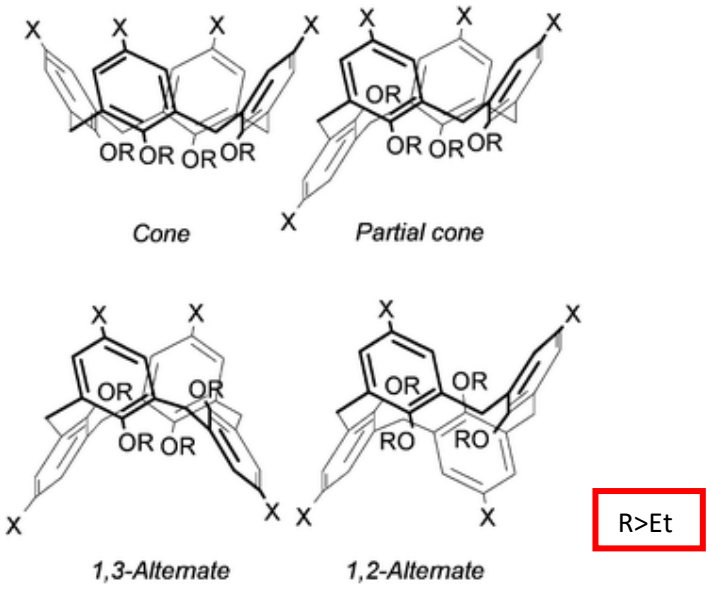


Fig.4: different conformations adopted by calix[4]arene and isolated when R>Et

Exploiting these macrocycles, an additional benefit is correlated to the possibility to functionalize upper and/or lower rim with different functional groups, with also selectivity among the equivalent positions. The chemistry of these cyclic derivatives is nowadays well known and efficient procedures are available to easily functionalize both the narrow (phenolic OH groups) and the wide (aromatic nuclei) rim with many functional groups (-CHO, -COOH, $-NH_2$, -NCS, -CCH…etc.) usually employed for the conjugation of the ligating units and of well suited spacers of different lengths. It is also possible to introduce elaborated ligating units such as carbohydrates, creating the class of the so called glycocalixarenes (see paragraph 1.4) or amino acids giving rise to peptidocalixarenes.[30,31]

Last but not least, the electron rich phenolic units and the aromatic region they define can play relevant roles in the binding processes that calixarenes establish with their targets. CH-π and π-π interactions can contribute to the complexation event and in the case of calix[4]arene derivatives blocked in the cone conformation the inclusion of positively

charged groups like ammonium ions or lipophilic groups like phenyl or alkyl chains can occur particularly in aqueous environment.[32]

1.4 Calixarene-based multivalent ligands for proteins

Ligands built on calixarenes have demonstrated very interesting capability to interact with proteins, evidencing high efficiency and selectivity in the studied recognition phenomena, thanks to their intrinsic multivalent nature, the wide variety of functionalities that can be linked to their structure at different distances from the aromatic units, the different exposition geometries they can furnish.

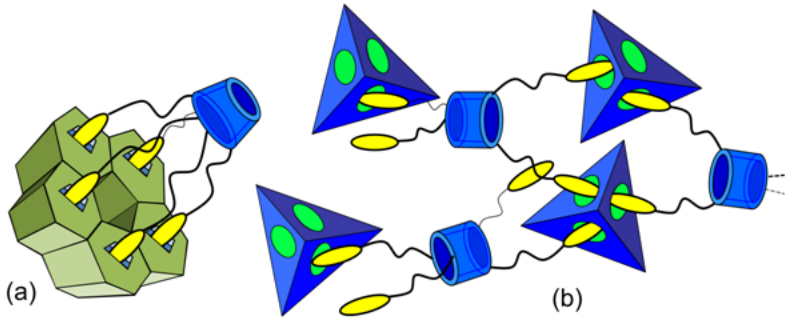


Fig.5: different complexes formed with cone and 1,3-alternate calixarenes (figure from Chem Comm, 2015, 51, 14140-14159)

The different shapes the calixarene-based ligands can assume bring to distinct modes of binding to the biomacromolecules. For example a cone derivative can be able to interact with large surfaces or with species having the binding sites relatively close to each other and disposed on the same spatial regions, giving rise to multivalent 1:1 complexes (figure **5a**). Furthermore the cone calix[4]arene has a right geometry for the inclusion in the aromatic cavity of both lipophilic residues and positively charged groups. On the other hand the 1,3 alternate geometry arranges the ligating units in two opposite regions of the space and for this reason is more suitable to connect different receptors, creating multivalent cross-linked aggregates as shown in figure **5b**.

Instead the conformationally mobile calixarenes can take advantage from the possibility of an induced arrangement of the ligating units in the presence of the protein with an optimization of the interactions that can counterbalance a higher lost in entropy.

It is important to say that the recognition phenomena between amino acids integrated in peptide chains and several calixarene ligands are mainly driven by electrostatic,[33] by hydrophobic interactions, with binding affinity governed by multivalent (cooperative) interactions. In this way the design of multivalent macrocycles bearing complementary functionalities to the protein features is fundamental for enhancing these recognition processes.

As seen above, proteins can interact with their counterparts through surface interactions even involving relatively extensive areas. In this context, examples of calixarene derivatives have been reported able to interact with protein surfaces through three different modes of binding, schematically reported in figure **6**:

1) interactions with the so-called “hot-spots”;

2) single point recognition, that means substantially complexation of the side chain of a selected amino acid in the peptide sequence;

3) recognition and stoppage of lipophilic pockets.

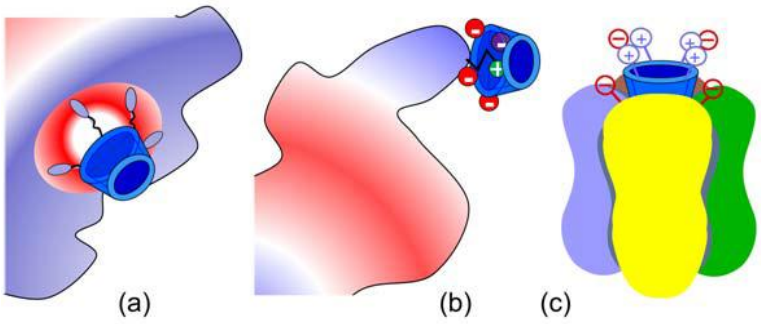


Fig. 6: schematic representation of three different binding modes of calixarene-based multivalent ligands: a) recognition of “”hot spots”; b) interaction with a specific residue in the protein backbone; c) recognition and stoppage of pocket in a protein. (figure from Chem Comm, 2015, 51, 14140-14159)

Regarding the first mode, the “hot spots” in a protein are key binding regions, providing most of the binding energy actually involving much larger areas. They are constituted by a few pivotal residues and define a surface area significantly smaller than the entire interacting interface[34] For this reason they represent an interesting and realistic target for the design of ligands able for instance to disrupt PPIs.

The cone geometry of the calix[4]arene plays an important role to this end, due to its structural features. For instance Hamilton and co-workers[35] reported a series of cone calix[4]arenes able to interact with the "hot spots" of a protein surface of several hundreds of Å. In fact the prepared synthetic agents contained a functionalized and large surface area at the upper rim in order to recognize the target exterior region. The protein targeting was indeed possible thanks to the simultaneous presence of a lipophilic region represented by the calixarene cavity surrounded by polar groups possessing all the same charge/polarity complementary to the charge of groups exposed by the macromolecules around a lipophilic patch.

In a very interesting project, they[36] designed calix[4]arenes with at the upper rim four cyclopeptides (reported in figure **7**) capable to interact with a large surface of cytochrome c, characterized by a lipophilic region, with a positively charged belt that surrounds the heme edge, correlated to its enzymatic activity.

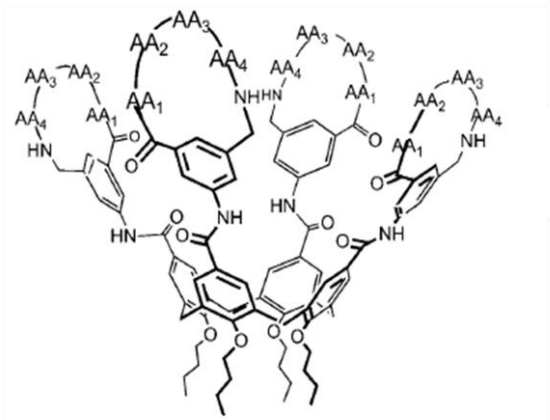


Fig. 7: Calix[4]arenes bearing cyclopeptides at the upper rim studied by Hamilton and co-workers[36]

For this reason at the upper rim negatively charged residues, namely aspartates, were introduced for the interaction whit the ammonium group of lysine residues. In order to increase the lipophilic contacts, an extension of the aromatic character of the calixarene scaffold with 3-aminomethylbenzoyl-based dipeptide was carried out.

Concerning the "single point" recognition, there are several examples in literature. This mode of recognition phenomenon is correlated to the ability of calixarene to interact with a single amino acid, or a limited number of amino acids, in of the protein sequence.

Not only the functionalities added to the calixarene scaffold are important to recognize these "single points". In fact, the aromatic cavity, well defined in the cone calix[4]arene derivatives, can trap lipophilic and positively charged side chains, through π- π, CH- π and cation- π interactions, providing the ligand with the capability to selectively bind this points in the peptide sequence.

For instance Mc Govern et al.[37] described this type of interaction for tetra-sulfonato-calix[4]arene with the outer surface of the cytochrome c protein, selected as a model protein since it is lysine-rich. This derivative, fixed in a cone geometry thanks to the hydrogen bonding array at the lower rim,[38] is able to include the side chain NH_3^+ group of lysines. The proof of this binding mode was supported by NMR studies in solution and the X-ray resolved structure of the complex. In the latter it was observed that the asymmetric unit contained two protein molecules and three sulfonato calixarenes, each complexing lysine side chains. Particularly, as it is possible to observe in figure **8**, the aromatic cavity of the macrocycle includes the amino acid side chain in a hook-like shape, projecting the positively charged ammonium groups towards the anionic sulfonate groups.

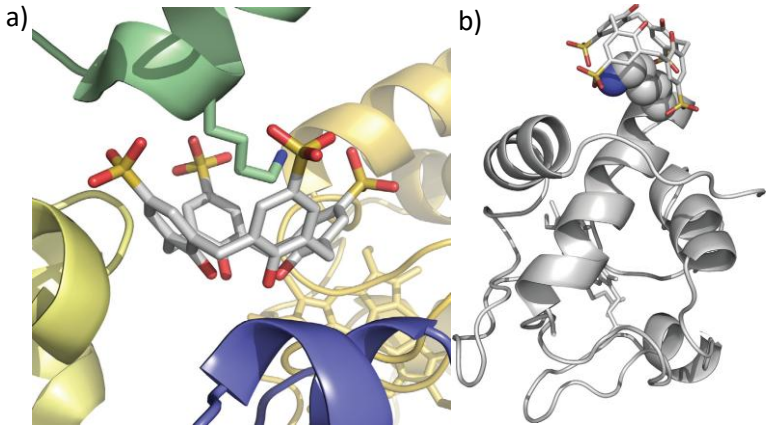


Fig. 8: a) Crystal packing around the molecule of sulfonato calix [4]arenes bound to Lys4 involving three additional protein chains and b) calixarene-bound Lys side chains shown as CPK (Figures from Nat. Chem., 2012, 4, 527-533).

Together with these electrostatic interactions also hydrophobic contacts were established between part of the amino acid carbon side chain and the aromatic units of the calixarene contributing to the stabilization of the complex. This is an example of protein camouflage

mediated by calixarene, giving rise to alteration in the cytochrome c interaction properties with natural counterparts.

The same tetra sulfonato calixarene was used to study its interactions with lysozime. It was observed complexation, in this case towards the arginine side chain and in particular, with high selectivity towards only Arg128 among the eleven arginines present in the sequence. This amino acid discrimination is probably due to the higher accessible surface area of this residue respect to the others. As observed in the X-ray crystal structure of the complex, the Arg side chain is included in the cavity of the macrocycle similarly to the lysine side chain of cytochrome c, with the guanidinium group pointing towards two of the sulfonate units that participate to the complexation forming salt bridge interactions with this. Also in this case charge-charge salt bridges and CH-π interactions are involved together.

Another interesting case again was explained by Crowley and co-workers[39] where modified lysozime (bis-methylated on lysines) was able to interact with the sulfonato calix[4]arene presented in the previous example. Of the six possible dimethyllysine sites only one was recognized and especially respect to the previous example, the selectivity moved on from Asp to Lys. Different was also this protein-ligand complex compared to that with cytochrome c: the methylated lysine side chain penetrated in the aromatic cavity, establishing cation- π interactions, with methylene units excluded, remaining at the exterior of the cavity.

Calixarenes are able also to function as a blocking cap of lipophilic pockets or as glue to keep together the quaternary structure of a protein. To this end, two very interesting examples were given by De Mendoza and co-workers, using guanidino calixarenes.

Guanidinium group is known to be a very efficient binding unit for anions, in particular carboxylate and phosphates, thanks to its geometry, the presence of multiple hydrogen bonding centres and because of it very low acidity it results protonated in a wide range of pH.

In the first example[40] they designed a series of calix[4]arene fixed in the cone geometry, with guanidinium groups at the upper rim, to block the activity of potassium channels. The conical shape effectively could be accommodated into the channel, from the side of the lower rim. At the same time the guanidinium moieties could interact via charge-charge interactions with the negatively charged residues situated at the external surface surrounding the channel entrance.

Additionally, when lipophilic chains are present at the lower rim, the ionic current was decreased in a irreversible way, while the presence of more hydrophilic chains caused a reversible process. This example evidences the importance of the simultaneous presence of the guanidinium units and of the calixarene conical shape for the interactions with potassium channels. In fact the same calixarene scaffold functionalized at the upper rim with negatively charged groups such as carboxylates or even positively charged primary ammonium moieties did not give appreciable activity.

In the second example from De Mendoza's group, [41,42] the studied protein was the homotetrametic p53, a tumor suppression protein that induces the expression of DNA repair machine when DNA is damaged or apoptosis if the damage of DNA is irreversible. The mutated form of p53 is characterized by the replacement of Arg337 with Hys in each of the four peptide chains of the wild-type protein, and as a consequence it shows defective association properties. In fact the imidazole group of Hys, since it is not always protonated at physiological pH, such as the guanidine on the Arg, is not able to strongly interact with Asp on the opposite protein units. Therefore this mutation weakens the tetramerization domain, which is typical of wild-type p53 and strictly related to the protein activity against tumour cells. Since the presence of two Glu on each of the two opposite monomers, above a hydrophobic pocket, allows squared disposition of negative charges at both the bases of the tetramer, cone calix[4]arenes (reported in figure **9a,b**), bearing guanidinium groups were prepared. The biophysical study of the interaction between tetraguanidinium calix[4]arenes and tetramerization domain of the natural mutant p53TD-R337H showed that the ligands are able to stabilize the tetramer through salt bridges and hydrophobic interactions (figure **9c**). The formers are given by the interactions between guanidinium and carboxylate units, the latter ones by the hydrophobic contacts between the protein pocket and the calixarene aromatic skeleton together with the chains at the lower rim. It is interesting to note that a flexible guanidinocalixarene (figure **9b**) resulted in a higher affinity to the mutated protein domains respect to the conformationally rigid analogue (figure **9a**). Normally flexibility is considered a thermodynamic drawback, but not in this case. Probably this major mobility allowed an increased induced fit binding of guanidinium groups with the negative carboxylates, maximizing electrostatic interactions. Moreover, the lipophilicity of the linear

alkyl chains at the lower rim of the flexible ligand gave additional hydrophobic interactions, since they can deeply penetrate into the pocket.

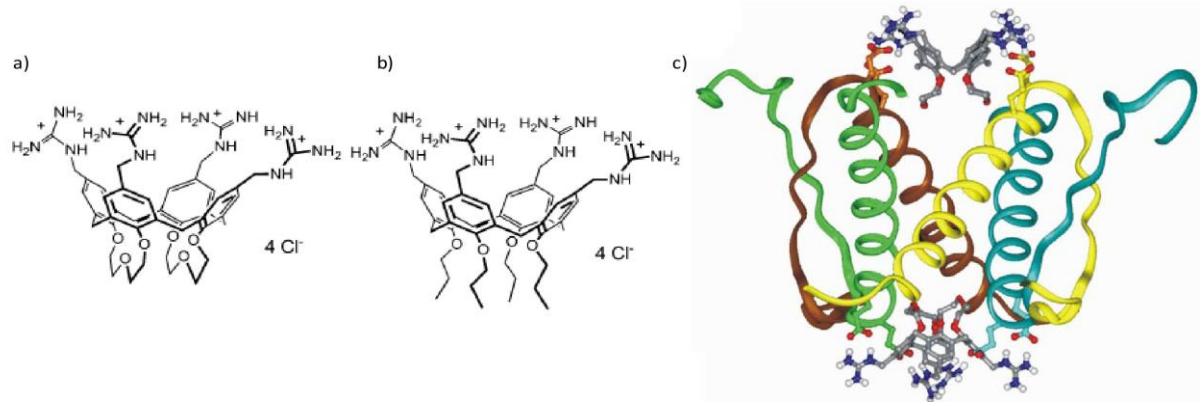


Fig. 9: a) more rigid and b) more flexible guanidinium calix[4]arene studied by De Mendoza;[42,41] c) Tetramerization domain with two molecules of more flexible calixarene showing the interactions of the ligand guanidinium groups and glutamates; it is possible to note the fitting of the lower rim of the ligand into the hydrophobic pocket of the protein.

This last example is a case that well shows the ability of a synthetic ligand to stabilize a protein assembly that it was otherwise very weak in consequence of a mutation.

The second approach, more classical, to pursuit protein-ligand interactions and then to result for instance in protein inhibition, is to target the binding sites used by the protein to complex its substrates. A typical example is given by the carbohydrate binding proteins (CBPs).

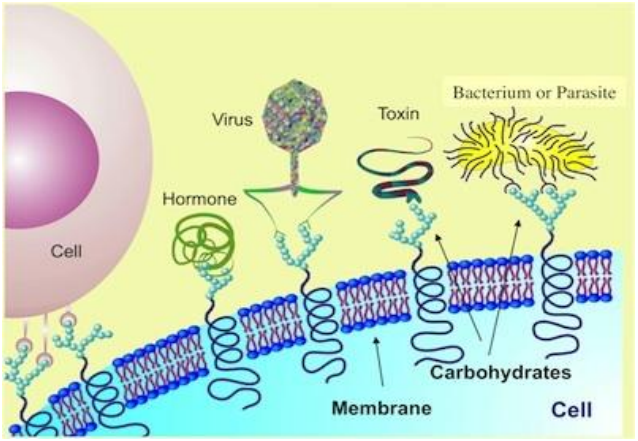


Fig. **10**: carbohydrate-protein interactions in biological processes (from https://www.griffith.edu.au/science-aviation/institute-glycomics/research)

Since the affinity of a single carbohydrate unit for its receptor is usually low, the strong binding observed in biological recognition events involving these species is determined by the simultaneous complexation of several identical glycoside residues by protein receptors bearing several equivalent binding sites. This is a particular case of multivalency, also named *glycoside cluster effect.*[43]

Being protein-carbohydrates interactions at the basis of many fundamental biological processes (as schematically reported in figure **10**), in the early 1980s chemists started to focalize their attention on the synthesis of glycoclusters,[44] with the aim of producing molecules that exploiting multivalency could inhibit these phenomena or recognize selectively the CBPs ,and in particular among them lectins, with high selectivity and potency.[45]

There are several families of glycoclusters but among them glycocalixarenes,[46,47] calixarenes bearing at the upper and/or at the lower rim sugars, already demonstrated the capability to recognize efficiently these type of macromolecules, thanks to all the advantages formerly listed. One of the key features that distinguishes these macrocycles is the possibility to modulate easily the valency and the topology of the sugar presentation together with their straightforward preparation and functionalization.

Often the best shape and valency of these calixarene-based ligands to obtain the best inhibition and protein recognition cannot easily be predicted. However when the biomacromolecule has a particular structure the choice of the calixarene molecules is simpler. This is the case of cholera toxin (CT), that possess 5 bindings units for the pentasaccharide of the GM1 ganglioside. In this way a calix[5]arene scaffold[48] was designed and since the receptor sites are exposed to the same part, a suitable spacer was designed to in principle allow the simultaneous complexation of the five saccharide units by the five CT recognition sites and form a 1:1 complex. Particularly a 31 carbon chain was connected to the calixarene through click chemistry to project in the space the saccharide portion of the GM1 ganglioside, natural substrate of CT. The inhibition activity of this macrocyclic ligand resulted very high with a gain of 20 000 fold per single GM1 ganglioside.

Small libraries can be also prepared with glycocalixarenes having different shape, valency in order to evaluate the best multivalent ligand towards the target protein. Once defined the

proper ligand, with the achievement of a best multivalent presentation of carbohydrate units towards the target protein, it can be studied more extensively.

The possible role played by different structural characteristics of these glycocalixarenes was for example observed by Andrè et al[49], when a family of 14 glycosylthioureidocalix[n]arenes was prepared to obtain effective inhibitors of human galectins, important biological targets since involved in tumor progression. The designed ligands, some bearing galactose units some lactose, had different properties in terms of valency, ligating units and geometry of the epitope display. It was observed a correlation between their inhibition potency and their structural features. In particular for lactose derivatives, resulted in general more efficient and interesting than those with galactose it was found that more conformationally mobile calix[6]- and calix[8]arenes were highly efficient in the inhibition of one type of galectin, as well as the 1,3-alternate calix[4]arene. Particularly the cone analogue, with facial display of the saccharide units, demonstrated to be poorly active against gal-1, but more reactive towards gal-3. Interestingly the 1,3-alternate analogue had the opposite behaviour in the inhibition potency.

Another example is given by Vidal and co-workers who prepared a series of galactosyl glycocalix[4]arenes against PA-IL lectin (Lec A) of *Pseudomonas Aeruginosa,* to create potential diagnostic and therapeutic agents towards this bacterium. Again these macrocycles presented different geometries and valency. The most efficient derivative resulted the 1,3-alternate calixarene, since it was able to form a complex with the lectin characterized by a stoichiometry of 1:4 between the calixarene and the monomeric form of PA-IL. This meant that all the four galactose units were simultaneously involved in the recognition of a lectin monomer for each one.[50]

All the reported examples show the relevance of calixarenes as interesting ligands for the recognition and inhibition of proteins. Several of these biologically active macromolecules, with different structures and activity were shown to interact with calixarenes with high affinity and selectivity through surface recognition and interactions with binding site or single residues. This is due to the synthetic versatility of these macrocycles, the possibility of controlling their conformational properties and three- dimensional shape, their multivalent nature.

Definitely in literature are present many other works where the versatility of these macrocyclic ligands is used for the creation of systems for the protein targeting. It is possible to say that these interesting scaffolds can be used as platform for the design of new bioactive ligands, carriers, sensors, activators and stabilizing agents.

1.5 References

(1) Medzhitov, R. *Nature* **2007**, *449* (7164), 819–826.

(2) Pawson, T.; Nash, P. *Genes Dev.* **2000**, *14* (9), 1027–1047.

(3) Norton, J. D. *J. Cell Sci.* **2000**, *113*, 3897–3905.

(4) Ringli, C.; Keller, B.; Ryser, U. *Cell. Mol. Life Sci.* **2001**, *58* (10), 1430–1441.

(5) Wilkinson, D. *FASEB J.* **1997**, *11* (14), 1245–1256.

(6) Liszewski, M. K.; John, P. *Mol. Microbiol.* **1997**, *25* (4), 639–647.

(7) Morin, P. J. *Cancer Res* **2005**, *65* (21), 9603–9607.

(8) Varambally, S.; Dhanasekaran, S. M. *Nature* **2002**, *419*, 388–390.

(9) Peczuh, M. W.; Hamilton, A. D. *Chem. Rev.* **2000**, *100*, 2479–2494.

(10) Casnati, A.; Sansone, F. *Calixarene ligands biomacromolecules recognition, Ref. Modul. Chem. Mol. Sci. Chem. Eng., Redijik, J. ed. Elsevier Waltham, MA, 2015; doi10.1016/B978-0-12-409547-2.10827-3*.

(11) Sheng, C.; Dong, G.; Miao, Z.; Zhang, W.; Wan, W. *Chem. Soc. Rev.* **2015**, *44*, 8238–8259.

(12) Yin, H.; Hamilton, A. D. *Angew. Chemie - Int. Ed.* **2005**, *44* (27), 4130–4163.

(13) Cohen, F. E.; Kelly, J. W. *Nature* **2003**, *426*, 905–909.

(14) Chen, L.; Yin, H.; Farooqi, B.; Chen, L.; Yin, H.; Farooqi, B.; Sebti, S.; Hamilton, A. D.; Chen, J. *Mol Cancer Ther* **2005**, *4* (6), 1019–1025.

(15) Kim, J. E.; Lee, M. *Biochem. Biophys. Res. Commun.* **2003**, *303*, 576–579.

(16) Kiessling, L. L.; Gestwicki, J. E.; Strong, L. E. *Angew. Chemie - Int. Ed.* **2006**, *45*, 2348–2368.

(17) Qureshi, S. A.; Kim, R. M.; Konteatis, Z.; Biazzo, D. E.; Motamedi, H.; Rodrigues, R.; Boice, J. A.; Calaycay, J. R.; Bednarek, M. A.; Griffin, P.; Gao, Y.; Chapman, K.; Mark, D. F. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (21), 12156–12161.

(18) Fasting, C.; Schalley, C. a.; Weber, M.; Seitz, O.; Hecht, S.; Koksch, B.; Dernedde, J.; Graf, C.; Knapp, E. W.; Haag, R. *Angew. Chemie - Int. Ed.* **2012**, *51* (42), 10472–10498.

(19) Ercolani, G. *J. Am. Chem. Soc.* **2003**, *125* (51), 16097–16103.

(20) Badjić, J. D.; Nelson, A.; Cantrill, S. J.; Turnbull, W. B.; Stoddart, J. F. *Acc. Chem. Res.* **2005**, *38* (9), 723–732.

(21) Jencks, W. P. *Proc. Natl. Acad. Sci. U. S. A.* **1981**, *78* (7), 4046–4050.

(22) Kitov, P. I.; Bundle, D. R. *J. Am. Chem. Soc.* **2003**, *125* (52), 16271–16284.

(23) Mammen, M.; Choi, S.-K.; Whitesides, G. M. *Angew. Chemie, Int. Ed. English* **1998**, *37* (20), 2754–2794.

(24) Kiessling, L. L.; Gestwicki, J. E.; Strong, L. E. *Curr. Opin. Chem. Biol.* **2000**, *4* (6), 696–703.

(25) Mendoza, J. De; Martos, V.; Castren, P. *Curr. Opin. Chem. Biol.* **2008**, *12*, 698–706.

(26) Gutsche, C. D. *Calixarenes - an introduction*; 2008; Vol. 34.

(27) Sansone, F.; Baldini, L.; Casnati, A.; Ungaro, R. *New J. Chem.* **2010**, *34*, 2715–2728.

(28) Lehn, J.-M.; Lehn, J.-M. *Chem. Soc. Rev.* **2007**, *36* (2), 151–160.

(29) Gutsche, C. D. *Calixarenes 2001* **2001**, 1–25.

(30) Casnati, A.; Sansone, F.; Ungaro, R. *Acc. Chem. Res.* **2003**, *36* (4), 246–254.

(31) Gaeta, C.; Troisi, F.; Talotta, C.; Pierro, T.; Neri, P. *J. Org. Chem.* **2012**, *77* (7), 3634–3639.

(32) Arena, G.; Casnati, A.; Contino, A.; Lombardo, G. G.; Sciotto, D.; Ungaro, R. *Chem. - A Eur. J.* **1999**, *5*, 738–744.

(33) Sheinerman, F. B.; Norel, R.; Honig, B. *Curr Opin Struct Biol.* **2000**, *10* (2), 153–159.

(34) Keskin, O.; Gursoy, A.; Ma, B.; Nussinov, R.; Aviv, T. *Chem. Rev.* **2008**, *108*, 1225–1244.

(35) Sebti, S. M.; Hamilton, A. D. *Oncogene* **2000**, *19*, 6566–6573.

(36) Hamuro, Y.; Calama, M. C.; Park, H. S.; Hamilton, A. D. *Angew. Chemie Int. Ed. English* **1997**, *36* (23), 2680–2683.

(37) Mc Govern, R. E.; Fernandes, H.; Khan, A. R.; Power, N. P.; Crowley, P. B. *Nat. Chem.* **2012**, *4*, 527–533.

(38) Taylor, P.; Arena, G.; Cali, R.; Lombardo, G. G.; Rizzarelli, E.; Sciotto, D.; Ungaro, R.; Casnati, A. *Supramol. Chem.* **1992**, *1*, 19–24.

(39) McGovern, R. E.; Snarr, B. D.; Lyons, J. a.; McFarlane, J.; Whiting, A. L.; Paci, I.; Hof, F.; Crowley, P. B. *Chem. Sci.* **2015**, *6* (1), 442–449.

(40) Martos, V.; Bell, S. C.; Santos, E.; Isacoff, E. Y.; Trauner, D.; Mendoza, J. De. *Chem. Biol.* **2009**, *106* (26), 1–5.

(41) Salvatella, X.; Martinell, M.; Gairí, M.; Mateu, M. G.; Feliz, M.; Hamilton, A. D.; De Mendoza, J.; Giralt, E. *Angew. Chemie - Int. Ed.* **2003**, *43*, 196–198.

(42) Gordo, S.; Martos, V.; Santos, E.; Menéndez, M.; Bo, C.; Giralt, E.; de Mendoza, J. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 16426–16431.

(43) Lundquist, J. J.; Toone, E. J. *Chem. Rev.* **2002**, *102* (2), 555–578.

(44) Chabre, Y. M.; Roy, R. *Design and creativity in synthesis of multivalent neoglycoconjugates.*; 2010; Vol. 63.

(45) Pieters, R. J. *Org. Biomol. Chem.* **2009**, *7* (10), 2013.

(46) Sansone, F.; Casnati, A. *Chem. Soc. Rev.* **2013**, *42*, 4623–4639.

(47) Sansone, F.; Casnati, A. *Chem. Soc. Rev.* **2013**, *42* (11), 4623–4639.

(48) Garcia-Hartjes, J.; Bernardi, S.; Weijers, C. a G. M.; Wennekes, T.; Gilbert, M.; Sansone, F.; Casnati, A.; Zuilhof, H. *Org. Biomol. Chem.* **2013**, *11* (26), 4340–4349.

(49) André, S.; Sansone, F.; Kaltner, H.; Casnati, A.; Kopitz, J.; Gabius, H. J.; Ungaro, R. *ChemBioChem* **2008**, *9* (10), 1649–1661.

(50) Cecioni, S.; Lalor, R.; Blanchard, B.; Praly, J. P.; Imberty, A.; Matthews, S. E.; Vidal, S. *Chem. Eur. J.* **2009**, *15*, 13232–13240.

Chapter 2

Mixed glycocalix[4]arene-DOPC liposomes for targeted drug-delivery

2.1 Introduction

2.1.1 Carbohydrates and their interactions with proteins

Among natural ligands involved in important recognition processes with proteins, which use to this aim specific binding sites, there are carbohydrates.[1] These in fact have an important role in Nature not only as energetic reserve and structural motif in cells but also as substrates whose complexation is at the basis of many biological processes.

The relevant role played by carbohydrates is significantly related with their structural complexity that excides that of amino acids and nucleotides. Each monosaccharide in fact presents several hydroxyl groups with different orientations in the space and each of them can in principle be engaged in a glycosidic bond with the anomeric position of another sugar. Additionally, each glycosidic bond can be in α- or β-configuration (see figure **1**). Thus, carbohydrates are capable of forming many different combinatorial structures, including highly branched ones.

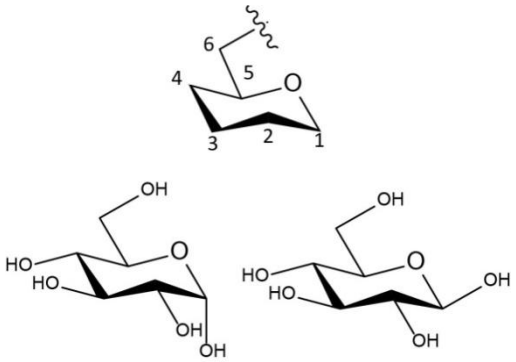


Fig. 1: different attachment points in a carbohydrate and stereochemistry on the anomeric carbon

Moreover, different functionalization of the hydroxyl groups, such as acetylation, phosphorylation, sulfation, can further increase the complexity of these molecules. For these reasons Nature uses carbohydrates as recipient of information, as an alphabet, necessary for many physiologically important functions such as differentiation, growth, cell-cell recognition, and pathological events like toxins, bacteria, viruses adhesion to cells, tumor progression and diffusion of metastases.

This enormous structural variability represented in the glycome, the equivalent of the genome or proteome, associated with so many information generated the so-called “Sugar

code"[2], that contains detailed levels of biomedical information which can be decrypted by suitable carbohydrates binding proteins.

Glycobiology[1,3] is the interdisciplinary science, that combining chemistry, biology and medicine studies all the aspects of glycans and in particular how they are involved in these recognition processes.[4]

2.1.2 Carbohydrate Binding Proteins: the Lectins

Among the carbohydrate binding proteins (CBPs) lectins are the translators of the Sugar code. Lectins are proteins that act as receptors for saccharides belonging to glycolipids and glycoproteins,[5,6] to oligo- and polysaccharides but also to free-floating glycans including monosaccharides. Lectins do not possess enzymatic activity and are able to bind mono- and oligo-saccharides with high specificity and without cause chemical modifications. These proteins are also not products of an immune response but they descript the information hold in the sugar units to convert them in signals and give rise to biochemical processes. In fact the interaction of lectins with particular carbohydrates can be as specific as the interaction antigen-antibody or substrate-enzyme.[7]

Lectins are widespread in nature: they were initially found and described in plants, but in following years lectins were isolated from animals and also from microorganisms such as viruses and bacteria.[8] Interestingly plant and animal lectins show no primary structural homology and they demonstrate similarities in the preferential binding to carbohydrates, suggesting that plant and animal lectin genes may have co-evolved. However, during the past several years, many primary and three dimensional lectins structures have been elucidated. It was observed that lectins from different sources lacked primary sequence similarity but in their tertiary structures they shared similarities.[8] There are regions in these proteins, the so called carbohydrate recognition domains (CRDs) that typically recognize the branch ends of cell membrane glycoconjugates.[9] To the selectivity for the substrate, lectins add the ability of discriminating between anomeric isomers of the same sugar. For example Concanavalin A (Con A), one of the lectins known since longer time and widely used in laboratory for its commercial availability and cheap prize, specifically binds the α-anomer of glucoside and mannoside in natural substrates, but not the β-anomer of either[9]. For these

reasons these macromolecules were called lectins, that hails from latin word “legere” that means “to choose” or “to select”.

The broad occurrence of lectins in different species, tissues or cells indicates the importance of these proteins in nature.

According to their origin, lectins are classified in plant, animal, and microbial lectins.[10]

Plant lectins have been more largely explored compared to the animal ones. They are distributed in many tissues and many plants contain them, including different food crops such as wheat, rice, potato, tomato, soybean and bean.[11] The classification was given by Van Damme et al.,[11] who divided the plant lectins into seven structurally and evolutionarily related protein families. More recently,[12] they introduced new proteins and redistributed the lectins in 12 families with different carbohydrate binding domains.

These lectins are often used for their biological value in preliminary experiments to verify the binding capacity toward designed molecules as potential ligands with biological activity. In this way they can be used, for instance, as reliable biochemical, cyto-chemical, and histochemical probes for the recognition of malignant cells.[9] Concanavalin A is a common plant lectin employed as a model for the determination of cell agglutination mediated by the carbohydrate-binding proteins.[13]

Animal lectins[14] are divided in five groups, based on the monosaccaride for which they exhibit the highest affinity: 1) N-acetyl-D-glucosamine, 2) N-acetylneuraminic acid, 3) D-galactose/ N-acetyl-D-galactosamine, 4) D-glucose and 5) D-mannose. This classification is however not so exhaustive and correct since several lectins have a significantly high specificity not towards simple monosaccharides but only to di-, tri- or tetra-saccharides. For these reasons, lectins are preferably classified on the basis of sequence homologies. The accepted classification evaluates the CRD of these proteins and in this way they result divided in six families: C-type, S-type, I-type, P-type, R-type and F-type.[14]

The differences between these lectin families are related to the cation-dependent binding capacity. For example C-type A are large, asymmetric, have one or more CRDs and exist as Ca^{2+} dependent proteins found in secreted or bound forms.[8] On the contrary, the S-type lectins are generally small, non-glycosylated, soluble and are Ca^{2+} independent proteins found intracellularly and extracellularly.[15]

C-type lectins (CTLs) are the most abundant among animal lectins, and their superfamily is grouped into three families: selectins, collectins and endocytic lectins. For instance selectins are a group of carbohydrate-binding proteins residing in cell membranes, that are known to mediate the initial recognition of immunologically important events such as lymphocyte routing.

Among the S-lectins the most important are galectins.[16,17] They show a consistent amino acid sequence, constituted by approximately 130 residues for the CRD. Galectins recognize selectively glycoconjugates exposing β-galactoside residues. They are important biomedical targets since are involved in cancer growth and diffusion, contributing to tumour cell survival, angiogenesis and tumour metastasis.

Also microbial lectins are widespread in Nature and play an important role that often is at the origin of physiological and pathological events into cells. Numerous bacterial strains produce surface lectins, commonly in the form of fimbriae and pili that are filamentous assemblies of protein subunits. Even if this is not in compliance with the definition of lectins[18], but because they are sugar specific and of non immune origin, it is convenient to call them in this way.

It is evident that lectins are important biomedical targets, because of their involvement in many biological processes such as cell-cell communication, cell-cell self-recognition, cell-extracellular matrix (ECM) interactions, cell growth[19], cell differentiation, cell signalling, cell adhesion and migration, apoptosis[20], host-pathogen interactions. In plants they are involved in the protection against pathogens.[21] In microbial organisms, lectins are often involved in host recognition and tissue adhesion in infections. Some of those may act as virulence factors by enhancing the bacterial adherence to the cells and by increasing phagocytosis.[22] Since prerequisite for the initiation of the majority of infectious diseases is the adhesion of the pathogenic organisms to host tissues lectins they have increased the attention of the scientific community .

For these reasons, the inhibition of the interactions between lectins and their counterparts is considered a relevant way to develop new biologically active compounds. But the specificity of interaction and the wide localization make lectins also an interesting target for site specific drug delivery and, on the other site, identify the sugars as important tools to make smart a drug delivery system. It is in fact possible to think about microbial lectins as the target to address a drug against the pathogen and think about cell surface lectins of humans

as the target to exploit for the selective transport of a therapeutic agent towards specific organs, tissues or degenerated cells.

Independently from their origin, many lectins have a structure presenting several domains for the recognition of an equivalent number of units of the same glycoside or tend to oligomerize forming self-assembled species with multiple CRDs. This means that lectins act exploiting multivalency. Then, the design and synthesis of multivalent ligands appears as a advantageous strategy to develop molecules able to efficiently and selectively bind, inhibit, target these proteins.

2.1.3 Sugar transporter proteins: GLUT's family

There are also proteins, not belonging to the class of lectins, that are important for the transport of small polar molecules inside the cells, since cell membranes are essentially impermeable to these solutes, such as sugars.

Since glucose is the principal energy source for mammalian cells, the presence of membrane proteins, able to vehicle this sugar from the external environment, is essential to maintain normal functions. One of the families of glucose transporters is the GLUT gene family.

This class of integral membrane protein transporters facilitates passive diffusion of glucose across tissue barriers by energy-independent stereo-specific mechanisms.[23] For example transfer across the blood- brain barrier (BBB) and into brain cells is mediated by members of this family.[24] Several proteins belong to this family and they have been detected in brain, liver, muscle/fat, and small intestine.[25] Particularly there are 7 isoforms (GLUT1-GLUT7), and among them, facilitative glucose transporter protein type 1 GLUT1 is known as a basic, high-affinity glucose transporter. GLUT1 is an integral membrane protein, highly expressed in erythrocytes and brain,[26] but it is also present in almost every tissue. This is the most extensively studied of all mammalian membrane transporters.[27]

It is important to say that enhanced tumor uptake of glucose is facilitated by the overexpression of these glucose transporter proteins, observed widely in tumor tissues. Elevated GLUT1 expression has been described in many cancers, including hepatic, pancreatic, breast, esophageal, brain, renal, lung, ovarian, and cervical carcinoma. [28] These different mammalian glucose transporters represent important reagents that will provide

better understanding of glucose uptake and metabolism in normal and altered metabolic states. If fact they can be used to understand the internalization levels of glucosilated molecules in tumoural cells. This is important in the drug delivery field, in order to explore which systems are better internalized, reaching easily the targets, thanks to the recognition process between those derivatives and these receptors.

2.1.4 Drug delivery and liposomes

Drug delivery is a field of vital importance to medicine and healthcare. Controlled drug delivery improves bioavailability by preventing premature degradation and enhancing uptake, maintains drug concentration within the therapeutic window by controlling the drug release rate. With the possibility to create "smart" systems able to target drugs to disease site and target cells, the side effects are also reduced.[29]

There are several nanocarriers that can be use for drug delivery purposes, such as micelles, nanoparticles, polymers. One of the most important classes is that of liposomes. Alec D. Bangham, in the 1964, at the Babraham Institute, University of Cambridge discovered for the first time liposomes.[30] They are vesicles characterized by the presence of a lipid bilayer and an internal aqueous core (figures **2a** and **2b**). Their formation is due to the amphiphilic character of acylic chains that in the presence of aqueous solutions start to aggregate in polar shells.

Liposomes can have different size and composition depending on which approach is used for their preparation. They are divided in 5 group, based on lamellarity (uni-, oligo-, multilamellarity), composition and size (small, intermediate, large), as explained in **Table 1**.

Type of liposomes	Size (nm)	composition
Small unilamellar vescicles (**SUV**)	20-100	Traditional liposomes
Large unilamellar vescicles (**LUV**)	>100	Long-circulating liposomes
Giant unilamellar vescicles (**GUV**)	>1000	Cationic liposomes
Oligolamellar vescicles (**OLV**)	100-1000	Stimuli sensitive liposomes
Multilamellar vescicles (**MLV**)	>500	Immunoliposomes

Table 1: classification of liposomes based on lamellarity, composition and size.

For example the formation of unilamellar or multilamellar vesicles depends on the synthetic

methods and postformation processing used for their preparation.[31] The differences in the ability to capture drugs, the release rate are all features based on the number of lipidic bilayers. For instance the ULVs are more suitable for hydrophilic drugs, whereas MLVs preferentially encapsulate lipophilic drugs.[32] Additionally ULVs have smaller entrapped volume than MLVs, but they possess a faster release rate compared to MLVs.[31]

Another characteristic of liposomes is that they can be also positively, negatively or neutrally charged depending on the presence of suited additives in their formulation. Furthermore the addition of these substances, such as surfactants, can influence also their encapsulation and release properties. Even cholesterol can influence the preparation: normally it is added to reinforce more the lipidic bilayer, raising the encapsulation properties.

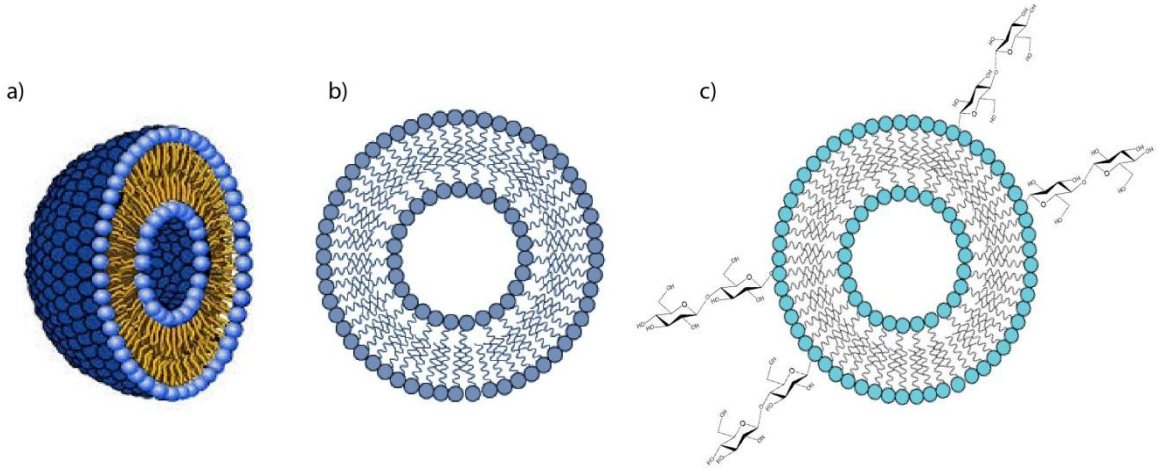


Fig. 2: schematic representation of a liposome section: a) 3D and b) 2D visualization and c) glycoliposome

One of the potentiality of these systems, as said above, is exactly their use as nanocarriers for transport of drugs.[33] The discovery of liposomes employment as delivery systems occurred in the 1970s.[34] These vesicles are a very important tool in the creation of drug delivery systems because of their structural features. In fact they can incorporate both hydrophilic and hydrophobic drugs. The possibility to vehicle biologically active compounds is important since many drugs, including nucleic acid-based drugs, have a lack in stability, solubility and controlled toxicity in physiological conditions. [35]

Actually liposomes presented several advantages[36], summarized below:

- improvements in the solubility of incorporated drugs;

- toxicity and side-effects reduction of the encapsulated drug;
- prevention of chemical and enzymatic degradation, immunological inactivation of drugs under storage conditions and patient administration;
- possibility to functionalize the surface with ligands for the creation of targeted drug-delivery systems.

Furthermore they are considered safe, although the addition to their assembly of non physiological molecules can introduce additional factors that in turn can change the toxicity profile.

However it is important to keep in mind that drugs entrapped in liposomes become bioavailable when there is release. Hence the ability of liposomes to increase the local bio-accessible drugs is obtained when the release rate of the entrapped drug is optimized.[37] Furthermore liposomes must maintain their structure intact until the cell/organ target is reached; hence the enhancement of rigidity of the lipid bilayer can help to create more resistant drug-delivery systems.

Definitively it is well-known that liposomes functionalized with bola-amphiphile molecules yield less permeable and more rigid membranes respect to those composed of monopolar lipids.[38,39] A bola-amphiphile[40] is, for definition, a compound that have two polar heads covalently connected through one or more hydrophobic chains.

The main natural source of these type of molecules are *Archaebacteria*,[41] whose membrane lipids have a bolaamphiphiles shape (as reported in the figure below).

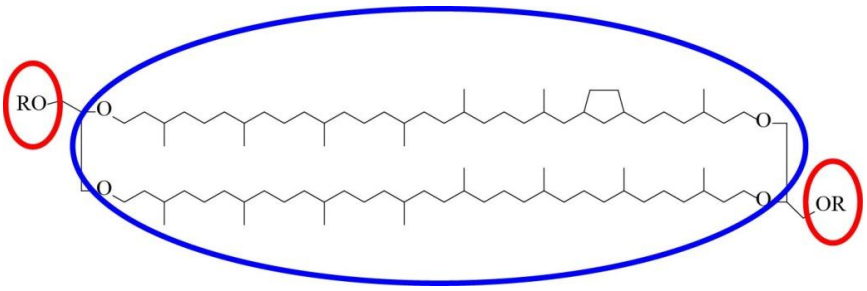


Fig. 3: example of bolaamphiphiles from archaebacteria; in red ellipses polar heads and in blue ellipse the hydrophobic chains, respectively.

The extreme physical conditions to which they are able to live, such as very high temperatures (up to 90 °C) and salinity, low pH (1-1.5) or absence of oxygen are correlated to the presence of these bolaamphiphiles compounds in their membrane.[42] This unique

combination of properties has raised the attention for the potential use of bolaamphiphiles as membrane-stabilizing agents for the preparation of more stable drug delivery vesicles in physiological conditions.[43,44] For example the liposomial formulations can carry the drugs for a longer period, reaching the cell/organ targets without loss of efficacy.

As mentioned above, another point to consider is the importance to create nanocarriers that can reach the desired tissues, producing higher and more selective therapeutic activity. For that reason it has been a goal for the scientists to develop liposomes that can recognize selectively the targets with the help of specific ligands or receptors on their surface.[45] These are the so called actively-targeted or targeted-drug delivery liposomes. The role of the ligands is owed mainly when liposomes arrive at the target tissue, where they are able to bind the carbohydrate receptors and subsequently internalize into the cell (see figure **4**).

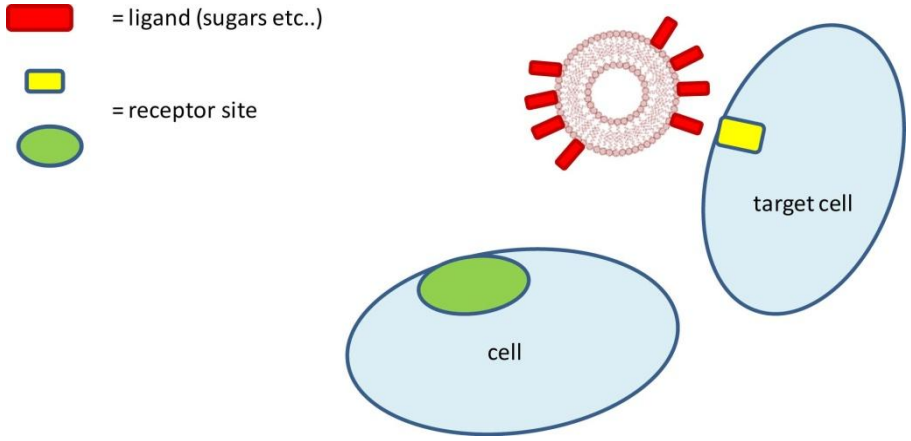


Fig. 4: binding of ligand liposomes on the target cells: systems are able to selectively recognize and bind to the complementary receptors on the target cell surface.

Therefore, it is necessary for these systems to have a longer blood circulation times by allowing the liposomes to evade immune detection as well as enhanced binding efficiencies respect to the natural ligands.[33]

The possibility to employ multivalent ligands in the functionalization of liposomes could help in the interaction phenomena with protein receptors. In fact multivalent interactions tend to be much stronger than the corresponding monovalent ones[46] and multiplying the number of existing interactions the binding affinities increase.

Sugars can be used as active units exposed on the liposome exterior surface (as schematically reported in figure **2c**). In fact, since carbohydrate-proteins interactions are fundamental for many biological processes, the idea to use these interactions as strategy to develop targeted drug-delivery systems can have therapeutic applications.[47,48] Examples in literature of glycoliposomes[49] are already known.[50] One is that of mannose-coated liposomes were found to be compatible in pulmonary tissue alveolar macrophages, since are specifically recognized by mannose receptors present in phagocytes.[51] Another interesting approach has been the use of glycosylated liposomes to target lectins.

For instance Kitano and co-workers created DMPC liposomes with amphiphile sugars, with long alkyl chains, able to fit the bilayer. These systems demonstrated their ability to recognize Con A, with very high K_a due to multivalent interactions between lectin and polyglycosylated liposome.

Analogously, Mauceri et al.[52] demonstrated the ability of a glucosylated amphiphile of interacting specifically with Con A when included in DMPC (,2-dimyristoyl-sn-glycero-3-phosphocholine) liposomes.

Nonetheless carbohydrates can be used also as polar head for the creation of bola-shaped molecules.[53,54] There are indeed some examples of liposomes functionalized with carbohydrate-based bolaamphiphiles,[38] with the production of stable and efficient drug delivery systems. In fact the combination of the structural properties of bolaamphiphiles with the selective binding properties of carbohydrate-based ligands can have an important application in nanomedicine.

In this context, few years ago in a joint project between the group in Parma, where I performed my PhD research project, and the group of Dr Giovanna Mancini at CNR in Rome, it was designed and synthesized the glucocalix[4]arene **1** (fig. **5a**)[55] with the idea of using it for the functionalization of liposomes.

Aleandri et al.[55] described the preparation of the derivative **1**, blocked in a 1,3-alternate geometry, intentionally selected. This conferred to the macrocycle the shape of a bolaamphiphile and the glucose units, introduced as polar heads, potentially provided the compound with interaction properties towards glucose binding proteins.

It could be possible to prepare stable mixed liposomes by combining 1,2-Dioleoyl-*sn*-glycero-3-phosphocholine (DOPC) and **1** in 90/10 percentage. They have a structure as that schematically reported in Figure **5b**. This represented the first example of the

functionalization of liposomes with saccharides by exploiting glycocalixarenes.

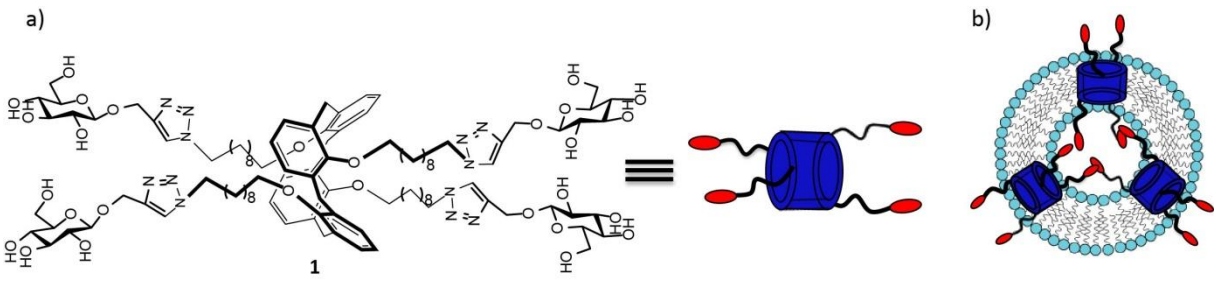


Fig. 5: a) 1,3-alternate glucocalix[4]arene [55] and b) schematic representation of DOPC/**1** mixed liposomes

These systems showed enhanced properties compared to liposomes simply made with DOPC. Effectively the calixarene behaved as a solidifying agent of the DOPC liposomes, conferring higher rigidity and stability. This was confirmed by permeability and entrapping measurements, with both lipophilic and hydrophilic drugs. In fact, the presence of **1** in the DOPC lipid bilayer reduced the permeability of the liposomes and the mobility freedom of an embedded lipophilic fluorescent probe. The hydrophobic m-THPC was effectively entrapped with a raised efficiency of 35% in DOPC/**1** mixed liposomes respect to DOPC liposomes. The reorganization of a lipid bilayer, due to the presence of **1**, gave rise also to a remarkable increase of the encapsulation efficiency of hydrophilic calcein in the inner pool. The features of liposomes resulted then improved. Moreover, these systems were capable to interact with the plant lectin Concanavalin A (Con A), with multivalent and strong interactions.

These first encouraging results with the mixed DOPC/glucocalixarene liposomes, brought us to develop the project as discussed in detail below, preparing another glycocalixarene, analogue of **1**, equipped with cellobiose units and a monomeric model of **1.**

The monomeric analogue (fig. **6**) was synthesized, in order to evaluate the possible importance of the macrocyclic structure of the calixarene in the formation of these liposomes and in determining their improved properties. The other glycocalixarene **2** (reported in figure **6**) was prepared to have a compound able to project the β-glucose units at longer distance in the space from the liposome surface with a better exposition of the carbohydrate units and a possible facilitate approach to the cell surface receptors.

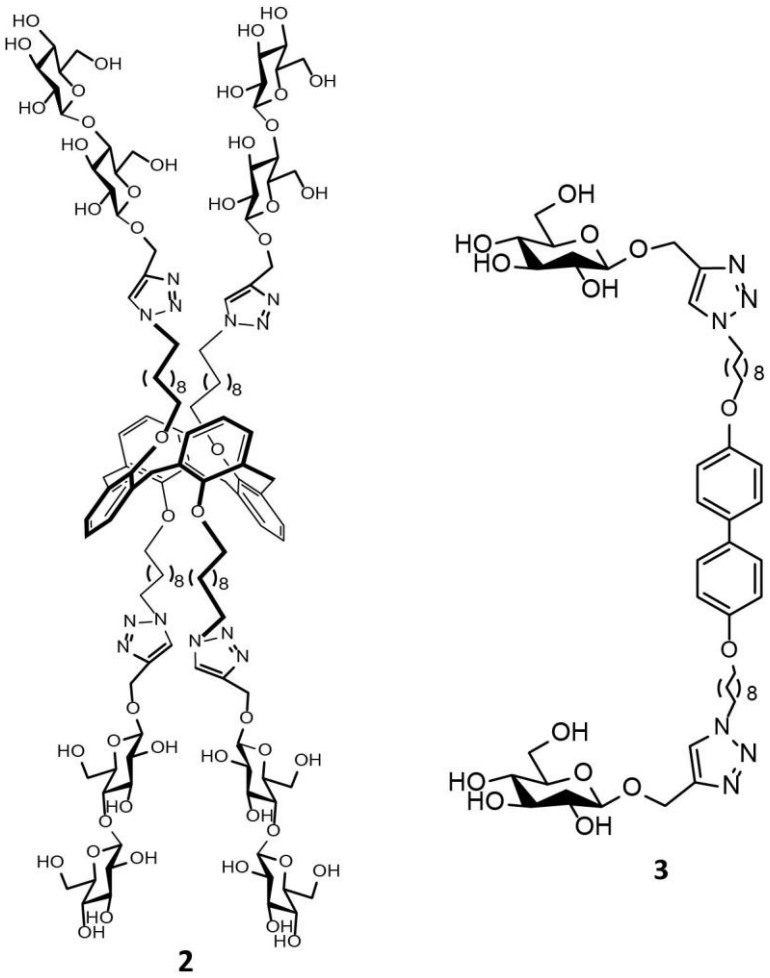


Fig. 6: new bolaamphiphiles cellobiose-based calix[4]arene **2** and non macrocyclic derivative **3**

2.1.5 1,3-dipolar cycloaddition: CuAAC

The chemistry of calixarenes is well consolidated and to attach the sugar units to the scaffold there are different synthetic methodologies.[56,57] Particularly the method that was used in the work of Aleandri and co-workers,[58] and even in this section for the conjugation of carbohydrate moieties to the calixarene scaffold, is the 1,3-dipolar cycloaddition reaction[59] between an alkyne and an azide. It is known also as Huigsen reaction,[60,61] with its "new" more regioselective Cu(I)-catalyzed azide-alkyne cycloaddition (known as CuAAC) [62]. This is one of the most important examples of the reactions of the so-called "click" chemistry,[63] defined by Sharpless et al. [63] as reactions that: ''are modular, wide in scope, high yielding, create only inoffensive by-products (that can be removed without chromatography), are stereospecific, simple to perform and that require benign or easily removed solvent''.

Many glycocalixarenes were prepared using this approach, [64,65] due to the high yield and the possibility to employ microwaves irradiation.[66]

Especially the Copper-catalyzed azide alkyne- cycloaddition [67] has raised the attention for the preparation of carbohydrate-base conjugates. This is a modification of the Huisgen pericyclic reaction, since this is not so selective, giving rise to 1,2,3-triazoles as a mixture of 1,4- and 1,5- substituted regioisomers. Sharpless discovered that the used of Cu(I) salts as catalysts for 1,3-dipolar cycloaddition reactions, gave only the 1,4-triazolic isomer.[63]

This procedure was then used very frequently for the covalent assembly of complex molecules[68] and also with multivalent ligands and multiglycosilated structures.[69,64]

Not only advantages in the synthetic procedure can be considered: in fact the presence of triazole units in the molecules gives some attractive properties. This group has a strong dipolar moment and has aromatic character, is capable to form hydrogen bonds, that are important for the solubility in aqueous solution of the final compounds. Furthermore it has high chemical stability, due to the inertness to strong oxidizing, hydrolytic and reducing conditions.[70]

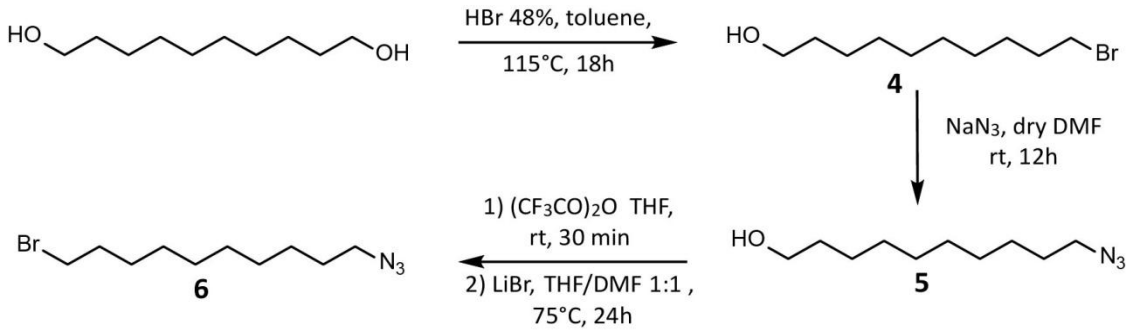
2.2 Results and discussions

Due to the previously described good results on calixarene **1**, the new bolaamphiphile calixarene **2** was synthesized. The idea was to keep constant the scaffold and the linker, since the hydrophobic part was able to interact with the lipidic bilayer and afford more rigid liposomes. The longer saccharidic units, from glucose to cellobiose, was selected for a better exposition of the glucose units from the external liposomial membrane to possibly increase the strength of the interaction with Con A already evidenced with glucocalixarene **1**. We have to say to this end that, although Con A is in nature selective towards α-glucosides (or α-mannosides), in the studies previously performed with compound **1** the β-glucoside units were shown to interact with the lectin. Also in other cases,[71,72] β-glucocalixarenes of different types showed the same ability. These results were in agreement with the data collected by other researchers who could conclude that, when the saccharide epitope presents non natural spacer at the anomeric position for the linkage to a scaffold, the natural anomer selectivity of the corresponding lectins can be missed.[73] Then, to avoid the

introduction of further modifications respect to **1**, for the new glycocalixarene we maintained the anomeric configuration of the terminal glucose units by using cellobiose.

2.2.1 Synthesis of the linker and conjugation to the calixarene scaffold

Considering the thickness of the lipidic bilayer, in the previous work on these glycocalixarene containing liposomes a 10-carbon linker was used that were re-synthesized for the new studies.[55] Starting from 1,10-decandiol (**Scheme 1**), the first reaction was the monobromination[74,75] of the hydroxyl group, using HBr (48%in aqueous solution), allowing the formation of derivative **4**. This was the main product and to reduce the formation of di-bromide compound, the ratio HBr/1,10-decandiol 1:1 was indeed used. Nucleophilic substitution of the bromide with azide was then performed using NaN_3 in dry DMF and compound **5** was obtained. In the ^{1}H-NMR spectrum the shift at higher fields of methylene protons in α position respect to azide confirmed that substitution took place. The hydroxyl group of 10-azidodecanol **5** was then converted into bromide, giving rise to the final linker **6** through a two steps reaction. The OH group was transformed in good leaving group by reaction with trifluoroacetic anhydride, and as a consequence the carbon atom could be attacked by the nucleophile (Br^- from LiBr). The two ends of **6** bore at that point two complementary functional groups, Br for the alkylation of the calixarene scaffold and azide for the functionalization with the sugar units.

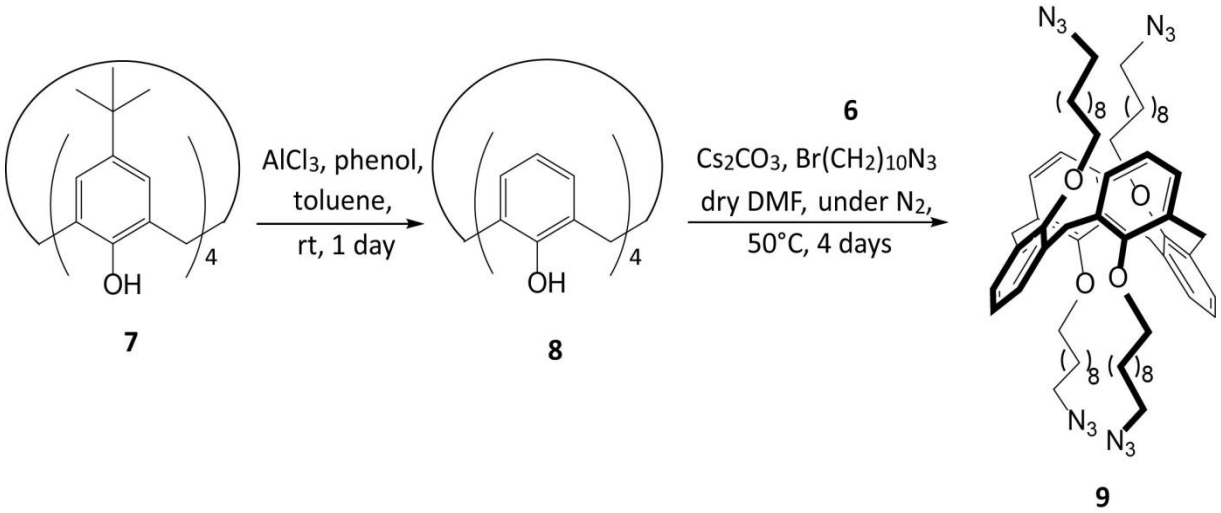


Scheme 1: synthesis of the 10-carbon linker for the alkylation of calixarenes and monomer.

The further step was the connection of the linker **6** to the calix[4]arene **8** (**Scheme 2**) obtained from **7** by removal of the *t*-butyl groups from the upper rim **8** was functionalized with the 10-carbon chain **6** using the typical conditions necessary to achieve the 1,3-alternate structure of the fully alkylated macrocycle **9**. Cesium carbonate was then used as

base because of the well-known templating effect of the cation that establishes interaction with two opposite aromatic rings forcing the calixarene to the 1,3-alternate geometry. Nevertheless, in part because of the low reactivity of such long chains, also the derivative with partial cone structure formed as by-product during alkylation but it was completely separated from the desired product by flash column chromatography on silica gel.

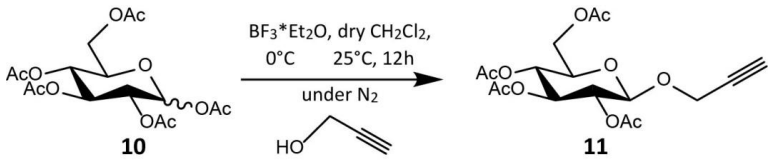
The 1,3-alternate geometry of derivative **9** was confirmed by the presence in the ^{1}H-NMR spectrum of a singlet at 3.59 ppm for the methylene bridge, typical of this geometry, corresponding to 8 hydrogen atoms.



Scheme 2: synthesis of calix[4]arene scaffold in 1,3-alternate geometry **9** from **7**

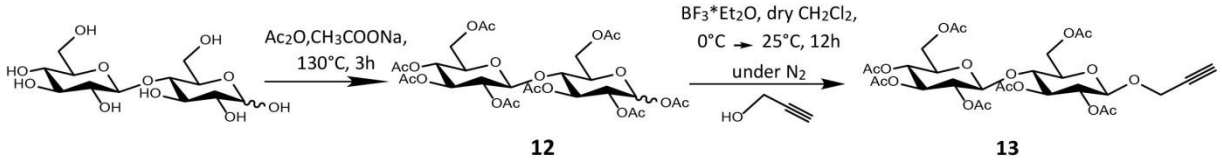
2.2.2 Functionalization of the saccharide units

Derivative **9** resulted in this way equipped with four terminal azide groups necessary for the copper catalysed cycloaddition reaction with an alkyne moiety located on sugars. To introduce this functional group, glucose derivative **11** was prepared from peracetylated glucose **10** (**Scheme 3**) using BF_3*Et_2O as promoter for this glycosidation step with propargyl alcohol. The presence of the participating protecting group in position 2 drove the reaction towards the β-anomer as product. Furthermore the reaction temperature was initially kept at 0 °C and only then left to slowly reach room temperature in order to further favour the formation of the β-anomer. Actually, the reaction produced only β-glucoside **11**, obtained in 95% yield, without further purifications. The structure of the desired product was confirmed by the J value of 9 Hz for H_1, due to the axial-axial coupling with H_2.



Scheme 3: synthesis of propargyl glucose **11** from D- peracetylate glucose **10**

The same sequence of reactions used for the synthesis of **11** was perfomed on cellobiose after its peracetylation (**Scheme 4**).



Scheme 4: synthesis of propargyl-β-cellobiose **13** from D-cellobiose

The peracetylation of D-cellobiose was carried out using acetic anhydride and sodium acetate at 130 °C, with the formation of **12** as a mixture of α and β anomer, in a 6:4 ratio.

In fact in the ^{1}H-NMR spectrum there are two doublets at 6.23 ppm and 5.65 ppm, corresponding to $H_{1\alpha}$ ($J_{ax\text{-}eq}$ = 5.6 Hz) and $H_{1\beta}$ ($J_{ax\text{-}ax}$ = 8.4 Hz), respectively.

The peracetylated cellobiose **12** was then functionalized with propargyl alcohol. The desired β-anomer was afforded in 97% yield and the ^{1}H-NMR spectrum confirmed the presence of **13**. As reported in figure **7**, the J constant of H_1 with H_2 was 7.8 Hz, typical of an axial-axial coupling and also there were the diagnostyc peaks of the terminal alkyne hydrogen at 2.45 ppm and of the methylene group at 4.33 ppm.

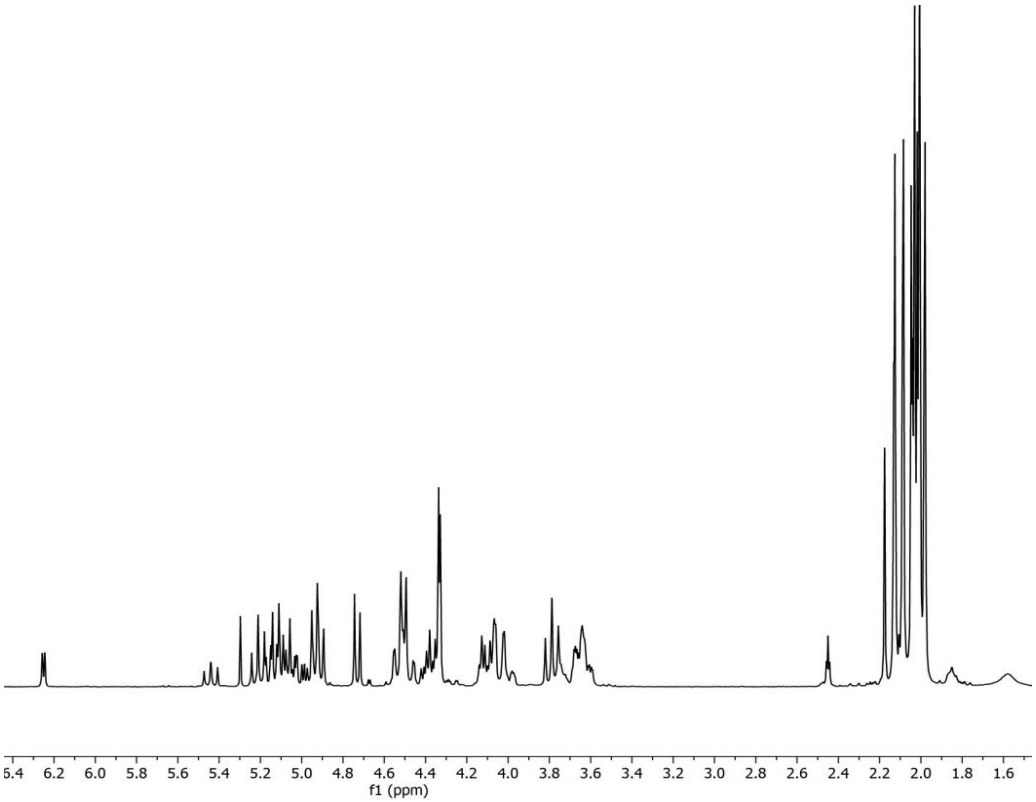
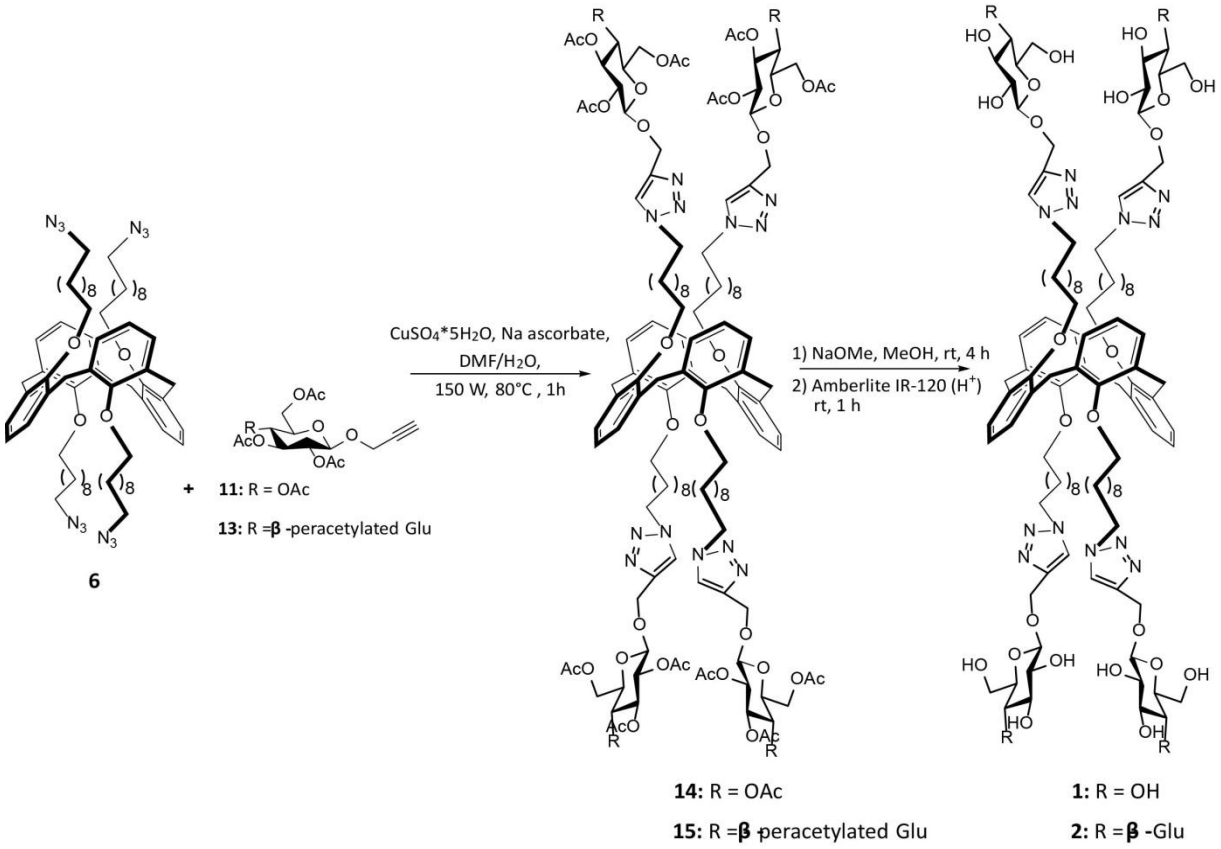


Fig. 7: ^{1}H-NMR spectrum (300 MHz, $CDCl_3$) of cellobioside **13**

The sugar units were then ready to react with calixarene **9** for the formation of final glycocalix[4]arenes **1, 2** and the monomer **3.**

2.2.3 Conjugation to the calixarene scaffold and deacetylation step

The conjugation of the sugar units to the macrocycles was carried out using the CuAAC "click" reaction. Both glucoside **11** and cellobioside **13** were then attached to the 1,3-alternate tetraazido-calix[4]arene **9**. On the basis of the previous experience on the synthesis of **1**, the reaction was performed under microwave irradiation to enhance reaction speed and yield. In fact, despite its "click" characteristics, in our case the reaction was not so efficient as expected when performed in traditional conditions. Probably, the long chains of the calixarene could roll up and make the azido terminal groups less available for the reaction. The tetraglycosylated calix[4]arenes **14** and **15** were obtained pure after flash column chromatography to remove the excess of sugars, since no evidences of partially functionalized calixarenes were found.



Scheme 5: synthesis of bolaamphihiles **1** and **2**.

The glucocalixarene **14** was well characterized previously[55] and the same compound re-synthesized presented the same spectroscopic data. On the other hand the cellobiose containing derivative **15** was new and then a complete characterization was needed.

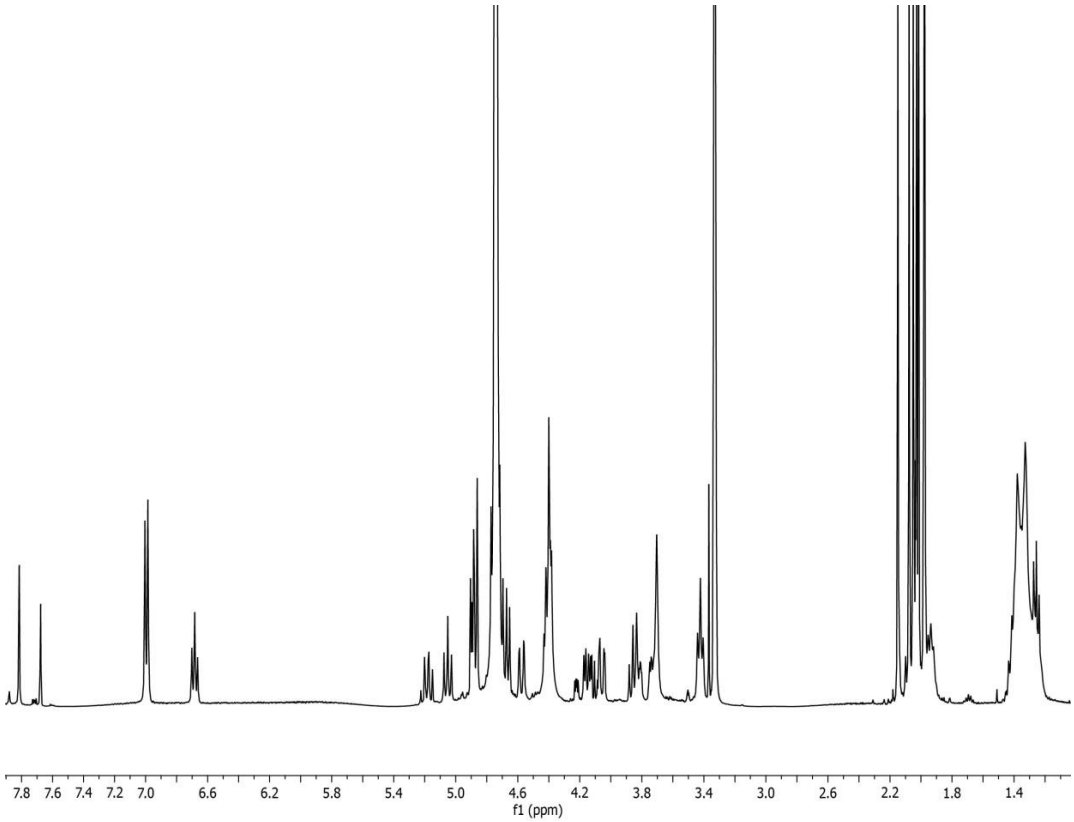


Fig. 8: ^{1}H-NMR spectrum (400 MHz, $CDCl_3$) of peracetylated glycocalixarene **15**

The presence of the triazole unit was confirmed by the singlet, corresponding to 4 hydrogen atoms of the CH of the rings, at 7.68 ppm (Fig. **8**). Also in the ^{13}C-NMR spectrum the presence of the peaks at 144 ppm and 124.5 ppm, for the quaternary and tertiary carbons of the ring respectively, gave a further confirmation of the successful reaction.

The subsequent step was the deprotection of the peracetylated glucosyl and cellobiosyl calix[4]arenes **14** and **15** to give final products **1** and **2** respectively. The Zemplèn method,[76] basic conditions at room temperature, by addition of sodium methoxide to methanol, was used for both the compounds and the reactions were monitored via ESI-MS and NMR, until the total disappearance of acetyl groups. For a full characterization of **2** also 2D-NMR experiments were exploited. The scarce solubility in all solvents the ^{1}H NMR spectra were registered in DMSO-d6 (figure **9**). The complete absence of the acetyl signals between 2.15 and 1.90 ppm showed that the deacetylation was successful. Also with ^{13}C-NMR spectrum the confirmation of the complete deprotection was achieved, since 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) were not present.

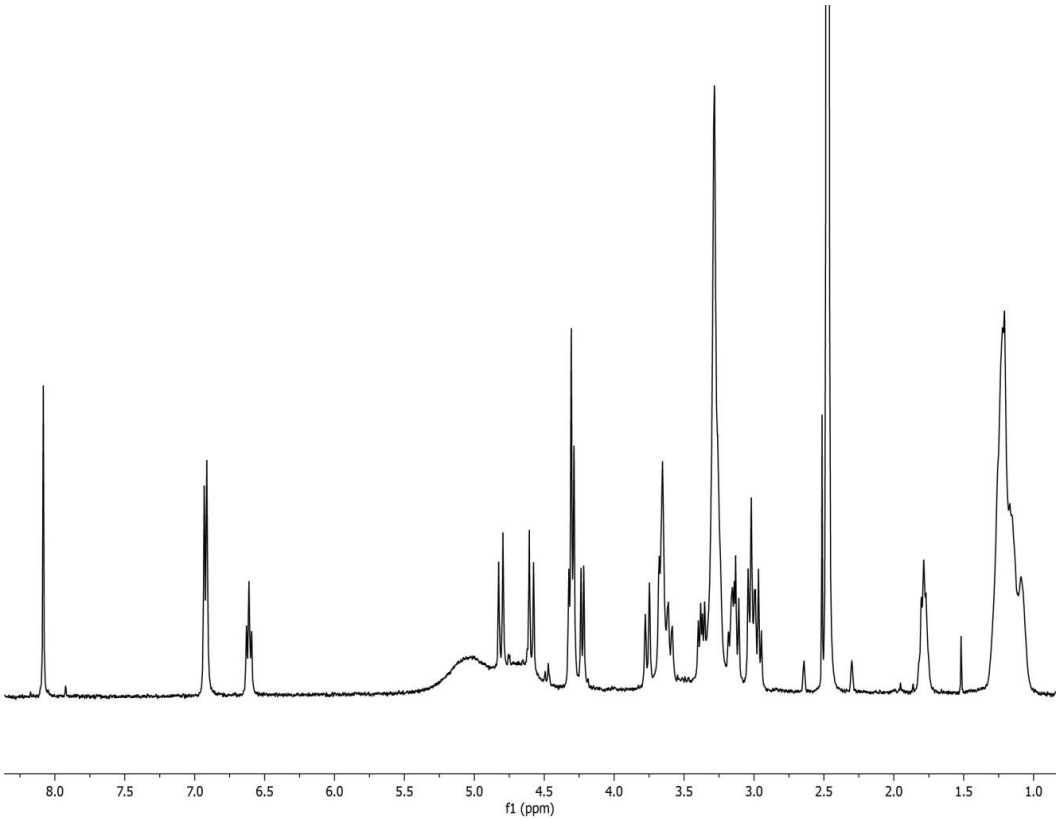
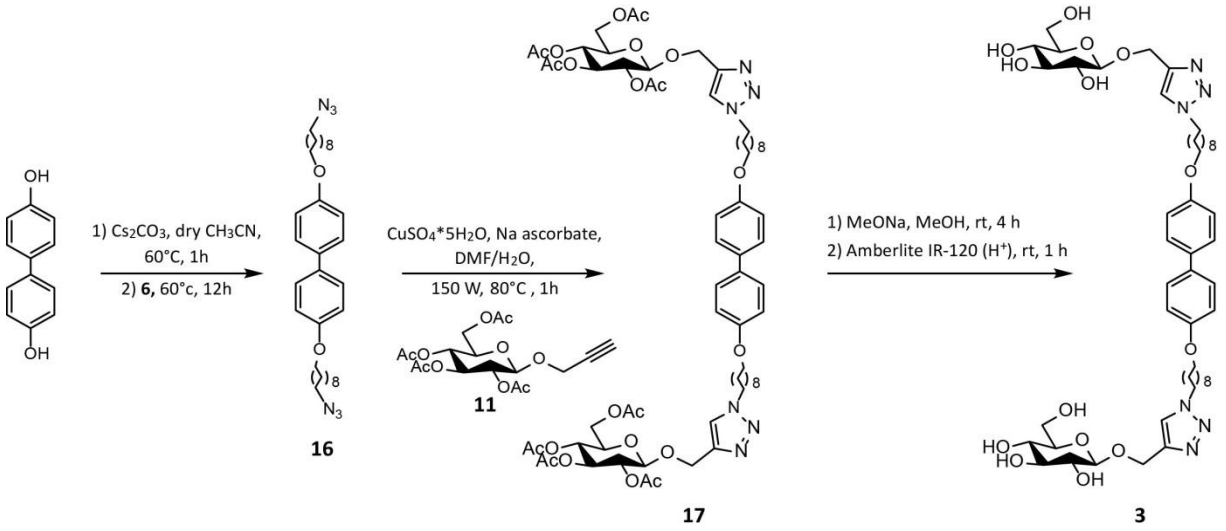


Fig. 9: ^{1}H-NMR spectrum (400m MHz, DMSO-d_6) of target compound **2**

2.2.4 Synthesis of the glucosylated monomer 3

To evaluate if the macrocyclic structure of the calixarene plays an important role in defining the properties of the mixed liposomes (higher stability and rigidity etc..), the linear bolaamphiphile analogue **3** was synthesized as reported in **Scheme 6**.



Scheme 6: synthesis of glycosyl monomer **3**

This monomer is based on a 4,4'-biphenolic core to reproduce as much as possible the reciprocal disposition and distance that sugar units have in the calixarene derivatives **1** and **2**. It was prepared then with the same aliphatic chains and with two units of glucose as **1** to be comparable.

The first reaction was the alkylation of the phenolic hydroxyl groups with chain **6**, using Cs_2CO_3 as a base for the activation of OH. The reaction allowed the formation of the desired derivative **16**, even if required long reaction times and gave low yields (25%). The reason could be correlated to the higher mobility and flexibility of **6**, that can wrap on itself. This could afford the camouflage of terminal CH_2Br groups, reducing the possibilities of attack by the phenate ion. The presence of the di-functionalized product **16** was confirmed by ESI-MS analysis and NMR spectroscopy.

The further step was the conjugation of two units of the β-glucoside **11**, via CuAAC, to **16**. The product was easily obtained and a purification step was required only for the removal of sugar in excess.

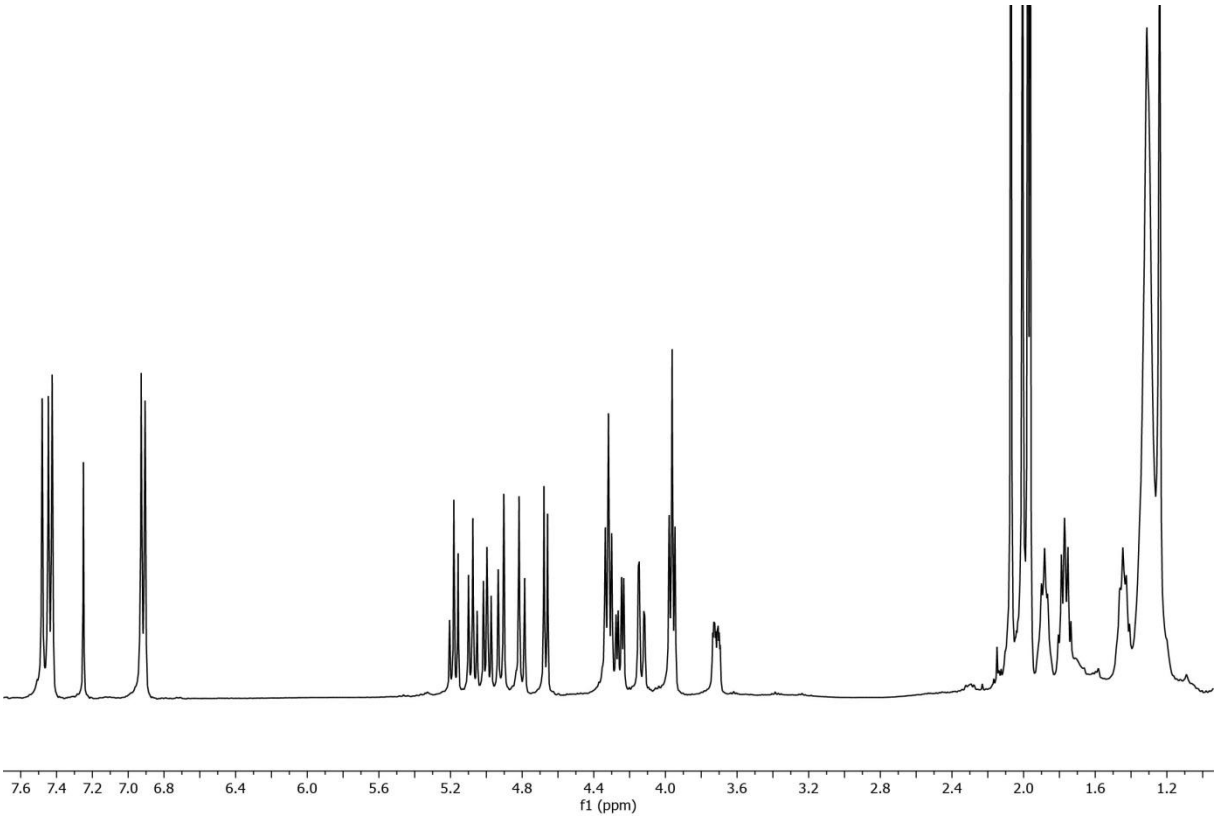


Fig. 10: ^{1}H-NMR spectrum (400 MHz, $CDCl_3$) of compound **17**.

As it is possible to observe in the ^{1}H-NMR spectrum reported above, the two douplets at 4.80 and 4.94 ppm, related to the diastereotopic protons of the methylene between triazolic units and the oxygens on anomeric position, integrate for 2 protons respectively. This integral value, as all those belonging to the glucoside, demonstrates the double functionalization, as well as the two aromatic peaks, related both to 4 protons, that indicate the equivalence of the two aromatic units. At ≈ 7.5 ppm the signal of the triazole hydrogen, a singlet integrating for 2 H, further confirms the di-functionalization of the linear scaffold. Furthermore the stereochemistry of the anomeric carbon was maintained (J coupling H_1-H_2 with a value of 7.2 Hz).

The removal of acetyl groups was once again achieved with Zemplén[76] method, MeONa in methanol, with further neutralization with Amberlite IR-120 (H^+). The final compound gave rise to the spectrum reported in figure **11**, where it is possible to observe the absence of acetyl groups in the region around 2 ppm. The signal of H_1 at ≈ 4.25 ppm had a J coupling higher than 6 Hz (8.0 Hz), indicating the maintenance of the β-stereochemistry on the anomeric carbon, even with basic treatment.

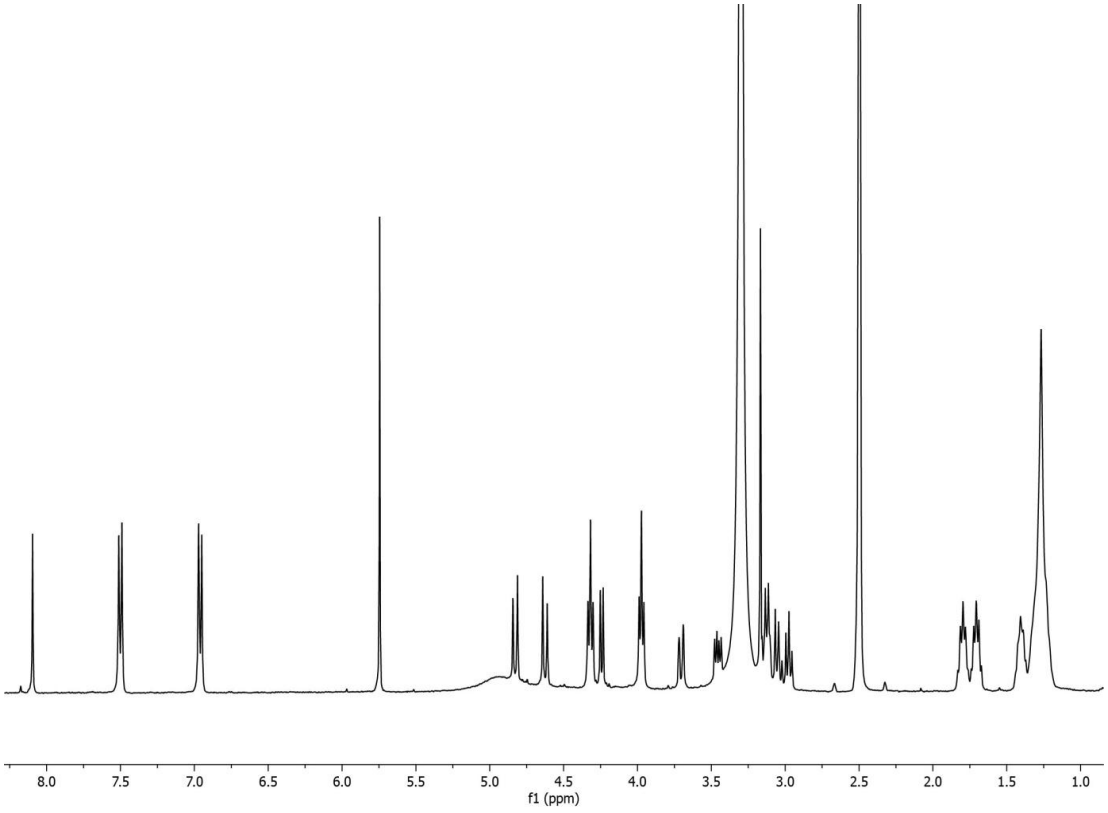


Fig. 11: ^{1}H-NMR spectrum of the deprotected di-glucose derivative **3** (400 MHz, DMSO-d_6).

2.2.5 Preparation of mixed liposomes

The preparation of the mixed liposomes and the agglutination experiments were carried out at University "La Sapienza" and CNR in Rome, by the group of Prof. Giovanna Mancini.

Lipid films containing either 95 or 90% of different diacylphosphatidylcholines and 5 or 10% of **1/2** were hydrated with PBS, freeze-thawed and extruded through a polycarbonate membrane to obtain aqueous suspensions of 100 nm liposomes. Pure diacylphosphatidylcholines were used instead of mixtures to better rationalize the molecular parameters controlling lipid-bolaamphiphile interactions. Lipid films composed of DOPC (1,2-Dioleoyl-sn-glycero-3-phosphocholine) and containing 5 and 10% of **1** or **2**. For **1** both yielded relatively stable liposome suspensions. However for the cellobiose derivative **2** with 10% the results were not complete clear on the nature of the formed aggregate. Furthermore seemed that with this formulation the aggregates contained less calixarene after the extrusion procedure.[77] Afterwards a lower concentration of **2** (5%) was required to form stable mixed liposomes. For **1** the creation of functionalized liposomes was described in the article by Aleandri et al..[55] The mixed DOPC/**2** liposomes formed are schematically represented in figure **12**. The bolaamphiphile **1** gave the positive effects on the liposome stability as previously observed,[55] suggesting that the calixarene scaffold strongly perturbs the bilayers formed by saturated lipids. The hypothesis was that in **1** the higher conformational freedom of alkyl tails respect to the more rigid aromatic scaffolds creates in the bilayer destabilizing empty spaces that are easily filled only by unsaturated chains. In order to determine or confirm these properties for DOPC/ **2** liposomes further characterization studies will be performed in the group of Prof. Giovanna Mancini.

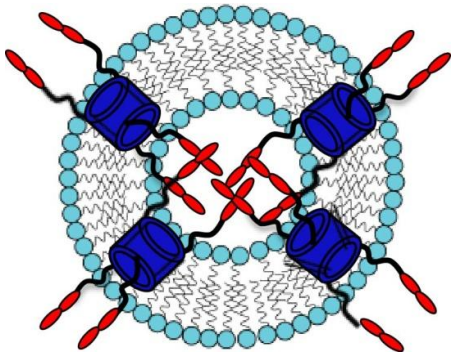


Fig. 12: Proposed schematic structure of a mixed DOPC/**2** liposome

DOPC was also mixed with the monomer **3** to study their assembly properties, but it was not possible to obtain stable mixed liposomes as on the contrary obtained with **1** and **2**. This result, of course, did not allow to verify possible differences between mixed liposomes with the glycocalixarenes and mixed liposomes with the non macrocyclic derivative. However, it already evidenced a marked difference between the two types of molecular structures with the clear advantage offered by the macrocycle over the non macrocyclic one.

2.2.5.1 Agglutination tests

It was already verified that mixed liposomes DOPC/**1** with calixarene at 10% were able to interact in a multivalent fashion with Con A.[44] Agglutination experiments were then performed with DOPC/**2** and DOPC/**1** mixed liposomes both at 5% of calixarene. The agglutination experiments were carried out by turbidimetric measurements in the presence of Con A. The turbidimetric method of analysis of the liposome suspensions in the presence of Con A is based on measuring time-dependent sample turbidity, determined as optical density (OD) at 525 nm, where the absorption of the sample is negligible. Therefore the light attenuation by the particle suspension is only due to scattering and OD variations are directly related to agglutination phenomena. 30 OD changes observed over six hours after the addition of Con A to the liposome solution (Fig. **13a**), and complete aggregate precipitation observed in 24 hours indicated agglutination according to the recognition mechanism between sugar and protein (Fig. **13b**). In the absence of Con A, neither OD changes nor precipitation was observed. The lower absorbance observed along all the monitored period for DOPC/**1** mixed liposomes in comparison with the DOPC/**2** ones, suggests a minor extent of agglutination. This result can be explained with a better exposition of the terminal glucose units of the cellobiose containing derivative **2** to the external space. In this way the new liposomial formulation consisting in the presence of **2** would seem able to interact more efficiently with the tetravalent Con A, in agreement with the initial idea that prompted the preparation of this second glycocalix[4]arene. These measurements need to be repeated at least three times to be sure that is the final result, in order to have comparable and reproducible results.

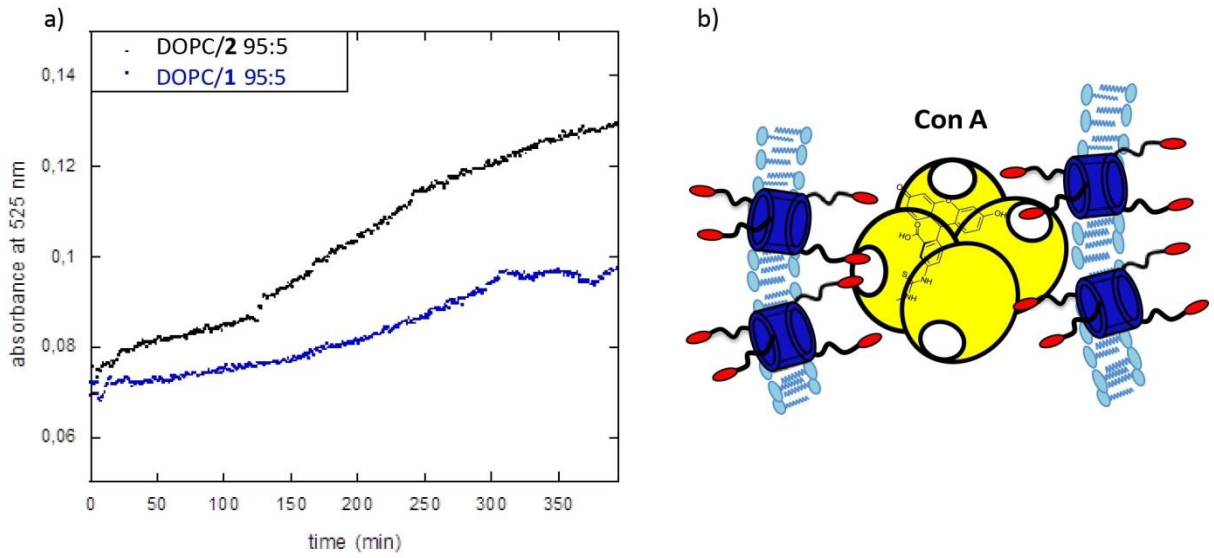


Fig. 13: a) Time-dependent changes of the specific turbidity of a 3 mL sample of 0.58 mM DOPC/**1** (blue dots) and DOPC/**2** (black dots) liposomes upon addition of Con A (final concentration = 0.24 mg/mL); b) schematic representation of the glycocalixarene-based liposome and Con A interactions.

These results encouraged additional experiments on these systems, in particular with some in vitro studies on cell uptake reported in the next paragraphs.

2.2.5.2 MTT assay: toxicity

Preliminarily to the internalization experiments, it was necessary to determine the toxicity of the mixed liposomes towards cells, in order to establish the feasibility of the planned studies and the right conditions. The toxicity tests and laser scanning confocal microscopy experiments were done at "Istituto Superiore di Sanità" in Rome, by the group of Dr. Agnese Molinari.

Previously the bolaamphiphiles-based liposomes were checked to determine their toxicity and to perform some preliminary cell uptake experiments. Since DOPC/**1** liposomes were already tested and no evidences of toxicological effects were found, the only mixed liposomes studied at this stage were DOPC/**2**.

The MTT assay was used, based on 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide, a dye that is reduced by enzymes of living cells to formazan characterized by a purple colour.

Three human breast cancer cell lines, MDA-MB-231, MCF7 WT and SKBR3, selected for the following uptake experiments, were treated with three different concentrations of liposomes, 2, 10 and 20 μM, to determine the toxicity. The time of the treatments were 4, 24 and 48 h. It was observed no time- and formulation-dependent toxicity for all the three liposomial concentrations towards all the considered lines, therefore in the further experiments even the highest one was used.

2.2.5.3 Uptake experiments

Unfortunately, for a series of reasons, the experiments with cells evidenced more problems than expected, then the results we can report here are very few and very preliminary.

The qualitative analysis of the uptake and intracellular distribution of the DOPC/**1** and DOPC/**2** mixed liposomes was carried out by laser scanning confocal microscopy (LSCM), in all the three cell lines.

In order to improve the detection of fluorescent signal by the optical apparatus, cultures were treated with the formulations tested with the MTT assay experiment, using more concentrated liposome suspensions: 10 μM and 20 μM. The analysis was performed after 30 min, 1 h, 4 h and 24 h of incubation.

The used mixed liposomes were detectable by fluorescence microscopy and confocal microscopy thanks to the presence in the bilayer of phopholipids bearing covalently linked a unit of the NBD fluorophore. This molecule is able to emit light in the green, that allows to follow easily the internalization into the cells. The three selected human breast cancer cell lines, MDA-MB-231, MCF7 and SKBR3, overexpress the glucose receptor GLUT1 and then allow to evidence if an improvement in the internalization process occurs in the presence of the glucose units linked to the calixarene included in the liposome bilayers. The increase in the fluorescence intensity (ratio between the medium channel of fluorescence for the treated samples and the control) is a measurement of the internalization of the liposomes: strong green fluorescence indicate efficient uptake of the studied systems and vice versa.

These experiments used DOPC liposomes as control (Figure **14**, left). Firstly DOPC/glicocalixarene **1** were tested but they were not able to significantly penetrate into the cells. In fact no green fluorescence more intense respect to the cells treated with DOPC liposomes was observed. This behaviour was rather unexpected considering the

agglutination experiments. In fact DOPC/**1** liposomes had shown to be able of interacting with Con A, but apparently, when placed in contact with cells, no interactions with the glucose receptor overexpressed in these lines took place. Probably the glucose units were not well and sufficiently exposed from the external liposomial membrane to be recognized and complexed by GLUT1.

On the contrary, mixed liposomes DOPC/**2** resulted effective in the internalization process. Here it is reported in particular the result found in the experiment with MDA-MB-231 cells. The green fluorescence observed in cells upon treatment with DOPC/**2** mixed liposomes (Figure **14**, right) was significantly more intense than that registered in the experiment with simple DOPC liposomes. This was due to a better uptake of these functionalized liposomes. Fluorescence was especially confined in cytoplasmic vesicles and in the Golgi apparatus

On the basis of this evidence we could say that the internalization of DOPC/**2** mixed liposomes appears mainly based on a selective recognition event between the glucose receptor GLUT1 and vesicles functionalized with cellobiose. In fact the absence of glycocalixarene in the liposomial structure determines a significantly less efficient uptake in the tumour cells, since no carbohydrate units are exposed to the external membrane for the binding with GLUT1 receptor.

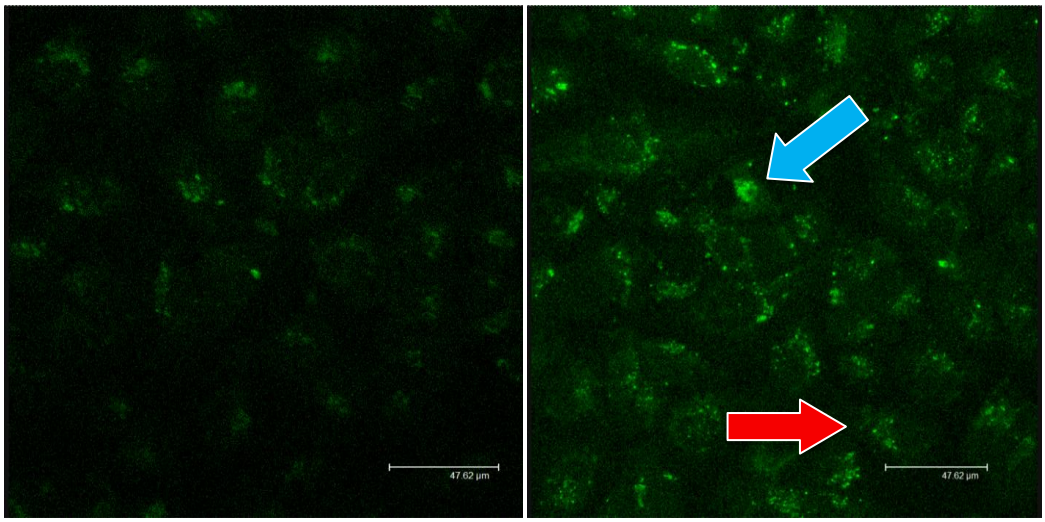


Fig. 14: Confocal microscopy images of cells treated with a) DOPC liposomes and b) mixed DOPC/**2** liposomes. The green fluorescence is due to the presence of a NBD based tag in liposomes. Red and blue arrows indicate fluorescence in cytoplasmatic vescicles and Golgi apparatus respectively.

These results are indeed interesting and encourage further study with these DOPC/**2** liposomes, in particular in relation with the possibility to verify a real capability of delivering medically relevant species, drugs in particular, into the cells and to make it in a selective way.

2.3 Conclusions

As previously reported,[55] a glycocalixarene in 1,3-alternate geometry (bola-shaped), **1,** was able to create more stable mixed liposome. This particular space exposition allowed the hydrophobic part to interacted with the lipidic bilayer and the sugar units (glucose) were involved in the interactions with a plant lectin, Con A, taking advantage of multivalent interactions.

A non macrocyclic glucose-based analogue **3**, with the same linker and sugar unit, was prepared to evaluate the role of calixarene scaffold in the formation of more rigid liposomes. However this compound was not able to create stable liposomial formulations as **1**, confirming the positive effect of calixarene structure in the self assembly of DOPC mixed liposomes.

In order to improve this recognition process, a new bolaamphiphiles calix[4]arene **2** was synthesized, analogue of the previous **1,** but bearing four longer saccharidic units: cellobiose. The better exposition of the polar heads to the external liposome membrane may allow the formation of stronger interactions with lectins.

Firstly assembly properties of this new compound **2** were determined and even in this case mixed liposomes were formed. Agglutination tests were performed and **2**, compared to **1,** had higher OD changes, indication of the occurrence of stronger protein-ligand interactions. This results is a confirmation of the increased binding with lectin Con A in DOPC/**2** liposomes respect to glucose-based liposomes, thanks to the presence of more exposed cellobiose units.

Since glycocalixarene **2** gave these interesting results, cell uptake studies were done to better understand if the carbohydrate-lectin interactions are at the base of liposomes internalization in the cells.

Data obtained using laser scanning confocal microscopy on human breast cancer cells (three different lines) were so obtained. DOPC/ **2** liposomes, labelled with a fluorophore emitting

green light, were studied on these cells able to overexpress the glucose receptor.

First outcomes demonstrated the improvement in the cell uptake using mixed liposomes respect to DOPC liposomes. In fact, thanks to fluorescence detection, a low internalization with these was observed while with DOPC/**2** mixed liposomes the uptake resulted much more significant. This is a confirmation of that the recognition process is based on carbohydrate-protein interactions between the GLUT1 glucose receptor and the glycocalixarene-based systems.

2.4 Experimental part

General information. All moisture sensitive reactions were carried out under nitrogen if specifically required) 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 aluminium) and then, according to the functional groups present on the molecules, revealed with UV light or using staining reagents: H_2SO_4 (5% in EtOH), ninhydrin (5% in EtOH), basic solution of $KMnO_4$ (0.75% in H_2O), Pancaldi solution (molybdatophosphorus acid and Ce(IV)sulphate in 4% sulphuric acid). Reverse phase TLC were performed by using silica gel 60 RP-18 F-254 on aluminium sheets. Merck silica gel 60 (70-230 mesh) was used for flash chromatography. ^{1}H NMR and ^{13}C NMR spectra were recorded on Bruker AV300 and Bruker AV400 spectrometers (observation of ^{1}H 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, δ_{TMS} = 0), as internal standard. All ^{13}C NMR spectra were performed with proton decoupling. For ^{1}H NMR spectra recorded in D_2O at values higher that 25°C the correction of chemical shifts was performed using the expression $\delta = 5.060 - 0.0122\times T(°C) + (2.11\times10^{-5})\times T^2(°C)$ to determine the resonance frequency of water protons.[78] Electrospray ionization (ESI) mass analyses were performed with a Waters spectrometer in both positive and negative mode with MeOH/CH_3CN as solvents. Melting points were determined on an Electrothermal apparatus

in closed capillaries. Microwave reactions were performed using CEM Discovery System reactor.

Nomenclature of calix[n]arene compounds.

In this thesis the simplified nomenclature proposed by Gutsche is used to name the calix[n]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.

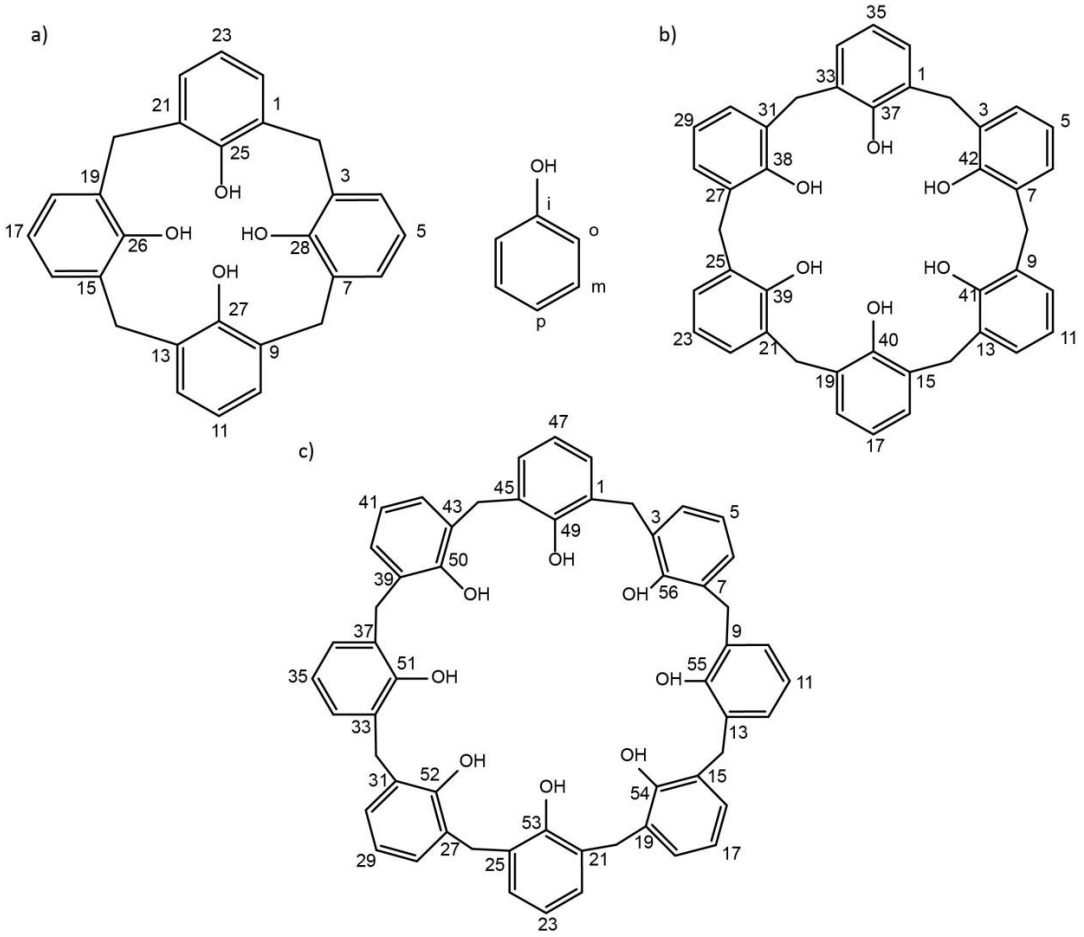


Fig.14: nomenclature of calix[n]arenes given by Gutsche for a) calix[4]arenes, b) calix[6]arenes and c) calix[8]arenes

Synthesis

1-bromo-10-decanol (4)

To a solution of 1,10 decandiol (6 g, 34.43 mmol) in dry toluene (85 ml), HBr 48% (4.6 ml, 40.28 mmol). The reaction was left to react under reflux (115°C), stirring for 12 h.

After monitoring with TLC (eluent: Hexane/Et_2O), the two phases were separated and the organic layer was diluted with Et_2O, then washed with NaOH 1M (50 ml), brine (3x50 ml), till neutral pH. The organic phase was dried over anhydrous Na_2SO_4 and then evaporated under reduced pressure. The removal of the di-bromide byproduct was achieved by flash column chromatography (eluent: Hexane/EtOAc 8:2) to give mono-bromide derivative as colorless oil (86% yield).

^{1}H-NMR (300 MHz, $CDCl_3$): δ (ppm) 3.55 (t, 2H, *J* = 6 Hz, CH_2Br); 3.35 (t, 2H, *J* = 6 Hz, HOCH_2); 1.84-1.77 (m, 4H, $HOCH_2CH_2$); 1.42-1.24 (m, 12H, $HOCH_2CH_2(CH_2)_9$).

The spectroscopic data are the same of those reported in literature.[79]

10-azide-1-decanol (5)

The derivative **4** (5.20 g, 21.93 mmol) was dissolved in dry DMF (25 ml), under N_2 atmosphere and then NaN_3 was added. After stirring for 12 h, at rt (TLC eluent: hexane/EtOAc 7:3), the reaction was quenched with water (30 ml) and extracted with cyclohexane (3x30 ml). After dying over anhydrous Na_2SO_4 the solvent was concentrated under vacuum to afford the compound **5** as a colorless oil, in 73% yield.

^{1}H-NMR (400 MHz, $CDCl_3$): δ (ppm) 3.55 (t, 2H, *J* = 6.3 Hz, HOCH_2); 3.19 (t, 2H,*J* = 6.9 Hz, CH_2N_3); 1.53-1.50 (m, 4H, $HOCH_2CH_2$); 1.24 (s, 12H, $HOCH_2CH_2(CH_2)_6$).

The derivative shows the same spectroscopic data present in literature.[79]

10-Azido-1-bromodecane (6)

10-Azido-1-decanol (3.60 g, 18.07 mmol) and trifluoroacetic anhydride (2.6 ml, 19.88 mmol) were dissolved in dry THF (15 mL) and stirred at room temperature for 30 min. After that the solvent and the excess of anhydride were removed using a rotary evaporator. A 1 M solution of LiBr in THF-DMF (20 mL, 1/1, v/v) was added to the oily residue and the mixture was stirred at 75 °C overnight. A further portion of solid LiBr (880 mg, 10 mmol) was added and the heating continued for another 24 h. The reaction was monitored by gas-chromatography and, when finished, it was quenched by addition of water (40 mL) and extracted with cyclohexane (3 × 35 mL). The combined organic phases were dried over anhydrous Na2SO4 and, after filtration, the solvent was removed under vacuum to afford the product **6** in 85%as a yellow oil without further purification.

^{1}H-NMR (300 MHz; $CDCl_3$): δ (ppm) 3.42 (t, 2H, *J* = 6.9 Hz, CH_2Br); 3.23 (t, 2H, *J* = 6.9, CH_2N_3); 1.91-1.79 (m, 2H, CH_2CH_2Br); 1.73-1.53 (m, 2H, $CH_2CH_2N_3$); 1.51-1.27 (m, 12H, $N_3(CH_2)_2(CH_2)_6(CH_2)_2Br$).

The spectroscopical data are in agreement with those reported in literature. [55]

25,26,27,28-Tetrahydroxycalix[4]arene (8)

In a two neck round bottom flask, containing 300 ml of toluene, calixarene **7** (30 g, 46.36 mmol), $AlCl_3$ (33.6 g, 254.47 mmol) and phenol (5.2 g, 55.47 mmol) were added. The mixture was stirred with magnetic stirrer at room temperature for 1 day and controlled via TLC (eluent: hexane/EtOAc 8:2). The reaction was quenched adding a ice/HCl 1 M solution (400 ml) and then stirred for 1h. The organic phase was then washed with water till neutral pH and dried over Na_2SO_4, evaporating after the solvent under vacuum. The crude was crystallized with EtOAc/MeOH 95:5 to afford compound **8** as a yellowish solid.

^{1}H-NMR (300 MHz, $CDCl_3$):δ (ppm) 7.05 (d, 8H, *J* = 7.5 Hz, H para); 6.73 (t, 4H, *J* = 7.5 Hz, H meta);4.14 (br s, 4H, $ArCH_2Ar$ eq.); 3.48 (br s, 4H, $ArCH_2Ar$ ax.).

The product shows the same chemical-physical features reported in literature.[80]

25,26,27,28-Tetrakis-(10-azido-decyloxy)calix[4]arene, 1,3 alternate (9)

Tetra-hydroxycalix[4]arene **8** (0.50 mg, 1.72 mmol) and Cs_2CO_3 (4.59 g, 11.75 mmol) were suspended in dry DMF (15 mL) and the mixture stirred for 1 h at 50 °C. 10-Azido-1-bromodecane **6** (3.08 g, 11.75 mmol) was added to the mixture and the reaction mixture stirred at the same temperature for 3-4 days. The reaction mixture was then quenched by addition of 1 M HCl till neutral pH. The mixture was extracted with EtOAc (3 × 30 ml), the organic phases combined, dried over anhydrous Na_2SO_4, filtered and the solvents evaporated under vacuum. The residue was initially by flash column chromatography (eluent: hexane-CH_2Cl_2 65:35) to remove the partial cone derivative and then by crystallization in MeOH. The desired product was obtained as a white solid in 58% yield.

^{1}H-NMR (400 MHz, $CDCl_3$):δ (ppm) 6.98 (d, 8H, *J* = 7.4 Hz, ArH meta); 6.64 (t, 4H, *J* = 7.4 Hz, ArH para); 3.59 (s, 8H, $ArCH_2Ar$); 3.55 (t, 8H, *J* = 7.6 Hz, OCH_2); 3.26 (t, 8H, *J* = 7.2 Hz, CH_2N_3); 1.60 (m, 16H, OCH_2CH_2, $CH_2CH_2N_3$), 1.38 (m, 48H, $OCH_2CH_2(CH_2)_6CH_2CH_2N_3$).

The pure alkylated calixarene presents the same chemical-physical features presented in literature.[58]

1-O-Propynyl-2,3,4,6-tetra-O-acetyl-β-D-glucopyranoside (11)

To a solution of peracetylated glucose **10** (1 g, 2.56 mmol) in dry DCM (15 ml), under N_2 atmosphere, propargyl alcohol (0.19 ml, 3.33 mmol) and $BF_3 \times Et_2O$ (0.52 ml, 4.01 mmol) were added at 0°C. The temperature was done increase up to rt and stirred for 2h, checking via TLC (eluent: hexane/EtOAc 1:1). It was quenched adding K_2CO_3 (0.35 mg) and stirred at room temperature for 30 min. The salt was filtered off, then mixture was washed with water (2x60 ml. The aqueous layers were extracted with DCM (3x50 ml) and the organic combined phases were dried over Na_2SO_4. The solvent was removed under reduced pressure to afford compound **11** as a white solid in 95%.

^{1}H-NMR (300 MHz, $CDCl_3$): δ(ppm) 5.10 (t, 1H, *J* = 9 Hz, H_4); 4.95 (t, 1H, *J* = 9 Hz, H_3); 4.85 (t, 1H, *J* = 9 Hz, H_2); 4.66 (d, 1H, $J_{1\text{-}2}$ = 9 Hz, H_1); 4.23 (d, 2H, *J* = 3 Hz, OC*H*$_2$CCH); 4.14 (dd, 1H, *J* = 12.6 Hz, H_{6a}); 4.00 (d, 1H, *J* = 12.3 Hz, H_{6b}); 3.63 (m, 1H, H_5); 2.43 (t, *J* = 3 Hz, H, C≡CH); 1.94 (s, 3H, CH_3CO); 1.91 (s, 3H, CH_3CO); 1.88 (s, 3H, CH_3CO); 1.86 (s, 3H, CH_3CO).

The derivative shows the same spectroscopic data present in literature.[81]

β-D-Glucopyranoside, 2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranosyl)-2,3,6-triacetate (12)

To a solution of acetic anhydride (45.4 ml) and CH_3COONa (9.59 g, 116.86 mmol) heated at 130°C, cellobiose (4.0 g, 11.69mmol) was added slowly and the reaction was allowed to react for 1 h under reflux. The reaction was monitored with TLC (eluent: hexane/EtOAc 1:1), quenched adding a ice/water bath (230 ml) and left under mechanical stirring at rt for 24 h. The solid was then filtered on Buchner funnel and then triturated with EtOH. The product was obtained without further purification as a white solid in 80% yield.

^{1}H-NMR (400 MHz, $CDCl_3$): δ (ppm) 6.24 (d, $J_{1\text{-}2}$ = 3.6 Hz, 1H, H1β); 5.65 (d, 1H, $J_{1\text{-}2}$ = 8.4 Hz, H1α); 5.02 (t, 1H, *J* = 9.4 Hz, H_3); 5.12 (dd, 1H, *J* = 7.9, 10.5 Hz, $H_2{}^1$); 5.02 (dd, 1H, *J* = 8.3 and 9.5 Hz, H_2); 4.92 (t, 1H, *J* = 10.5, $H_3{}^1$); 4.51 (m, 1H, $H_4{}^1$); 4.46 (d, 1H, *J* = 7.9 Hz, $H_1{}^1$); 4.36 (dd, 1H, $J_{6b\text{-}5}$ = 4.4 Hz, $J_{6b\text{-}6a}$ = 12 Hz, H_{6b}); 4.11 (dd, 1H, $J_{6a\text{-}5}$ = 4.4 Hz, $J_{6a\text{-}6b}$ = 12 Hz, H_{6a}); 4.09-3.98 (m, 2H, H_{6b1}, $H6_{a1}$); 3.85 (m, 1H, $H_5{}^1$); 3.80 (m, 1H, H_4); 3.74 (m, 1H, H_5); 2.17, 2.12, 2.08, 2.02, 1.98, 1.97 (s, 24H, 8xCH_3).

The spectroscopic data are the same of those reported in literature.[82]

β-D-Glucopyranoside, 2-propyn-1-yl 4-*O*- (2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranosyl)-, 2,3,6-triacetate (13)

Peracetylated compound **12** (6.38 g, 9.41 mmol) was added to a solution of dry DCM (20 ml) under N_2 atmosphere. The temperature was cooled down to 0°C and then propargyl alcohol (0.71 ml, 12.22 mmol) and $BF_3 \times Et_2O$ (1.91 ml, 15.05 mmol) were added. The reaction was stirred for 2 h and monitored by TLC (eluent: hexane/EtOAc 1:1). It was quenched with K_2CO_3 (2g, 138.2 mol) and stirred at room temperature for 30 min. The salt was filtered off, then mixture was washed with water (2x60 ml. The aqueous layers were extracted with DCM (3x50 ml) and the organic combined phases were dried over Na_2SO_4. The solvent was removed under reduced pressure to afford derivative **13** as a white foam in 97%.

^{1}H-NMR (300 MHz, $CDCl_3$): δ (ppm) 5.18 (t, 1H, *J* = 9.3 Hz, $H_3{}^1$); 5.03-4.95 (m, 2H, H_2, $H_2{}^1$); 4.90 (d, 1H, *J* = 7.9 Hz, $H_1{}^1$); 4.73 (d, *J* = 7.8 Hz, 1H, H_1); 4.55 (m, 1H, H_3); 4.49-4.38 (m, 2H, $H_{6b}{}^1$, H_{6b}); 4.33 (d, 2H, *J* = 2.4 Hz, O*CH*$_2$C≡CH); 4.14 (m, 1H, H_4); 4.14-3.98 (m, 2H, $H_{6a}{}^1$, H_{6a}); 3.79 (m, 1H, H_4); 3.68-3.61 (m, 2H, H_5, $H_5{}^1$); 2.45 (t, 1H, *J* = 2.2 Hz, C≡CH); 2.17-1.97 (m, 21H, 7 CH_3). The product shows the same chemical-physical properties found in literature.[83]

General procedure for the "click" reaction to form the glycocalix[4]arenes 14 and 15

To a solution of calixarene **9** (1 eq.), 2-propynyl-peracetylated derivatives **14-15** (1.5 eq for each azide group), $CuSO_4{*}5H_2O$ (0.3 eq) and Na-ascorbate (0.6 eq.) in DMF-H_2O (5 ml, 4/1) was stirred under microwave irradiation (150 W, 80 °C) for 1h-1h 30 min. The reaction was then quenched by addition of water (20 ml) and extracted with EtOAc (3×20 ml). The combined organic layers were dried over anhydrous Na_2SO_4, filtered and evaporated under vacuum.

25,26,27,28-Tetrakis-{10-[4-(methylen-2,3,4,6-tetra-O-acetyl-β-D-glucoside)triazolyl]decyloxy}calix[4]arene, 1,3-alternate (14)

The pure product was obtained via flash column chromatography (eluent: hexane/EtOAc 7:3 to EtOAc 100%) as a white solid (53% yield).

^{1}H-NMR (400 MHz, $CDCl_3$): δ(ppm) 7.86 (s, 4H, H triazole); 6.91 (d, 8H, *J* = 7.6 Hz, ArH meta); 6.58 (t, 4H, *J* = 7.6 Hz, ArH para); 5.14 (t, 4H, *J* = 9.6 Hz, H_3); 5.03 (t, 4H, *J* = 10 Hz, H_4); 4.95 (dd, 4H, *J* = 8.4 Hz,H_2); 4.87 (d, 4H, *J* = 12.4 Hz, OCH*H*N); 4.75 (d, 4H, *J* = 12.4 Hz, OC*H*HN); 4.63 (d, 4H, *J* = 7.6 Hz, H_1); 4.29 (t, 8H, *J* = 7.2 Hz, CH_2N); 4.22 (dd, 4H, *J* = 12.4 and 4.4 Hz,

H_{6a}); 4.09 (d, 4H, *J* = 12.4 Hz, H_{6b}); 3.70-3.66 (m, 4H, H_5); 3.54 (br s, 8H,$ArCH_2Ar$), 3.49 (br s, 8H, *J* = 12.4 Hz, $ArOCH_2$); 2.02, 1.96, 1.93, 1.88 (4×s, 4 ×12H, CH_3CO); 1.86 (br s, 8H, CH_2CH_2N); 1.35-1.15 (56H, m, CH_2).

The spectroscopic data are in agreement with those observed in literature.[55]

25,26,27,28-Tetrakis-{10-[4-methylen-peracetylated cellobiose triazolyl]decyloxy}calix[4]arene, 1,3-alternate (15)

The pure product was obtained after purification of flash column (eluent: hexane/EtOAc 1:9) as a yellowish solid (92%).

^{1}H-NMR (400 MHz, MeOD-$CDCl_3$ 3:1): δ (ppm) 7.67 (s, 4H, H triazole); 6.99 (d, 8H, *J* = 7.6 Hz, ArH meta), 6.70 (t, *J* = 7.6 Hz, 4H, ArH para); 5.22-5.15 (m, 8H, H_3^1, H_3); 5.05 (t, 4H, *J* = 10 Hz, H_4^1); 4.90-4.86 (m, 8H, H_2, H_2^1); 4.77 (m, 8H, OCH_2N); 4.71 (d, 4H, *J* = 7.9 Hz, H_1^1), 4.66 (d, 4H, *J* = 8 Hz, H_1); 4.57 (d, *J* = 10.4 Hz, 4H, H_{6a}); 4.43-4.38 (m, 12H, H_{6a}^1, CH_2CH_2N); 4.16 (m, 4H, H_{6b}); 4.05 (dd, 4H, $J_{6b^1\text{-}5}$ = 1.6 Hz, $J_{6b^1\text{-}6a^1}$ = 12.4 Hz, H_{6b1}); 3.85 (m, 4H, H_4); 3.81 (m, 4H, H_5^1); 3.74 (m, 4H, H_5); 3.70 (s, 8H, $ArCH_2Ar$); 3.42 (t, 8H, *J* = 7.2 Hz, OCH_2CH_2); 1.93 (m, 8H, CH_2CH_2N); 1.37-1.25 (m, 56H, $OCH_2CH_2(CH_2)_6$). **^{13}C-NMR** (100 MHz, MeOD): δ (ppm) 170.9, 169.5 ($COCH_3$); 155.3 (C ipso); 142.9 (C_q triazole); 133.8 (C orto); 129.5 (C meta); 121.5 (C para); 100.5 (C_1); 99.5 (C_1^1); 76.3 (C_4); 73.1-72.7 (C_3, C_3^1, C_5, C_5^1,); 71.7- 71.5 (C_2, C_2^1, OCH_2CH_2); 67.8 (C_4^1); 61.8-61.4 (C_{6a}, C_{6b}, C_{6a}^1, C_{6b}^1,OCH_2N); 51.4 (CH_2CH_2N); 38.5 ($ArCH_2Ar$), 30.1, 29.5, 29.4, 29.3, 28.9, 27.4, 26.3, 25.8 (8xCH_2 chain);19.9 ($COCH_3$).

ESI-MS: m/z 1305.06 $[M+Na]^{3+}$, 1924.78 $[M+H]^{2+}$, 1946.08 $[M+Na]^{2+}$

M.p.: 118.3-120 °C

General procedure for the deacetylation of glycocalixarenes 1 and 2

The fully protected calixarene was dissolved in MeOH and then solid CH_3ONa (freshly prepared) was added till pH 9-10. The reaction was checked via ESI-MS and NMR; when finished, the reaction was quenched by dipping in the reaction mixture a sintered glass vessel containing Amberlite IR-120 (H^+), under stirring till neutral pH. The solvent was removed under reduced pressure to obtain the pure product.

25,26,27,28-Tetrakis-{10-[4-(methylen-β-D-glucoside)triazolyl]-decyloxy}calix[4]arene, 1,3-alternate (1)

The pure glycocalixarene **1** was obtained as a yellowish solid (yield= 87%).

^{1}H-NMR (400 MHz; DMSO-d6): δ (ppm) 8.11 (s, 4H, C*H* triazole); 6.94 (d, 8H, *J* = 7.2 Hz, ArH meta); 6.65 (t, 4H, *J* = 7.2 Hz, AH para); 5.03 (d, 4H, *J* = 4.8 Hz, OH-2); 4.94 (d, 4H, *J* = 4.82 Hz, OH-3); 4.91 (d, 4H, *J* = 5.2 Hz, OH-4); 4.83 (d, 4H, *J* = 12.0 Hz, NC*H*HO); 4.61 (d, 4H, *J* = 12.0 Hz, NCH*H*O); 4.54 (t, 4H, *J* = 5.2 Hz, OH-6); 4.32 (t, 8H, *J* = 6.9 Hz, CH_2N); 4.25 (d, 4H, *J* = 7.6 Hz, H_1); 3.80-3.68 (m, 12H, $ArCH_2Ar$ and H_{6a}); 3.50-3.38 (m, 4H,H_{6b}); 3.26 (br s, 8H, CH_2CH_2O); 3.18-3.12 (m, 8H, H_3 and H_5); 3.10-3.06 (m, 4H, H_4); 3.05-2.94 (m, 4H, H_2); 1.83 (br s, 8H, CH_2CH_2N); 1.40-1.04 (m, 56H, CH_2 chain).

The product as the same spectroscopic data found in literature. [55]

25,26,27,28-Tetrakis-{10-[4-(methylen-β-D-cellobioside)triazolyl]-decyloxy}calix[4]arene, 1,3-alternate (2)

The pure product was obtained as a white oil (68% yield).

^{1}H-NMR (400 MHz, DMSO-d_6): δ (ppm) 8.08 (s, 4H, H triazole); 6.92 (d, 8H, *J* = 7.2 Hz, Ar*H* meta); 6.61 (t, 4H, *J* = 7.6 Hz, Ar*H* para); 4.90 (br s, 28H, O*H*); 4.84 (d, 4H, *J* = 12.4 Hz, OC*H*HN); 4.59 (d, 4H, *J* = 12.4 Hz, OCH*H*N); 4.47 (d, 4H, *J* = 8 Hz, $H_1{}^1$); 4.30 (t, 8H, *J* = 7.2 Hz, CH_2CH_2N); 4.22 (d, 4H, *J* = 8 Hz, H_1); 3.74-3.59 (m, 20H, H_{6a}, $H_4{}^1$, H_{6b}, $ArCH_2Ar$); 3.39-3.25 (m, 20H, H_4, H_5, $H_5{}^1$, $H_{6b}{}^1$, $H_{6a}{}^1$); 3.28 (m, 8H, OCH_2); 3.18-3.10 (m, 8H, $H_3{}^1$, H_3); 3.13-2.96 (m, 8H, H_2, $H_2{}^1$); 1.78 (m, 8H, CH_2CH_2N); 1.21 (m, 56H, $OCH_2CH_2(CH_2)_6$). **^{13}C-NMR** (100MHz, DMSO-d_6): δ (ppm) 157.2 (C ipso); 144.0 (C_q triazole); 134.0 (C orto); 129.6 (C meta); 124.5 (C_t triazole); 122.2 (C para); 103.6 ($C_1{}^1$); 102.2 (C_1); 81.0, 77.2, 76.93, 75.5, 73.5, 71.0, 71.6, 70.5 (C_3, $C_3{}^1$, C_5, $C_5{}^1$, $C_6{}^1$, C_2, $C_2{}^1$, OCH_2CH_2); 61.4, 60.8, 60.5, 60.1 (C_4, $C_4{}^1$, OCH_2N, C_6) 49.7 (CH_2CH_2N); 39.5 ($ArCH_2Ar$); 30.3-25.7 (CH_2 chain).

ESI-MS: m/z 1357.96 (100% $[M+Na]^{2+}$)

M.p.: dec. > 130°C

O,O'-bis(10-azidodecil)bisphenol (16)

To a solution of 4,4′-biphenol (0.1 g, 0.54 mmol) in dry CH_3CN (8 ml), Cs_2CO_3(0.71 g, 2.68 mmol) was added and the reaction was stirred at 60 °C for 1h. Afterwards 10-Azido-1-bromodecane **6** was dissolved and the monitoring was done via TLC (eluent: hexane/EtOAc

40:1), until the formation of the di-substituted derivative (12 h). It was quenched adding water and then the organic phase (EtOAc) was washed with HCl 1M till neutral pH and with brine (3x20 ml). The solvent was removed under vacuum and the crude was purified under flash column chromatography (eluent: hexane/EtOAc 40:1) and after with crystallization in hexane/EtOAc 8:2 to afford the product as a white solid (25%).

^{1}H-NMR (300 MHz, $CDCl_3$): δ (ppm) 7.45 (d, 4H, *J* = 8.7 Hz, ArH); 6.90 (d, 4H, *J* = 8.7 Hz, ArH); 4.12 (t, 4H, *J* = 6.6 Hz, OCH_2); 3.25 (t, 4H, *J* = 6.9 Hz, CH_2N_3); 1.90-1.72 (m, 4H, OCH_2CH_2); 1.65-1.52 (m, 4H, $CH_2CH_2N_3$); 1.50-1.20 (m, 24H, $OCH_2CH_2(CH_2)_6$). **^{13}C-NMR** (75 MHz, $CDCl_3$): δ (ppm) 158.3 (C ipso); 133.3 (C para); 127.7 (C meta); 114.7 (C orto); 68.1 (OCH_2); 51.5 (N_3CH_2); 29.3-26.1 (CH_2 chain).

ESI-MS: m/z 571.38 (100% $[M + Na]^+$)

M.p.: 89.7-91.8 °C

10-[4-(methylen-2,3,4,6-tetra-O-acetyl-β-D-glucoside)triazolyl]decyloxy-bisphenol (17)

In a MW tube **16** (0.17 g, 0.32 mmol) was added in DMF/H_2O (4 ml/500µl). Then the sugar **11** (0.49 g, 1.26 mmol), $CuSO_4$*$5H_2O$ (24 mg, 0.09 mmol) and Na ascorbate (37 m, 0.19 mmol) were added. The tube was irradiated under MW (150 W) for 1h at 80°C. After monitoring via TLC (eluent: hexane/EtOAc 2:8), the reaction was quenched with water (5 ml) and extracted with EtOAc (3x5 ml). After the removal of the solvent under vacuum, the crude was purified on flash column (eluent: hexane/EtOAc 2:8) to give pure product as a white solid in 89% yield.

^{1}H-NMR (400 MHz, $CDCl_3$): δ (ppm) 7.49 (s, 2H, H triazole); 7.43 (d, 4H, *J* = 8.4 Hz, ArH); 6.91 (d, 4H, *J* = 8.4 Hz, ArH); 5.18 (t, 2H, *J* = 9.6 Hz, H_4); 5.07 (t, 2H, *J* = 9.6 Hz, H_3); 5.0 (t, 2H, *J* = 9.6 Hz, H_2); 4.92 (d, 2H,*J* = 12.8 Hz, OC*H*HN); 4.80 (d, 2H, *J* = 12.8 Hz, OCH*H*N); 4.67 (d, 2H, *J* = 7.2 Hz, H_1); 4.32 (t, 4H, *J* = 6.6 Hz, $ArOCH_2$); 4.25 (dd, 2H, *J* = 12.4, 4.8 Hz, H_{6a}); 4.13 (dd, 2H, *J* = 12.4, 2.4 Hz, H_{6b}); 3.96 (t, 4H, *J* = 6.4 Hz, CH_2CH_2N); 3.78-3.68 (m, 2H, H_5); 2.07, 2.01, 1.98, 1.96 (4s, 4x3, CH_3CO); 1.95-1.82 (m, 4H, OCH_2CH_2); 1.82-1.70 (m, 4H, CH_2CH_2N); 1.50-1.40 (m, 4H, CH_2CH_2O); 1.40-1.15 (m, 20H, $OCH_2CH_2(CH_2)_5$). **^{13}C-NMR** (100 MHz, $CDCl_3$): δ (ppm) 170,6, 170,2, 169.4, 169.3 ($COCH_3$); 158.2 (C ipso); 144.1 (C_q triazole); 133.3 (C para); 127.6 (ArCH); 122.5 (CH triazole); 114.8 (ArCH); 99.9 (C_1); 72.8 (C_4); 71.9 (C_5); 71.3 (C_2); 68.4 (C_3); 68.1 (OCH_2N); 50.4 (OCH_2CH_2); 61.9 (C_6); 30.3-26.0 (CH_2 chain); 20.7-20.5 ($COCH_3$).

ESI-MS: m/z 1343.63 (100% $[M + Na]^+$)

M.p.: 118.3-120.2 °C

10-[4-(methylen-β-D-glucoside)triazolyl]decyloxy-bisphenol (3)

Compound **17** was suspended in CH_3OH and a methanol solution of CH_3ONa was added dropwise till pH 9-10. The reaction mixture was stirred at room temperature for 4-6 h, monitored vi ESI-MS and NMR. When finished Amberlite IR-120 (H^+) added and stirred at rt for 30 min-1 h, removing at the end the solvent under vacuum. The product was obtained without further purification, in 96% yield as a white solid.

^{1}H-NMR (400 MHz, DMSO-d_6): δ (ppm) 8.10 (s, 2H, H triazole); 7.50 (d, 4H, *J* = 8.0 Hz, ArH); 6.96(d, 4H, *J* = 8.8 Hz, ArH); 4.97 (br s, 8H, OH); 4.82 (d, 2H, *J* = 12.4 Hz, OC*H*HN); 4.63 (d, 2H, *J* = 12.4 Hz, OCH*H*N); 4.32 (t, 4H, *J* = 7.2 Hz, OCH_2CH_2); 4.24 (d, 2H, *J* = 8.0 Hz, H_1); 3.97 (t, 4H, *J* = 6.4 Hz, NCH_2CH_2); 3.70 (d, 2H, *J* = 10.0 Hz, H_{6b}); 3.46 (dd, 2H, *J* = 11.6 and 6.0 Hz, H_{6a}); 3.16-3.10 (m, 4H, H_4, H_5); 3.11-3.10 (m, 2H, H_3); 2.98 (t, 2H, *J* = 8.0 Hz, H_2); 1.8-1.83-1.77 (m, 2H, OCH_2CH_2); 1.76-1.68 (m, 4H, CH_2CH_2N); 1.50-1.40 (m, 4H, CH_2CH_2O); 1.38-1.14 (m, 20H, OCH_2CH_2(CH_2)$_5$).

^{13}C-NMR (100 MHz, DMSO-d_6): δ (ppm) 158.2 (C ipso); 144.2 (C_q triazole); 132.6 (C para); 127.6 (ArCH); 124.5 (CH triazole); 114.8 (ArCH); 102.6 (C_1); 72.8 (C_4); 71.9 (C_5); 71.3 (C_2); 68.4 (C_3); 68.1 (CH_2CH_2N); 62.0 (O*C*H_2N); 61.6 (C_6); 50.4 (O*C*H_2CH_2); 30.3-26.0 (*C*H_2 chain).

ESI-MS: m/z 1343.63 (100% $[M + Na]^+$)

M.p.: dec > 120 °C

Liposome preparation

Aqueous dispersions of phospholipid-**1/2** liposomes were prepared according to the procedure described by Hope et al.[77] Lipid films were prepared on the inside wall of a round-bottom flask by evaporation of $CHCl_3$ solutions containing a suitable amount of phospholipid and **1/2** to obtain the desired molar percentage mixture. The obtained lipid films were kept overnight under reduced pressure (0.4 mbar) and a suitable volume of PBS solution (Aldrich, 0.01 M phosphate buffer, 0.0027 M KCl, and 0.137 M NaCl, pH 7.4) was added to obtain a lipid dispersion at the desired concentration. The aqueous solutions were vortex-mixed and then freeze-thawed six times from liquid nitrogen to 60 °C. Lipid dispersions were then extruded (10 times) through a 100 nm polycarbonate membrane.

Optical density measurements

The agglutination of DOPC/**1** (95:5) and of DOPC/**2** (95:5) liposomes (0.58 mM total final conc.) in the presence of Con A (0.24 mg/mL final conc.) in PBS was determined from the time-dependent changes of specific turbidity of 3 mL of samples in a 1 cm quartz cell upon addition of Con A. PBS was used as blank. Scans were carried out at 525 nm immediately after mixing and every minute for 400 min.

MTT assay

The cell citotoxicity was determined using MTT assay. Cells were cultured and seeded in 96-well plate (200 µl per well of $1x10^5$ cells/ml) until the cells attached. Three concentrations were tried: 2 µM, 10 µM and 20 µM. The experiments were done in culture media (control) for 4 h, 24 h and 48 h. After the cells were treated with 100 µl of MTT for 1 h at 37°C. The MTT was removed and then 100 µl of DMSO were added, re-suspending the solution. Absorbances of the samples were measured at 570 nm by using UV-visible spectrophotometer.

Flow Cytometry.

The time course analysis of DOPC liposome/**2** uptake was performed on human breast cancer cells MDA, MCF7 and SKBR3, treated for 30 min, 1 h, 4 h and 24 h with the following liposome formulations: DOPC (1mM) and DOPC/**2** liposomes (1mM), with a $5x10^5$ cells/ml. After each treatment, cells were washed with ice-cold HBSS, detached with EDTA and trypsin, resuspended in icecold PBS and immediately analyzed for the PS content. Fluorescence signals were analyzed with a FACScan flow cytometer, equipped with a 15 mW, 488 nm, and air-cooled argon ion laser. The content of NBD was evaluated as fluorescence intensity and expressed in arbitrary units (AU), calculated as the ratio between the mean fluorescence channel (MFC) of treated samples and the MFC of untreated cells.

Laser scanner confocal microscopy

Living cells were analyzed by laser scanning confocal microscopy (LSCM) in order to investigate the intracellular distribution of the liposome formulations. Cells, grown on 12 mm glass coverslips, were inoculated with the different liposome formulations, under

stirring at 300 rpm. After incubation with liposome formulations (at 20 μM) for 30min, 1 h and 24 h at 310 K, cells were fixed in 4% paraformaldehyde in PBS, for 10 min at room temperature.

2.5 References

(1) Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683–720.

(2) Gabius, H.-J.; Siebert, H.-C.; André, S.; Jiménez-Barbero, J.; Rüdiger, H. *ChemBioChem* **2004**, *5* (6), 740–764.

(3) Ifuku, O.; Koga, N.; Haze, S. I.; Kishimoto, J.; Wachi, Y. *Eur. J. Biochem.* **1994**, *224* (1), 173–178.

(4) Solís, D.; Bovin, N. V.; Davis, A. P.; Jiménez-Barbero, J.; Romero, A.; Roy, R.; Smetana, K.; Gabius, H.-J. *A guide into glycosciences: How chemistry, biochemistry and biology cooperate to crack the sugar code*; Elsevier B.V., 2015; Vol. 1850.

(5) Gorelik E., Galili U, R. A. *Cancer Metastasis Rev.* **2001**, No. 20, 245–277.

(6) Bies, C.; Lehr, C.-M.; Woodley, J. F. *Adv. Drug Deliv. Rev.* **2004**, *56* (4), 425–435.

(7) Minko, T. *Adv. Drug Deliv. Rev.* **2004**, *56* (4), 491–509.

(8) Sharon, N., Lis, H. *Glycobiology* **2004**, *14* (11), 53R – 62R.

(9) Mody, R.; Joshi, S.; Chaney, W. *J. Pharmacol. Toxicol. Methods* **1995**, *33* (94), 1–10.

(10) Dias, R.; Machado, L.; Migliolo, L.; Franco, O. *Molecules* **2015**, *20* (1), 519–541.

(11) Van Damme, E. J. M.; Peumans, W. J.; Barre, A.; Rougé, P. *CRC. Crit. Rev. Plant Sci.* **1998**, *17*, 575–692.

(12) Van Damme, E. J. M.; Lannoo, N.; Peumans, W. J. *Adv. Bot. Res.* **2008**, *48*, 107–209.

(13) Kanai, M.; Mortell, K. H.; Kiessling, L. L. *J. Am. Chem. Soc.* **1997**, *119* (41), 9931–9932.

(14) Gabius, H. J. *Eur. J. Biochem.* **1997**, *243* (3), 543–576.

(15) Drickamer, K.; Street, W.; York, N.; Taylor, M. E. *Cell* **1993**.

(16) Barondes, S. H.; Cooper, D. N. W.; Gitt, M. a.; Leffler, H. *J. Biol. Chem.* **1994**, *269* (33), 20807–20810.

(17) Liu, F.-T.; Rabinovich, G. A. *Nat. Rev. Cancer* **2005**, *5* (1), 29–41.

(18) Goldstein, I. J.; Hughes, C. R.; Monsigny, M.; Osawa, T.; Sharon, N. *Nature*. 1980, p 66.

(19) Yu, L. *Methods Mol. Med.* **1998**, *9*, 379–384.

(20) Kilpatrick, D. C. *9*.

(21) Chrispeels, M. J.; Raikhel, N. V. *Plant Cell* **1991**, *3* (1), 1–9.

(22) Sharon, N. *FEBS Lett.* **1987**, *217* (2), 145–157.

(23) Heberts, D. N.; Carruthersg, A. *J. Biol. Chem.* **1992**, *267* (33), 23829–23838.

(24) Mueckler, M.; Hresko, R. C.; Sato, M. *Biochem. Soc. Trans.* **1997**, *25*, 951–954.

(25) Bell, G. I.; Kayano, T.; Buse, J. B.; Burant, C. F.; Takeda, J.; Lin, D.; Fukumoto, H.; Seino, S. *Diabetes Care* **1990**, *13* (3), 198–208.

(26) Maher, F.; Vannucci, S. J.; Simpson, I. A. *FASEB J.* **1994**, *8* (12), 1003–1011.

(27) Pascual, J. M.; Wang, D.; Lecumberri, B.; Yang, H.; Mao, X.; Yang, R.; Vivo, D. C. De. *Eur. J. Endocrinol.* **2004**, *150*, 627–633.

(28) Macheda, M. L.; Rogers, S.; Best, J. D. *J. Cell. Physiol.* **2005**, *202*, 654–662.

(29) Zhang, Y.; Chan, H. F.; Leong, K. W. *Adv. Drug Deliv. Rev.* **2013**, *65* (1), 104–120.

(30) Bangham, A. D.; Horne, R. W. *J. Mol. Biol.* **1964**, *8* (5), 660–668.

(31) Bozzuto, G.; Molinari, A. *Int. J. Nanomedicine* **2015**, *10*, 975–999.

(32) Immordino, M. L.; Cattel, L. *Nanomedicine* **2006**, *1* (3), 297–315.

(33) Torchilin, V. P. *Nat. Rev. Drug Discov.* **2005**, *4* (2), 145–160.

(34) Gregoriadis, G. *N. Engl. J. Med.* **1976**, *295*, 704–710.

(35) Allen, T. M. *Drugs* **1998**, *56*, 747–756.

(36) Pattni, B. S.; Chupin, V. V.; Torchilin, V. P. *Chem. Rev.* **2015**, *115*, 10938−10966.

(37) Allen, T. M.; Cullis, P. R. *Adv. Drug Deliv. Rev.* **2013**, *65* (1), 36–48.

(38) Clary, L.; Gadras, C.; Greiner, J.; Rolland, J.; Santaella, C.; Vierling, P.; Gulik, A. *Chem. Phys. Lipids* **1999**, *99*, 125–137.

(39) Gu, Q.; Zou, A.; Yuan, C.; Guo, R. *J. Colloid Interface Sci.* **2003**, *266* (2), 442–447.

(40) Wang, T.; Fuhrhop, J.-H. *Chem. Rev.* **2004**, *104*, 2901–2937.

(41) Bücher, C.; Grosse, X.; Rothe, H.; Fiethen, A.; Kuhn, H.; Liefeith, K.; Grosse, X.; Rothe, H.; Heiligenstadt, H. *Biointerphases* **2014**, *9* (1), 011002.

(42) Gliozzi, A.; Relini, A.; Chong, P. L. *J. Memb. Sci.* **2002**, *206*, 131–147.

(43) Grinberg, S.; Kipnis, N.; Linder, C.; Kolot, V.; Heldman, E. *Eur. J. Lipid Sci. Technol.* **2010**, *112* (1), 137–151.

(44) Wamberg, M. C.; Wieczorek, R.; Brier, S. B.; de Vries, J. W.; Kwak, M.; Herrmann, A.; Monnard, P.-A. *Bioconjug. Chem.* **2014**, *25* (9), 1678–1688.

(45) Lian, T.; Ho, R. J. *J. Pharm. Sci.* **2001**, *90* (6), 667–680.

(46) Badjić, J. D.; Nelson, A.; Cantrill, S. J.; Turnbull, W. B.; Stoddart, J. F. *Acc. Chem. Res.* **2005**, *38* (9), 723–732.

(47) Mcnicholas, S.; Rencurosi, A.; Lay, L.; Mazzaglia, A.; Sturiale, L.; Perez, M.; Darcy, R. *Biomacromolecules* **2007**, *8*, 1851–1857.

(48) Zhang, H.; Ma, Y.; Sun, X. *Chem. Commun.* **2009**, 3032–3034.

(49) Jayaraman, N.; Maiti, K.; Naresh, K. *Chem. Soc. Rev.* **2013**, *42*, 4640–4656.

(50) Zhu, J.; Xue, J.; Guo, Z.; Zhang, L.; Marchant, R. E. *Bioconjug. Chem.* **2007**, *18*, 1366–1369.

(51) Chono, S.; Kaneko, K.; Yamamoto, E.; Togami, K.; Morimoto, K. *Drug Dev. Ind. Pharm.* **2010**, *36* (1), 102–107.

(52) Mauceri, A.; Borocci, S.; Galantini, L.; Giansanti, L.; Mancini, G.; Martino, A.; Manni, L. S.; Sperduto, C. *Langmuir* **2014**, *30*, 11301–11305.

(53) Gerber, S.; Garamus, V. M. *Langmuir* **2005**, No. 14, 6707–6711.

(54) Ferna, G.; Robina, I. **2001**, *42*, 5413–5416.

(55) Aleandri, S.; Casnati, A.; Fantuzzi, L.; Mancini, G.; Rispoli, G.; Sansone, F. *Org. Biomol. Chem.* **2013**, *11*, 4811–4817.

(56) Dondoni, A.; Marra, A. *Chem. Rev.* **2010**, *110*, 4949–4977.

(57) Edition, I. *Russ. Chem. Bull. Int. Ed.* **2013**, *62* (3), 577–604.

(58) Aleandri, S.; Casnati, A.; Fantuzzi, L.; Mancini, G.; Rispoli, G.; Sansone, F. *Org. Biomol. Chem.* **2013**, *11* (29), 4811–4817.

(59) Moses, J. E.; Moorhouse, A. D. *Chem. Soc. Rev.* **2007**, *36* (8), 1249–1262.

(60) Lutz, J. *Angew. Chemie - Int. Ed.* **2007**, *46*, 1018–1025.

(61) Huisgen, R. *Angew. Chemie - Int. Ed.* **1963**, *2* (10), 565–632.

(62) Hein, J. E.; Fokin, V. V. *Chem. Soc. Rev.* **2010**, *39* (4), 1302.

(63) Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angewandte Chemie - International Edition*. 2001, pp 2004–2021.

(64) Bernardi, S.; Fezzardi, P.; Rispoli, G.; Sestito, S. E.; Peri, F.; Sansone, F.; Casnati, A. *Beilstein J. Org. Chem.* **2014**, *10*, 1672–1680.

(65) Moni, L.; Rossetti, S.; Scoponi, M.; Dondoni, A. *Chem. Commun.* **2010**, *46*, 475–477.

(66) Kappe, C. O. *Angew. Chemie Int. Ed.* **2004**, *43* (46), 6250–6284.

(67) Aragão-Leoneti, V.; Campo, V. L.; Gomes, A. S.; Field, R. a.; Carvalho, I. *Tetrahedron* **2010**, *66* (49), 9475–9492.

(68) Kolb, H. C.; Sharpless, K. B. *Drug Discov. Today* **2003**, *8* (24), 1128–1137.

(69) Garcia-Hartjes, J.; Bernardi, S.; Weijers, C. a G. M.; Wennekes, T.; Gilbert, M.; Sansone, F.; Casnati, A.; Zuilhof, H. *Org. Biomol. Chem.* **2013**, *11* (26), 4340–4349.

(70) Dalvie, D. K.; Kalgutkar, A. S.; Khojasteh-bakht, S. C.; Obach, R. S.; Donnell, J. P. O. *Chem. Res. Toxicol.* **2002**, *15* (3), 269–299.

(71) Sansone, F.; Chierici, E.; Casnati, A.; Ungaro, R. *Org. Biomol. Chem.* **2003**, *11*, 1802–1809.

(72) Sansone, F.; Baldini, L.; Casnati, A.; Ungaro, R. *Supramol. Chem.* **2008**, *20*, 161–168.

(73) Lis, H.; Sharon, N. *Chem. Rev.* **1998**, *98* (2), 637–684.

(74) Kang, S. K., Kim, W. S., Moon, B. H. *Synthesis (Stuttg).* **1985**, 1161–1162.

(75) Chong, J. M.; Heuft, M. A.; Rabbat, P. *J. Org. Chem.* **2000**, No. 65, 5837–5838.

(76) Ren, B.; Wang, M.; Liu, J.; Ge, J.; Zhang, X.; Dong, H. *Green Chem.* **2015**, *17* (3), 1390–1394.

(77) Hope, M.; Nayar, R.; Mayer, L. *Liposome Technol.* **1993**, *I*, 123–139.

(78) André, S.; Sansone, F.; Kaltner, H.; Casnati, A.; Kopitz, J.; Gabius, H. J.; Ungaro, R. *ChemBioChem* **2008**, *9* (10), 1649–1661.

(79) Yamakoshi, H.; Kanoh, N.; Kudo, C.; Sato, A.; Ueda, K.; Muroi, M.; Kon, S.; Satake, M.; Ohori, H.; Ishioka, C.; Oshima, Y.; Osada, H.; Chiba, N.; Shibata, H.; Iwabuchi, Y. *ACS Med. Chem. Lett.* **2010**, *1* (6), 273–276.

(80) Surov, O. V; Voronova, M. I.; Barannikov, V. P.; Shaposhnikov, G. P. *J. Phys. Chem. C* **2014**, *118*, 338–345.

(81) Rajaganesh, R.; Ravinder, P.; Subramanian, V.; Mohan Das, T. *Carbohydr. Res.* **2011**, *346* (15), 2327–2336.

(82) Su, Y.; Xie, J.; Wang, Y.; Hu, X.; Lin, X. *Eur. J. Med. Chem.* **2010**, *45* (7), 2713–2718.

(83) Ma, D.-L.; Shum, T. Y.-T.; Zhang, F.; Che, C.-M.; Yang, M. *Chem. Commun.* **2005**, No. 37, 4675–4677.

Chapter 3

Glycocalix[n]arenes for immunostimulation

3.1 Introduction

3.1.1 Carbohydrates: eliciting the immune response

Among the roles played by carbohydrates in biology, one of very high relevance is their involvement in the activation of the immune response.[1] In fact sugars are among the most abundant molecules on viruses and bacteria cell surfaces and altered cells can overexpress certain typologies of saccharides respect the physiological levels or even produce unique saccharide sequences different from those physiologically present on the cell coating defined as glycocalyx. Particularly the recognition of carbohydrates from proteins is the initial, fundamental step for infection and colonization. If saccharides belonging to glycoconjugates can constitute a tool for microorganisms to adhere to cells and start the infection or for damaged cells, such as tumor cells, to spread in the body, on the other hand they can also represent a specific signal of their presence, acting as antigen determinants, that elicits the activation of the defense system in humans. In fact pathogens are characterized by the presence of the so called Pathogen Associated Molecular Patterns (PAMPs), often including carbohydrate structures, that allow the host organism to sense their presence. PAMPs are structurally and chemically different compounds highly conserved in pathogens and absent in their hosts. The host organism owns PRRs (Pattern-Recognition Receptors), including the recently discovered TLRs (Toll-Like Receptors), which recognizing PAMPs, trigger the innate immune response, the first defence mechanism, which acts during the early stages of infection (minutes). This pathway is mediated by antigen-presenting cells (APCs), especially dendritic cells, plentiful of PAMPs and it is very sensitive and selective. In fact foreign molecules (antigens) can be recognized also in the pM range.

After that, the second line of defence, the adaptive immune response, starts with its action, mediated by B and T cells (lymphocytes), that needs long period (from days to weeks) to become effective, but it clears definitely the infection. These lymphocytes recognize pathogens with high affinity, providing the fine antigenic specificity required for complete removal of the infective agent and the creation of the immunological memory.

Antigen, for definition, is every compounds perceived by the immune system as a “foreign” or something potentially dangerous for the host. The immune system responds to antigens by enhancing a suitable immune response. Most antigens, such as proteins and derivatives,

are T-dependent: this means that they induce T-cell activation by interacting with B cells through MHC (Major Histocompatibility Complex restricted pathways).[2] The memory B cells survive for a long time and reply immediately to subsequent exposures of antigen by secreting high-affinity IgG antibodies. On the contrary immune response triggered by carbohydrates antigens is T-cell-independent. This means that they induce T-cell activation through MHC (Major Histocompatibility Complex restricted pathway), carbohydrates are T-cell-independent. This means that oligo- and polysaccharides are able to activate B cells, without cooperation by T cell[3] by binding to receptors of B-lymphocytes and inducing cross-linking of the Ig proteins, which leads to activation of the B cell and production of only low affinity antibodies, as immunoglobulin M (IgM) antibodies. These T cell-independent responses are less robust and short-lived, without creation of immunological memory.

It is evident that carbohydrates can be exploited as therapeutic agents in the activation of the immune response. Especially in the field of vaccines production, the possibility to enhance the potency of carbohydrates can be reached by converting T-cell independent polysaccharides in T-cell dependent immunogens .

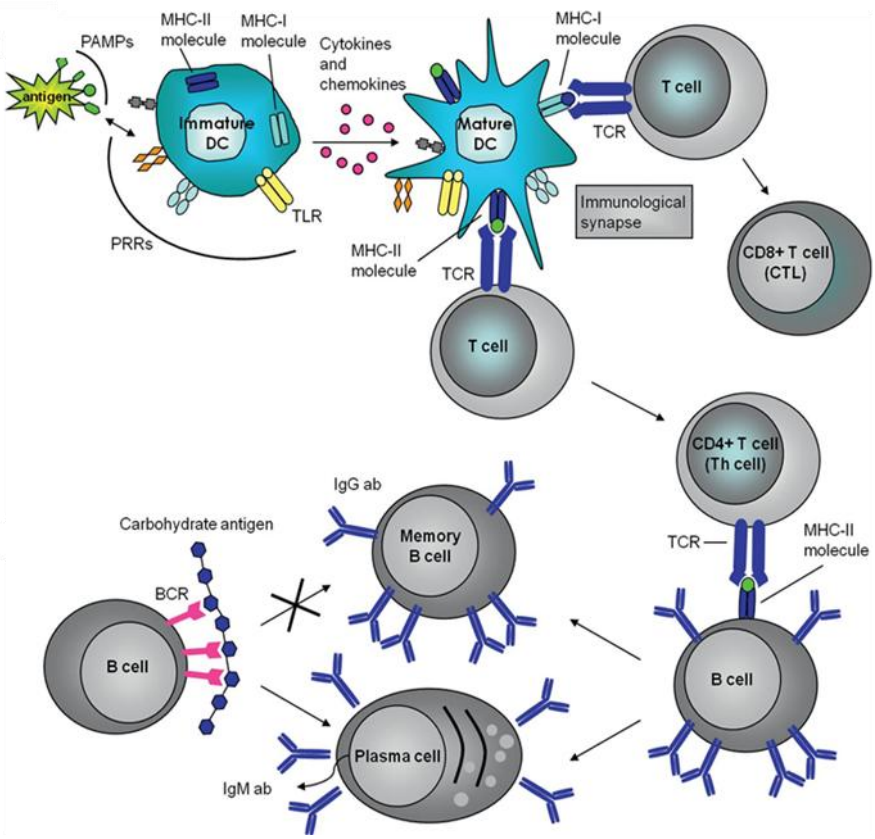


Fig. 1: schematic representation of the immune response (from Eur. J. Org. Chem. 2011, 5723–5777).

3.1.2 Vaccines based on carbohydrates

Vaccines[4] (from latin "vacca", cow) are one of the greatest goal of medicine against infectious diseases. Their introduction have greatly increased the levels of the public health. In fact the practice of vaccination is the key for the control of many infectious diseases and it was introduced in 1796 in the Western world by Edward Jenner with the smallpox vaccine, the first applied successfully. From this first example many efforts were and are currently done to fight fungal[5], bacterial[6,7,8], viral[9] infections and also cancers.[10,11]

Traditional vaccines include living attenuated or killed microbes, or microbial components. However the use of these conventional vaccines presents many drawbacks such as limited shelf life, potential side effects, poor immune responses and safety problems.

In order to overcome these limitations in the last decades the development of carbohydrate-based vaccines represents one of the most effective and safe routes for a preventative therapy against infections. Many sugar structures are pathogen-specific, which makes them important molecular targets for vaccine development, diagnosis of infectious diseases and pathogen recognition[12]. Nonetheless, also the use of saccharide-based vaccines is not free of problems. Firstly they are poorly immunogenic: in the elderly and in children (under 2 years of age) they induce only short-lasting immune response and do not create B-cell-mediated immunological memory.[1] Consequently the low efficacy of these traditional vaccines is widely due to the T-cell independent immune response, which is typically triggered by carbohydrate antigens. To enhance the polysaccharide immunogenicity there are different approaches. The first one was the conjugation of carbohydrates moieties to a carrier protein, such as mutant diphtheria toxin (CRM197), tetanus toxoid (TT), diphtheria toxoid (D), and outer membrane protein (OMP), since the T-cell dependent responses are typically created by these macromolecules and peptide epitopes.[13] Examples of this approach are many in literature[7,14,15] and the first glycoconjugate vaccine, composed of a glycan conjugated to a carrier protein, to be used in humans was a conjugate against *Haemophilus influenzae* type b (Hib),[16] and was licensed in the USA in 1987. These type of vaccines are the so-called semi-synthetic vaccines.[17]

Others drawbacks are the chemical heterogeneity of carbohydrates-based antigens necessary for the inclusion in vaccines and also the poor availability from natural sources. This means difficult isolation and purification techniques and lack of sufficient amount of

material for the development of this field. The creation of fully synthetic vaccines[2,14,17] can help to go beyond this problem, because they can have defined composition, high reproducibility and no significant issues in the isolation and purification steps, since they do not need extraction from natural sources. Furthermore the sugar synthesis have achieved important breakthrough in the last years.

A third disadvantage is the antibody insufficient low affinity for carbohydrates. In fact the carbohydrate-specific IgG antibodies, produced upon T and B cell activation, typically bind to the corresponding carbohydrates epitopes with dissociation constants (K_d) in the mM range. This is a very soft complex compared with protein-specific antibodies with dissociation constants in the range of nM. This problem could be overcome by the creation of glycoclusters bearing many units of the saccharide antigens to decrease the dissociation constant exploiting multivalency.[18,19]

The last drawback is correlated to the stability of these carbohydrate-based vaccines. In fact the lability of polysaccharide structures (for example cleavage of linkages in water) is one of the problems associated with conventional vaccines that necessitate transporting and storing in cold conditions due to their limited shelf-life at room temperature.[20] In order to exceed this obstacle, stable molecules with long shelf-lives are required. Synthetic analogues of natural polysaccharides, endowed both with the immunological properties of the natural compounds and with an increased stability in water, can be the key to overcome lability problems.

3.1.3 Semi-synthetic and synthetic carbohydrate-based vaccines

The reaction at the base of the creation of new glycoconjugates is the glycosylation, reaction that allows to bind the anomeric carbon of a carbohydrate (glycosyl donor) to the hydroxyl or other functional group of another molecule (glycosyl acceptor). In biology it is usually considered as an acceptor a protein (as in semi-synthetic vaccines), a lipid, or other organic molecules (as in synthetic vaccines).

Glycoconjugate vaccinology started from the discovery by Oswald Avery and Walter Goebel in 1929 that a specific polysaccharide could work as an antigen when covalently coupled with some other constituent of the cells, such as proteins. Two different sugars, glucose and galactose, were employed in the preparation of these glycoclusters as vaccines. Distinct

serological responses, depending on the carbohydrate component, were obtained and this was a demonstration of the specificity of the antigen-antibody interactions.[21]

The possibility to enhance the immunogenicity of polysaccharide antigens was also discovered by Avery and Goebel in the same year. They demonstrated that the poor immunogenicity of purified *S. pneumoniae* type 3 polysaccharide in rabbits could be raised by conjugation of the polysaccharide to a protein carrier.

Since the first glycoconjugate vaccine against *Haemophilus influenzae* type b[16] had a great success, reducing the incidence of this invasive disease in childhood,[22] significant progress has been made in vaccines research. [13,23]

This discovery has allowed the development of semi-synthetic and fully synthetic vaccines based on saccharides, where the common element of all is the presence of a synthetic carbohydrate-based antigen in the form of single glycan or displayed as glycocluster.

Belonging to the first one, there is a high number of carbohydrate vaccines in clinical use today composed of several copies of saccharidic antigen conjugated to an antigenic carrier protein, since the T-cell dependent responses are typically created by proteic and peptidic epitopes.[13]

In the case of very efficient vaccines the carbohydrate epitope is composed of fragments of bacterial capsular polysaccharides, for instance polysaccharides from *Streptococcus pneumoniae, Neisseria meningitidis and Haemophilus influenza* conjugated to the carrier protein that sometimes is itself a TLR ligand. These formulations have been found to provide protection against different strains of pathogens.[4]

Fully synthetic vaccines incorporate instead in the same molecule the synthetic glycan, a synthetic peptide as T cell epitope and, sometimes, a third chemical moiety with the function of internal adjuvant. This component called adjuvant is common in both the categories, with the aim to amplify the immune response, due to the poor immunogenicity of carbohydrates antigen. In this way to improve the quality and efficiency of therapeutic effects, the conjugation of this moiety to carbohydrates base vaccines is needed. Particularly, adjuvant role consists in being perceived as “danger signals”[2] after binding the PRRs, increasing the duration of both innate and adaptive immune response. Adjuvants can be also delivery systems to localize vaccines and to target the APCs. In semi-synthetic vaccines carbohydrates can play the double role of adjuvant/antigen and adjuvant is not

covalently bound whereas in fully synthetic vaccines, if present, there is a covalent binding to the epitope (Figure **2**).

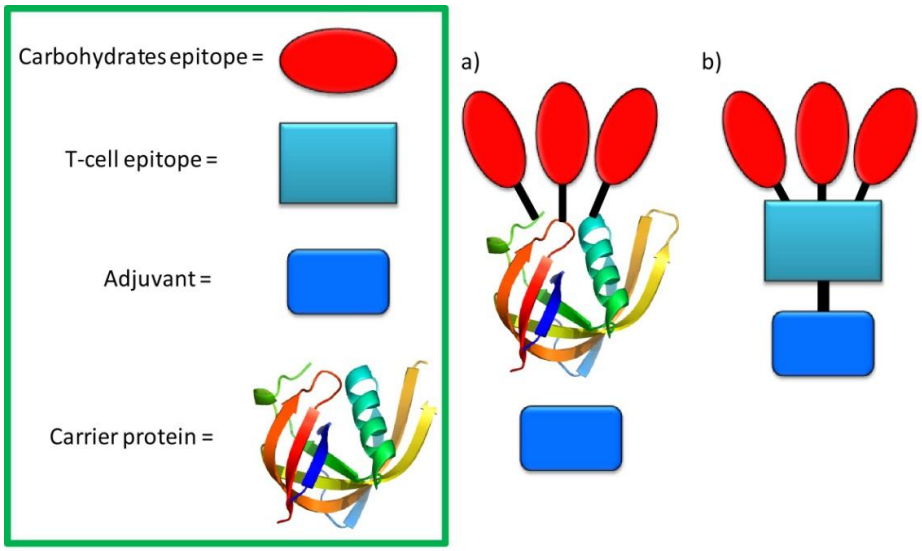


Fig.2: schematic representation of differences between a) semi-synthetic and b) fully synthetic vaccines

For the preparation of the acceptor protein a not simple procedure is often required: for example, the carrier protein CRM197 used to be isolated from the culture supernatant of *Corynebacterium diphtheria* and needs appropriate purification steps.[24]

Then, in order to connect polysaccharides to the carrier protein the accepted technique is chemical conjugation, where polysaccharides are chemically activated at random sites, or alternatively at the reducing end.[25]

Furthermore if natural polysaccharides are used, it is firstly required to strip those from the pathogenic organism against which the vaccine is prepared, removing contaminating endotoxins with long purification steps.

It is easy to understand that a total control on the structure can be achieved in fully synthetic vaccines, while protein conjugation, polysaccharide purification and even external adjuvant are still sources of variability in semi-synthetic vaccines.

In fact fully synthetic carbohydrate-based vaccines can overcome all the disadvantages of semi-synthetic analogues. Thanks to the enormous steps done in the carbohydrates synthesis, the construction even of complex oligosaccharides portions is then allowed.[26] In addition, they must have proper purity and reliable structural integrity and stability for clinical studies and administration.

For all these reasons fully synthetic vaccines were preferred in the last years to the semi-synthetic analogues.

3.1.4 Encapsulated bacteria: *Streptococcus pneumoniae* and its 19F serotype

As said in the first paragraph, the cell surface of many pathogens is covered by oligo- and polysaccharides, that act as antigens to elicit anti-carbohydrate immune response and virulence factors. Particularly many bacteria are coated by a thick layer of carbohydrates, called capsule polysaccharide (CPS), that is responsible of their pathogenicity, with protective functions against the host's immune defence.[27] These microorganisms are named encapsulated bacteria and are the third leading cause of death in the world. The capsule plays several roles in these bacteria: 1) helps the pathogens to invade the host taking advantage of the adhesive properties conferred by some sugar such as hyaluronates; 2) mimics molecules produced by the cells host, so the pathogen is not recognize by the host immune response as a danger (for example the case of Men B polysaccharide); 3) offers substantial protection to bacterial pathogens against phagocytosis by migrating phagocytes and tissue-fixed macrophages; 4) gives protection from desiccation due to the hydrophobic character. Furthermore the capsular polysaccharides are powerful TLRs activators.[28]

Among encapsulated bacteria one of the most diffuse is the *Streptoccoccus pneumoniae* (SP)[29], a Gram-positive, facultative anaerobic member (**Fig 3**). It is a principal agent for bacterial pneumonia, bacteremia, otitis media, meningitis and has caused great mortality throughout the world, especially in young children and the elderly.

Fig. 3: *Streptoccoccus pneumoniae* and its polysaccharide capsule

The polysaccharide capsules are composed of repeating saccharide units and CPS structure defines the capsular type of the organism (serotype). Actually in the same family of bacteria every serotype is characterized by the own CPS. On the basis of differences in CPS chemical structure, pneumococci can be divided in more than 90 serotypes. The prevalence of some serotypes to others depends on the geographical distribution and differ among the United States, Europe, and some Asian countries.[30]

In children, a few types are responsible for a large proportion of pneumococcal diseases. In one study, the most important pediatric serotypes (6A, 14, 19F, and 23F) were responsible for almost 60% of all infections. In adults, however, serotypes 3, 19F, and 6A counted for only 31% of the isolations.[31]

Among these serotypes, the 19F is one of the most virulent. It is responsible for a large number of infections of the upper respiratory system and meningitis, especially in children and immunodeficient subjects. It is characterized by the presence of a trisaccharide repetitive unit reported in Figure **4**, constituted by a units of N-acetylmannosamine, bounded by a β 1-4 linkage to a glucose, connected with α 1-2 linkage to a rhamnose with a phosphodiester bridge associating a residue of the last sugar at the reducing end of one unit to the non-reducing end of the following unit.

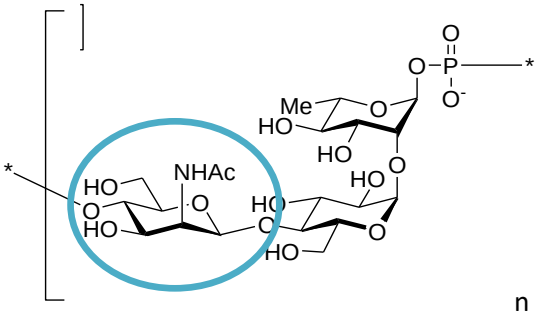


Fig.4: repeating units of Streptococcus pneumonia 19F with β-N-acetyl-mannosamine encircled.

3.1.5 Vaccines against Streptococcus pneumoniae

Many efforts in the development of vaccines against this bacterium were done during the last decades.[32] The use of a highly conserved, immunogenic, and protective pneumococcal protein antigen would be a solution, after choosing the right biomacromolecule, to create new vaccines against *Streptococcus pnuemoniae*.[33] Since the discovery that CPS fragments induce a response of the host's adaptive immune system, research has been focused on developing CPS-based vaccines as an alternative to antibiotic therapy.[34,35]

The idea to commercialize vaccines based on different polysaccharide capsules due to the large number of serotypes has increased the attention the last years. Firstly polysaccharide-based vaccines were produced: as first 14-valent product licensed in 1977[36] that was efficacious for people 2 years or older with certain chronic health conditions that increased the risk of pneumococcal infection. After that a 23-valent vaccine (PPV23) by both Merck and Lederle was licensed subsequently in 1983. The last one is composed of capsular polysaccharide purified from the 23 most dominant serotypes among clinical isolates of *S. Pneumoniae.* This vaccine does not provide sufficient protection in children younger than 2 years old, who are the most vulnerable to pneumococcal infection. Approximately the efficacy of this is 65% against pneumococcal infections but is only 0-60% against all serotypes.[33] Since polysaccharide antigens are alone poor immunogenic in infants and do not induce immunological memory, scientists moved on glycoconjugate vaccines (as explained in paragraph 3.1.2). In the semi-synthetic vaccines family it was created a 7-valent PS-protein conjugate vaccine, called PCV7, composed of nontoxic diphtheria toxin mutant protein CRM197 conjugated with CPS of seven serotypes that are predominant in children. Even if with the conjugation to a carrier protein the immunity is enhanced, there is a limit in the

efficacy of this vaccine: the specificity of the elicited antibody. Therefore it cannot protect sufficiently populations where the most present serotypes are different from those in the vaccine formulation. The scientists tried to discover more efficient vaccines: for example 9- and 11-valent conjugate vaccines were developed[37,38] even if the cost of vaccine formulation increased raising the number of serotypes in the same drug.

Therefore the possibility to prepare new vaccines able to provide optimum immunological protection in vaccinated individuals against a single serotype, could be a so far unused approach. As seen before the goal is the introduction of fully synthetic carbohydrate-based vaccines, that are more specific, selective to fight different bacterial infections.[39,2]

Against serotype 19F efforts to synthesize the repeating units were done,[40,41] using the trisaccharide group in figure **4** or structural analogues.[42,43] In fact it was observed that also the single trisaccharidic repeating unit shows a good ability to inhibit the binding between the 19F polysaccharide and the anti-19F human polyclonal antibody, even if with stability, efficacy and affinity lower than the natural polysaccharide.[44,45] The aim of the researchers is then the possibility to realize molecules endowed with the natural biological activity, but with higher efficacy and even stability against this virulent serotype of *S. pneumoniae*.

3.1.6 Glycocalixarenes as immunostimulants against SP serotype 19F

Glycoconjugate vaccines provide enormous health benefits and superior immunity compared with polysaccharides alone, especially in the less than 5-year-old age group, the elderly and the immune compromised.

Since the synthesis of oligosaccharides containing certain numbers of the repeating unit of the natural polysaccharide, needs very long time and attempts, multivalency can be used to overcome this problem. In fact instead of preparing complex polysaccharides, only a simplest part of the carbohydrate antigens can be chosen and it will be repeated n times depending on the valency of the multivalent structure.

Multivalency can give also benefits increasing the lower affinity that characterized the carbohydrate-antibody interactions. In fact it is well-known that antibodies have multiple binding sites and so they can interact with other molecules via polyvalent interactions[46]. If

this binding is strong enough the process of the activation of the immune response can be enhanced.

The design of multivalent ligands able to have a multivalent presentation of antigens for the elicitation of the immune response is the aim of this chapter. Connecting all the advantages of calixarenes as multivalent scaffolds, already described in chapter 1, and of fully synthetic carbohydrate-based vaccines (reported in paragraph 3.1.3), we decided to prepare a small library of glycocalix[n]arenes for the creation of synthetic vaccine against *Streptococcus pneumoniae* type 19F. We have focused the attention on the capsular polysaccharide of this serotype, because is one of the most commonly isolated serotypes causing pneumococcal disease.[20] The exposition of multiple copies of the same saccharide unit (antigen), given by glycocalixarenes, can allow a stronger binding with the antibody.

The research was focussed on calix[4]- [6]- and [8]arenes as multivalent scaffolds for the preparation of carbohydrate-based vaccines against SP 19F. A small library of compounds was prepared. Earlier it was effectively observed that changing the valency and the exposition geometry of the saccharide units, the binding efficiency between multivalent system and the corresponding receptor can be influenced significantly.[47,48] This could determine selectivity or specificity in the antigen-antibody interaction.

A monomeric analogue was also prepared to demonstrate the possible importance of the macrocyclic structure in the interaction with antibody.

Furthermore the exposition of the antigens in these glycoclusters will be parallel, not in series as happened in the natural polysaccharide (see Figure **5** below). In this way the number of antigenic molecules will be limited respect to the native polysaccharide to see if this parallel fashion can be able to counterbalance the high decrease of antigenic units. Same examples based on this logic are already present in literature.[49,50]

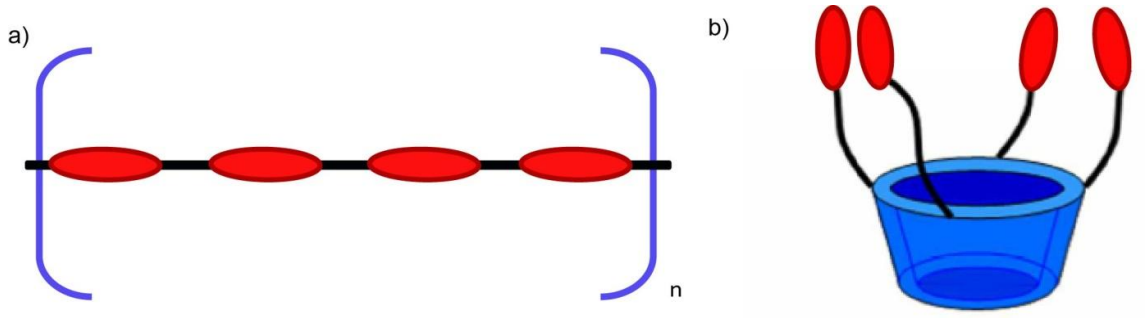


Fig.5: different exposition geometry of saccharidic units: a) in series, as in the natural polysaccharide and b) in parallel as in a glycocalixarene.

Since the synthesis of the trisaccharide unit requires many efforts and the amount of this natural antigen is not enough for the functionalization of all the prepared calixarene scaffolds, it was applied a further simplified approach to obtain in a rapid way multivalent glycocalix[n]arenes to be studied in the immunostimulation experiments. In this first phase of these studies, the calixarene structures were decorated with a minimal analogue structure of the natural polysaccharide: the monomeric unit of N-acetyl-D-mannosamine (highlighted in figure **4**). Effectively this sugar seems playing the pivotal role in the activation of the immune response compared with the other two sugars belonging to the trisaccharidic repeating unit. Hence, once understood the best glycocalixarene structure in terms of capability to enhance the immune response, this glycocluster will be functionalized with the trisaccharide.

The reaction used for the conjugation step is between the isothiocyanate group, present at the upper rim of the calixarenes, and the amino moiety presented on the sugar, to generate a thiourea linker, using a "click"reaction.[51] It arises the family of glycosylthioureidocalixarenes.[52,53] The employment of the thiourea-bridging strategy, actually one of the most conventional methodologies for the design and synthesis of multivalent glycoconjugates,[54] is due to several benefits.

Firstly this type of conjugation does not involve bond formation as the anomeric carbon of the glycoside, preventing stereochemistry problems that can occur during glycosylation reaction. Another advantage in using thiourea is the increase in the potential hydrogen bonding,[55] thanks to the two urea NH groups that could help in the binding with the antibody. Even the water solubility of the thiourea-bridged adducts can be raised.

In this way five thiourea-bridged derivatives were prepared and they are shown in Figure **6.**

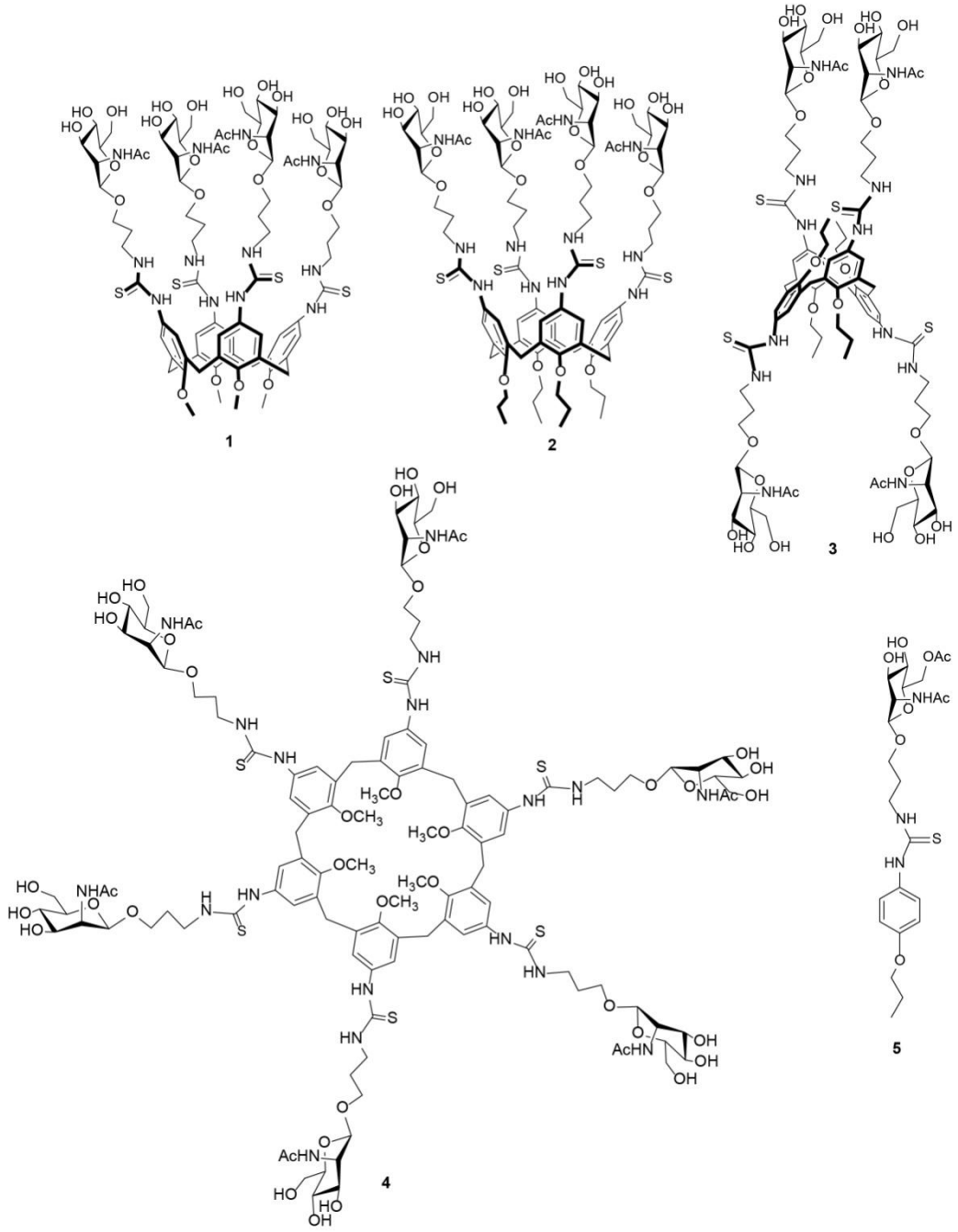


Fig.6: Schematic representation of glycosyl-thioureido calixarenes and the glycomonomer prepared for biological studies against *S. pneumoniae* 19F.

The glycocalix[n]arenes can be divided in three group 1) cone calix[4]arene, 2) 1,3-alternate calix[4]arene and 3) conformationally mobile calix[4]- and calix[6]arenes.

As the IUPAC name of these derivatives would be very long and complicated, in order to easily recognize these compounds in this chapter the code *conf*-mN-ManAc[n]alk will be used. This can be codified as: the conformation *conf* of the calixarene when well defined (in the case of cone and 1,3-alternate), the number m of saccharide units and their name, N-acetylmannosamine (N-ManAc), the size of the macrocycle [n] and the alkyl chain (alk) present on the phenolic oxygens.

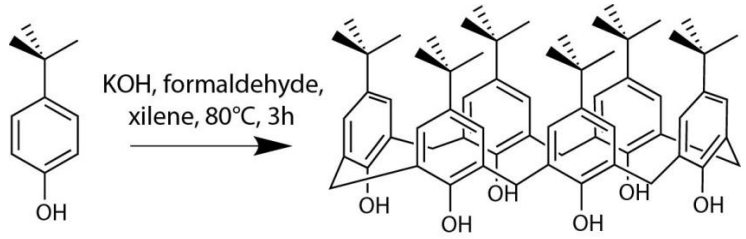
3.2 Results and discussion

3.2.1 Synthesis of the calix[n]arene scaffolds

As mentioned above, for our purposes calixarenes with different valency and conformational properties were necessary. Size of the macrocycle and functionalization at the lower rim influence the conformational mobility of the central scaffold, and consequently the directionality of the peripheral binding units. Compounds **10a-e** were then identified as proper calixarene scaffolds to be conjugated with the sugar moieties. The cone derivative **10b**, with propoxy groups at the lower rim can bring to a glycocalixarene exhibiting the ligating units in the same part of the space, while the 1,3 alternate isomer **10c** to a cluster exposing the sugars in two different and opposite positions, and they are both fixed in a defined conformation. The calix[4]arene **10a** bearing at the lower rim methoxy groups is instead conformationally mobile, as well as calix[6]- **10d** and calix[8]arene **10e**. In fact conformational interconversion process that is fast also on NMR time scale is possible thanks to the small methoxy groups. These are endowed with a reduced steric hindrance and allow rotation of the phenolic units through the annulus identifiable with the methylene bridges despite the size of the substituents at the upper rim. These flexible derivatives were prepared because they could exploit a beneficial induced fit effect in their approach to the target. This means that they can adapt their conformation in the presence of the antibody, optimizing the binding process, even if, on the other hand, this high mobility leads to a low degree of preorganization of the multivalent ligand, that could have an unfavourable effect on the total free energy ΔG of binding to the antibody. Furthermore compounds **10d** and **10e** should bring to glycoclusters with increased valency respect to the calix[4]arenes.

Moving to the synthesis, **Scheme 2** shows the steps employed to attain the isothiocyanate-functionalized macrocyclic cores **10a-e**.

For only p-*tert*-butyl calix[6]arene the reaction started from the beginning, hence from the condensation between p-*tert*-butyl phenol and formaldehyde, with a well-known procedure, using KOH as base (**Scheme 1**).[56]



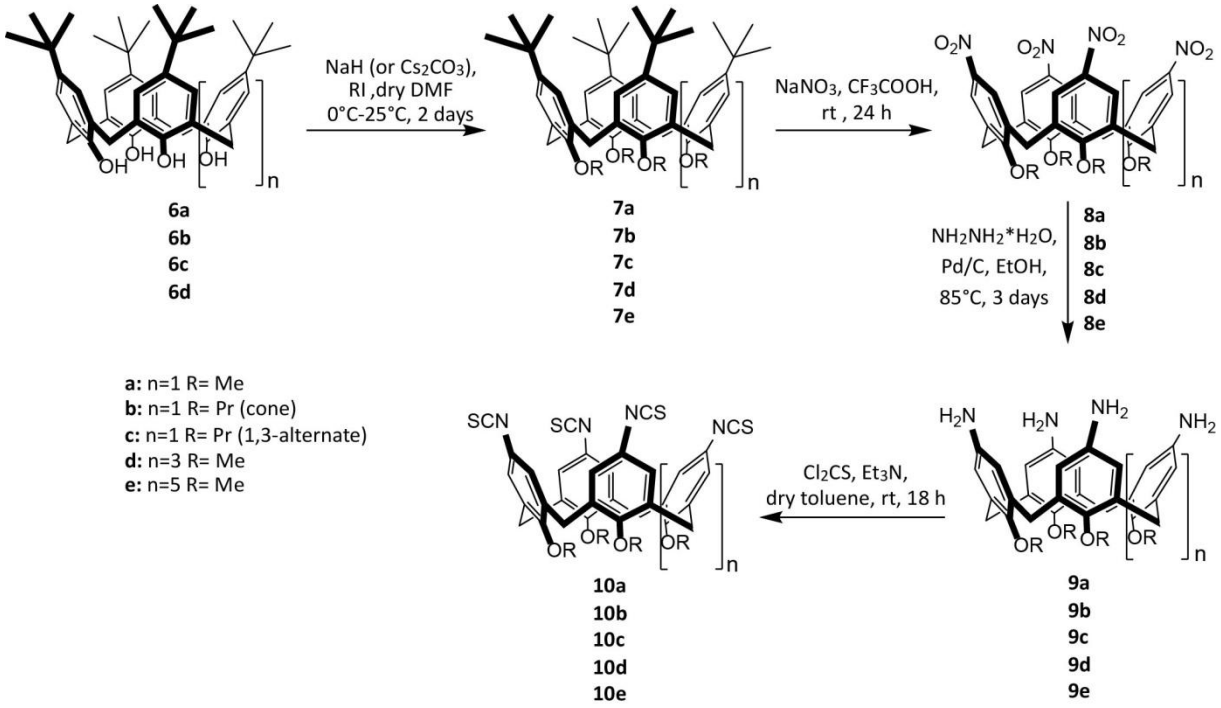
6d

Scheme 1: synthesis of p-*tert-butyl*calix[6]arene **6d**

To prepare all the calix[n]arenes derivatives **10a-e** necessary to anchor the sugar moieties, the synthesis started (except for derivative **7e** already present in the lab[57]) with the alkylation of the hydroxyl groups of the corresponding p-*tert*-butylcalix[n]arenes, changing the alkylating reagents, from methyl to propyl iodide, depending on the R group needed at the lower rim, and the base depending on the desired geometry for the calix[4]arene. So NaH was used as base for the preparation of all the derivatives apart the 1,3-alternate **7b** for which Cs_2CO_3 was necessary.

After that, the ipso-nitration, using $NaNO_3$ and TFA, gave rise to the nitro calix[n]arenes **8a-e**, confirmed by the disappearance of the *t*-butyl signals at around 1 ppm in their ^{1}H-NMR spectra. Diagnostic of the new functionalization at the upper rim is also the marked downfield shift of the singlet relative to the aromatic protons due to the presence of a strong electron-withdrawing group such as NO_2. It may be interesting to say that the ^{1}H-NMR spectrum in $CDCl_3$ of **8a** evidenced the presence in equilibrium and in slow exchange on the NMR timescale of two conformers identified as cone and partial cone in a 1:9 ratio. The subsequent total reduction of the nitro groups by using hydrazine and Pd/C in H_2 atmosphere, permitted the achievement of the amino compounds **9a-e**.[47] The nature of the products was mainly confirmed by the strong upfield shift of the ^{1}H NMR signal for the

aromatic protons respect to the frequency resonance observed for the corresponding nitrocalixarene precursors **8a-e**.



Scheme 2: Synthesis of isothiocyanate calix[n]arenes **10a-e.**

The last reaction at the upper rim of the macrocycles was the transformation of the amines of derivatives **9a-e** in isothiocyanate groups. The low nucleophilicity of aromatic amines could in fact limit the use of amino calixarenes, such as **9a-e**, in coupling reactions. Furthermore, as told before, this is a study at the first stage, where the calixarenes are functionalized only with the monomeric units of the SP 19F thrisaccharide, in order to simplify the synthetic efforts. Once choosen the best "glycocalix[n]arene" towards anti-19F antibody, the trisaccharide will be used as ligating unit to functionalize the macrocycle. However the formation of NCS on this more complex carbohydrate is very difficult and so the presence of isothiocyanate on the calixarene scaffolds was required to avoid this subsequent problems. For these reasons we opted for the coupling strategy involving a sugar functionalized with an amino group and calixarenes in the form of isothiocyanate. The synthesis of isothiocyanate calix[4]arenes **10a-e** was carried out using thiophosgene and triethylamine in dry toluene. Derivatives **10a-c** were easily obtained and their achievement was confirmed by ESI-MS and a new significant shift, this time to lower field, of the

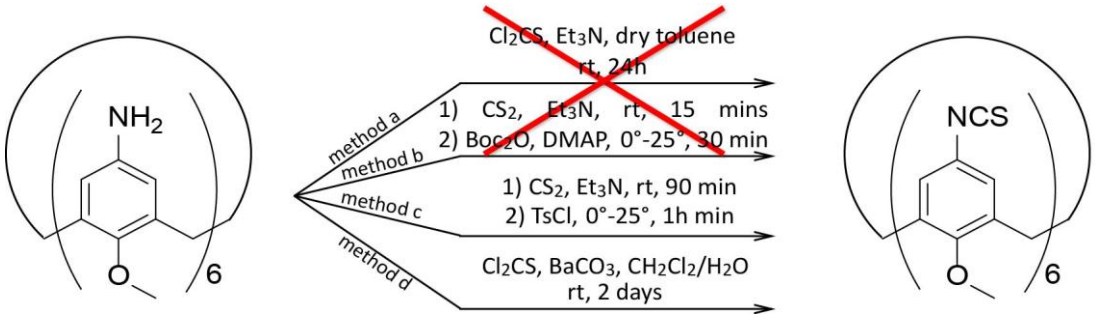
frequency resonance of the aromatic protons in the ^{1}H-NMR spectra confirmed the goodness of the reaction.

Actually, in the case of the tetramethoxy calix[4]arene **10a**, it is also interesting to mention that in $CDCl_3$ solution it is present in two conformations, cone and partial cone (observed even in the nitro calixarene **8a**) in 1:4 ratio, in slow exchange respect to the NMR time scale. (see appendix figure **1A**) To confirm this behaviour, spectra were acquired increasing the temperature from 25 °C to 80 °C, changing the solvent to DMSO, and, at higher temperature, it was observed that signals became broader, while not reaching the coalescence. The provided energy to the system allowed the rotational movements to be faster and so the situation found is an intermediate between the two previous conformations.

Going to the high-valency derivatives, unfortunately isothiocyanate calix[6]- and calix[8]arene **10d,e** were not obtained using the above mentioned reaction conditions. No evidence of the presence of compounds **10d,e** in the respective reaction crudes were found both by ^{1}H-NMR and ESI-MS analysis. There were only complex mixtures of compounds not possible to identify. Even doing a trial on the crudes with propylamine, similar to the sugar linker, to verify a possible reactivity no formation of the desired products of conjugation was achieved.

Since these derivatives with higher valency were not obtained using the procedure with thiophosgene, other approaches[58] were tried using the hexa-amino calix[6]arene **9d**, while we decided to leave the development of the octamer. The hexavalent macrocycle was chosen between the two because its functionalization resulted for us more consolidated respect to its octavalent homologous and it could furnish in any case important elements on the role of conformational mobility and valency increase and variation. Altogether four alternative synthetic strategies[59] , reported in **Scheme 3,** were explored and only methods c and d gave the desired hexa-isothiocyanate product **10d**. However, the biphasic reaction in CH_2Cl_2/H_2O, in the presence of thiophosgene and $CaCO_3$ as base (method d), gave derivative **10d** in higher yield compared to method c. The identity of the product was confirmed by NMR spectroscopy and ESI-MS analysis and it was fully characterized because synthesized for the first time. Particularly, the presence of the isothiocyanate groups at the upper rim was confirmed by the carbon signal of the -NCS group at 133.9 ppm in the ^{13}C NMR spectrum.

All the compounds **10a-d** were used as scaffolds for the synthesis of final glycocalixarenes **1-4** as described in paragraph 3.2.3.



Scheme 3: different synthetic approaches tried to obtain calixarene **10d**

3.2.2 Synthesis of the saccharide unit

The synthesis of the β-N-acetyl-D-mannosamine with a proper spacer (Figure **7**) required long efforts and was optimized trying different methodologies.

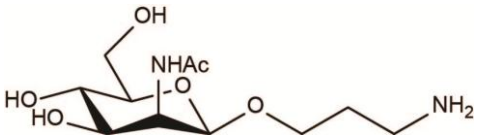


Fig.7: Aminopropyl-2-acetamido-2-deoxy-β-D-mannopyranoside

Three synthetic strategies were tried to obtain compound **19** and only the last one gave the desired carbohydrate. This was due to the general difficulty in the synthesis of β-mannosides: in fact mannopyranosyl and N-acetyl-mannopyranosyl derivatives are not easily accessible[60] and only by multistep synthesis, because the introduction of a group on the anomeric position, mainly or totally, gives rise to the α anomer. This is correlated to both steric and anomeric effect.[61] For that reason all the attempted synthetic pathways started from glucose for which easily it is possible the functionalization of the anomeric position to obtain the β anomer in good yields and after invert the configuration of C-2 (correction of stereochemistry) to afford the mannoside with the desired β linkage.

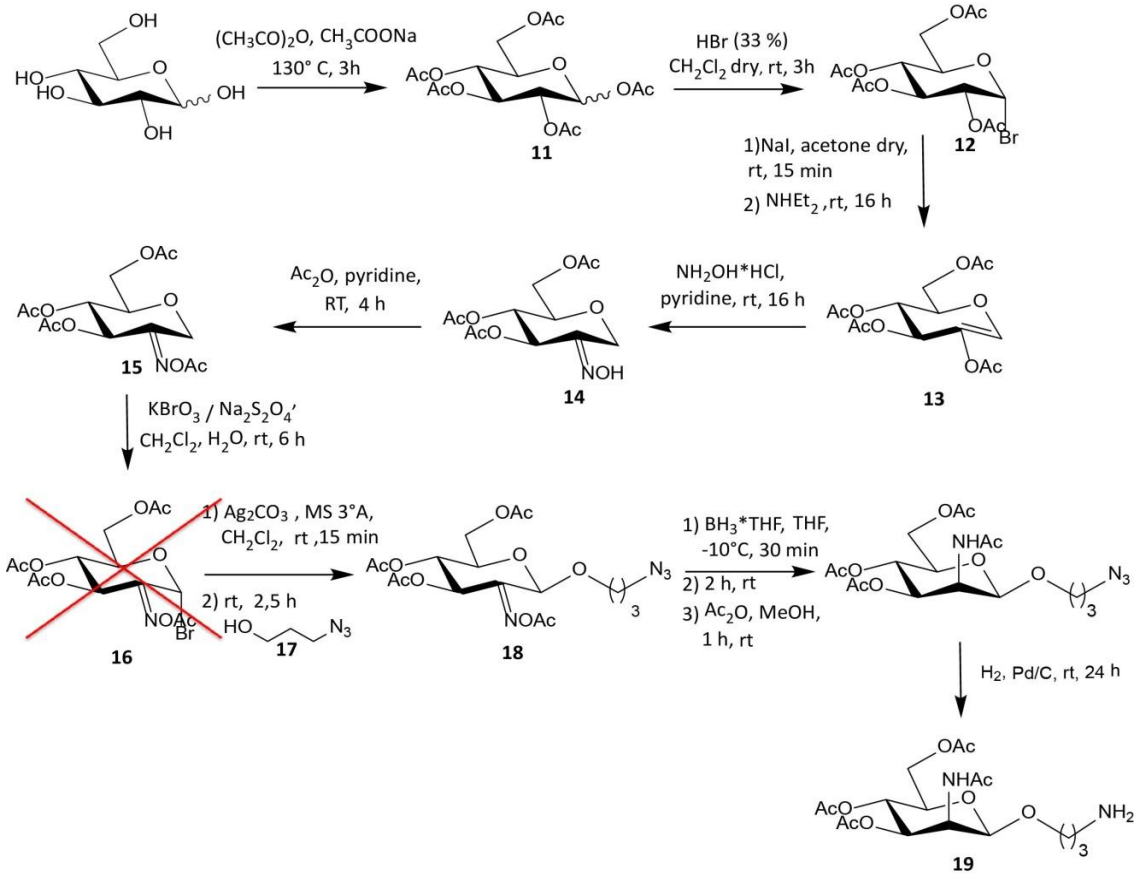
Another parameter to take in account for the preparation of final glyco-derivatives is the choice of the proper linker. In fact the distance between the ligating units and the

macrocycle is often considerably important and needs to be correctly designed prior to their synthesis to allow the simultaneous formation of all the possible supramolecular contacts.

It was also demonstrated by Fan and coworkers[62] that it is fundamental to determine the optimal linker length in a synthetic multivalent architecture because it is strictly connected to its potency. If the linkers are too long, during complexation there would be a loss of conformational freedom with a subsequent unfavourable entropic contribution to the binding event. On the other hand linkers too short make impossible for the glycoside groups of the multivalent ligand to simultaneously reach the binding sites of the antibody. In any case, both situations would lead to a lower binding affinity between antigen and receptor.

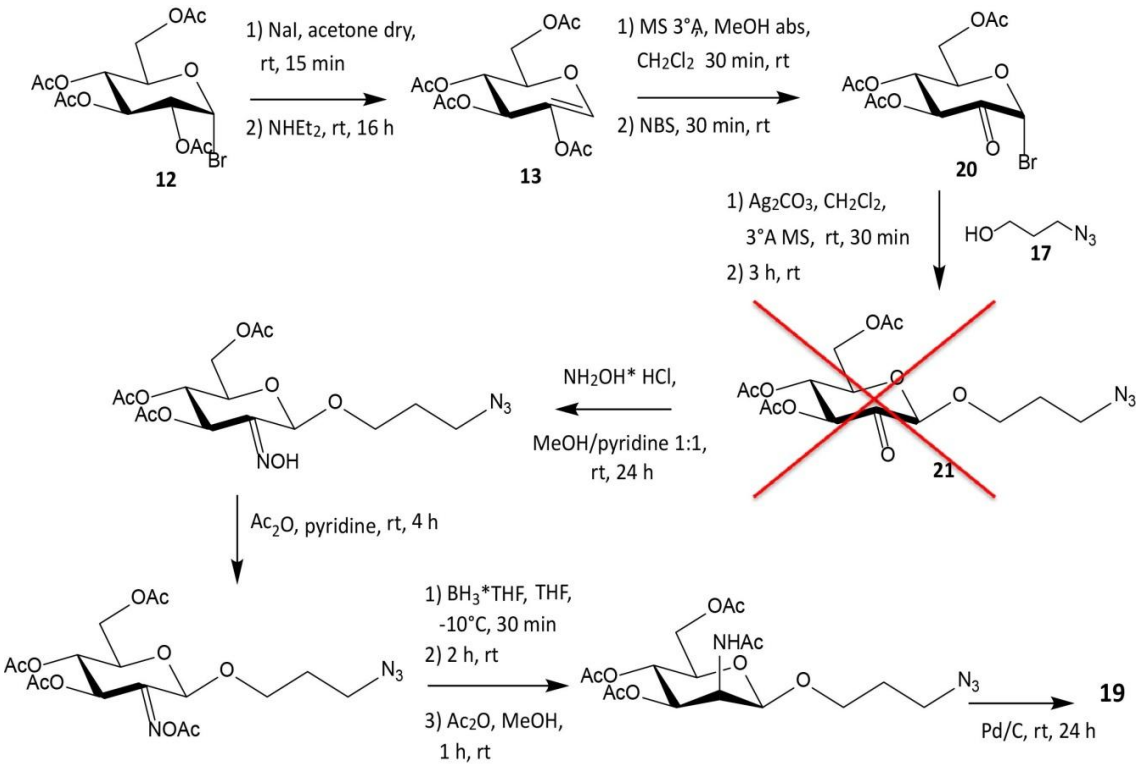
Actually, in our case there was not a specific knowledge about the potentially more suitable distance among the different saccharidic ligating units. However we decided to distance the sugar units from the macrocyclic scaffold to prevent crowding and ensure a certain degree of freedom to the sugars to reach the binding sites with hopefully beneficial effects on the recognition process. For these reasons the choice was for a propyl spacer.

The first tried synthetic pathway is reported in **Scheme 4**.



Scheme 4: 1st explored synthetic strategy to synthesize compound **19**

Fully deprotected D-glucose was then per-acetylated using a standard procedure[63] to prepare the pentaacetylglucose **11** in moderate yield (65%) as a mixture of α and β anomers. After that, the acetyl group in the anomeric position was replaced with bromide to give compound **12** as only α anomer for anomeric effect. **12** was transformed[64] using NaI and $NHEt_2$ in the 2-hydroxylglycal ester **13.** This is a good intermediate for the synthesis of the oxime **14**,[65,66] because can easily undergo hydroxylaminolysis of its enol ester group. The 2-oximino analogues from D-hexopyranosides have indeed figured prominently as key intermediates for the preparation of β-N-acetyl-mannosamine units.[67] The resulting 1-anhydro-D-glucose oxime **14** with further acetylation, using acetic anhydride in pyridine, gave compound **15**. The further reaction was the bromination of the anomeric position to obtain compound **16.** Normally the insertion of bromide is done by photobromination, but using acetylated saccharides it leads the formation of by-products such as bromoacetyl and dibromoacetyl derivatives, and even for example displacement of anomeric ester group.[64,68] An alternative procedure was found in literature[69] that exploits radical-mediated bromination using $KBrO_3$-$Na_2S_2O_4$ in CH_2Cl_2/H_2O. Unfortunately this method gave in our hands very different results÷ since the product was obtained in very low yield or not obtained at all. However a small amount of oximinoglycosyl bromide **16** was collected to go ahead with the glycosydation step with the three carbon atom linker **17**. Nonetheless, due to the difficulties found with the synthesis of **16** and the very low yield obtained in the preparation of **18,** for all these problems this route was left and to obtain compound **19** another one was tried (see **Scheme 5** below).

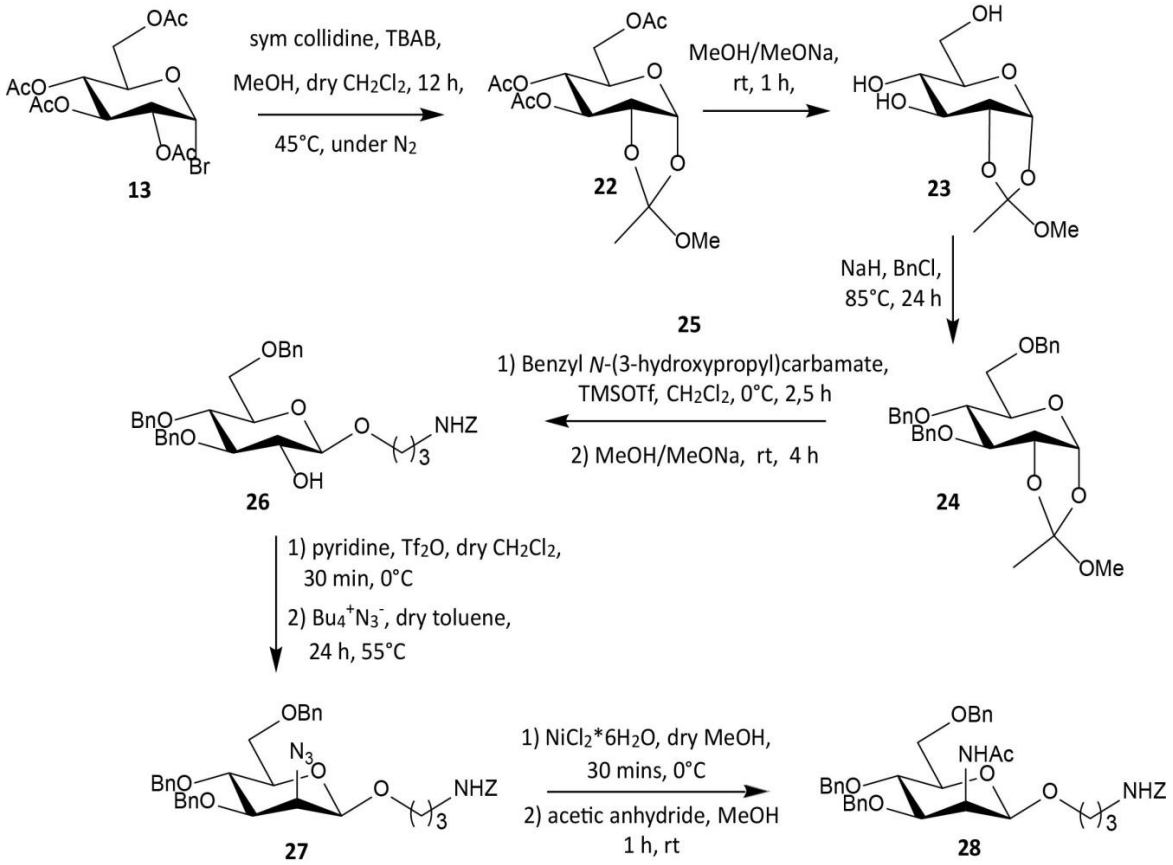


Scheme 5: 2nd explored synthetic strategy to synthesize compound **19**

Also in this second strategy (**Scheme 5**) the synthesis started from the α-acetobromoglucose **12** that was transformed in the glycal **13** using the same reaction condition. After that, the formation of acetylated 2-oxoglycosyl(ulosyl) bromide[70,71,72] **20** from hydroxyglycal ester was carried out in a single step. It comprises the exposure of **13**, in CH_2Cl_2 to N-bromosuccinimide (NBS) in the presence of methanol. The initial axial attack of the bromonium ion attempts the formation a 2-bromoacetoxonium salt intermediate in which the 2-O- acyl group is captured by methanol. The resulting formation of methyl acetate and a bromide ion led the formation of a ion pair between the stabilized carbocation on the sugar and Br^-. Their combination gave rise to ulosyl bromide **20**. We got a yield a little bit lower compared with the literature (65% against 70-90%). Derivative **20** was the glycosyl donor for the reaction with the linker **17** for the construction of the β-linkage via β-selective glycosidation, promoted by insoluble a silver catalyst. [67]

Unfortunately the isolation of the glycosidation product was not possible because when the crude was purified by flash chromatography the acidic character of the silica gel induced partial elimination in **20** producing a compound that, on the basis of literature data,[67] was identified as the respective enolone ester.

Due to this purification problem, this synthetic pathway too was discarded and a third alternative procedure was therefore designed on the basis of procedures reported in literature.[73] The synthesis of the spacer-bearing mannopyranoside **19** was achieved by the divergent approach illustrated in **Scheme 6**.



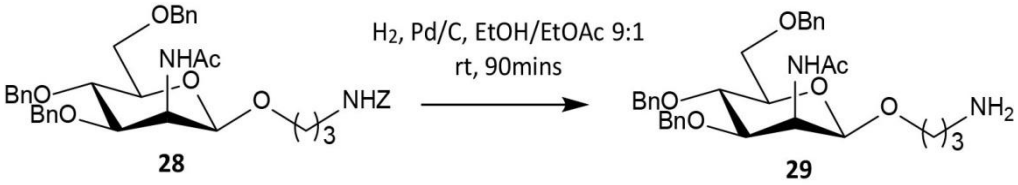
Scheme 6: 3rd synthetic strategy to synthesize compound **19**

The first step was the preparation of acetylated orthoestere **22** from α-acetobromoglucose **13**, using sym-collidine and tetra butylammonium bromide.[71] Subsequent exchange of acetyl with benzyl protecting groups was done firstly removing acetyl groups by Zemplen method, to obtain fully deprotected orthoestere **23.** Consequently the addition of benzyl chloride with NaH as base allowed the formation of benzylated orthoester **24**. Derivative **26** was synthesized *via* TMSOTf-catalysed opening of orthoester **24** with the previously prepared benzyl N-(3-hydroxypropyl)carbamate **25**, followed by 2-*O* deacetylation giving **26** as β-anomer with 62% overall yield. The configuration of the anomeric position was confirmed by the ^{1}H-NMR spectrum, where the $J_{1\text{-}2}$ coupling constant resulted of 7.2 Hz, due to the axial-axial coupling. Afterwards the 2-OH group was activated as triflate with

trifluoromethanesulfonic anhydride in DCM, followed by nucleophilic displacement with tetrabutylammonium azide and consequent inversion of configuration on C2, that provided β-2-azidomannopyranoside **27** in 40% yield. Reduction of the azide into the amine was done under Staudinger conditions (PPh_3 in THF, rt, 12h followed by addition of H_2O, 20h) and subsequent acetylation of the amino group provided compound **28** in 50% yield.

Nevertheless a new protocol[74] was later used for azide reduction using $NaBH_4$ and $NiCl_2$*$6H_2O$ in methanol and the reaction occurred smoothly increasing the yield up to 90% with a cleaner crude. The N-acetylation was carried out using acetic anhydride in pyridine and the overall yield was raised up to 80%.

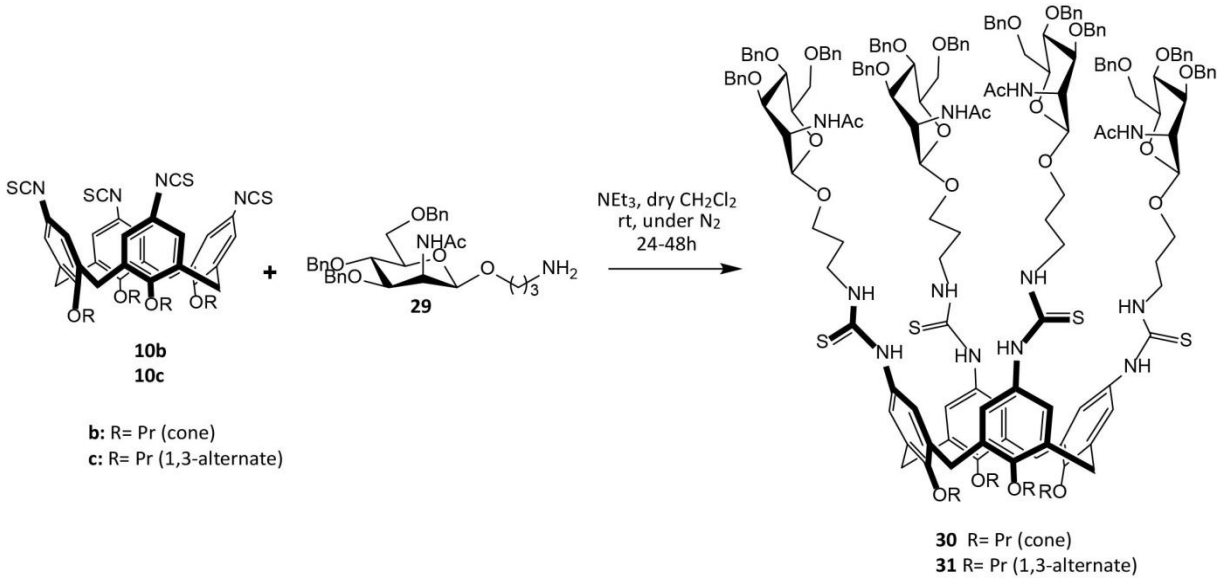
To conjugate the benzylated sugar **28** to the calixarene scaffolds a selective removal of the benzyloxycarbonyl (Cbz) group was necessary, to have available the amino group for the coupling reaction.[75] In fact, to remove O-benzyl-protected groups using hydrogenolysis the reaction times are longer. Instead it was observed that ester groups undergo hydrogenolysis with greater facility than ether groups.[76] In our case there was a carbamate to protect the amine, but this method was tried the same, since in literature no evidences of degradation can occur with these conditions for this functionality. In this way the reaction time was decreased to 90 min respect to 24-48 h required for benzyl groups and the desired product **29** was successfully prepared under these conditions (**Scheme 7**).



Scheme 7: selective removal of Cbz protecting group in **28**

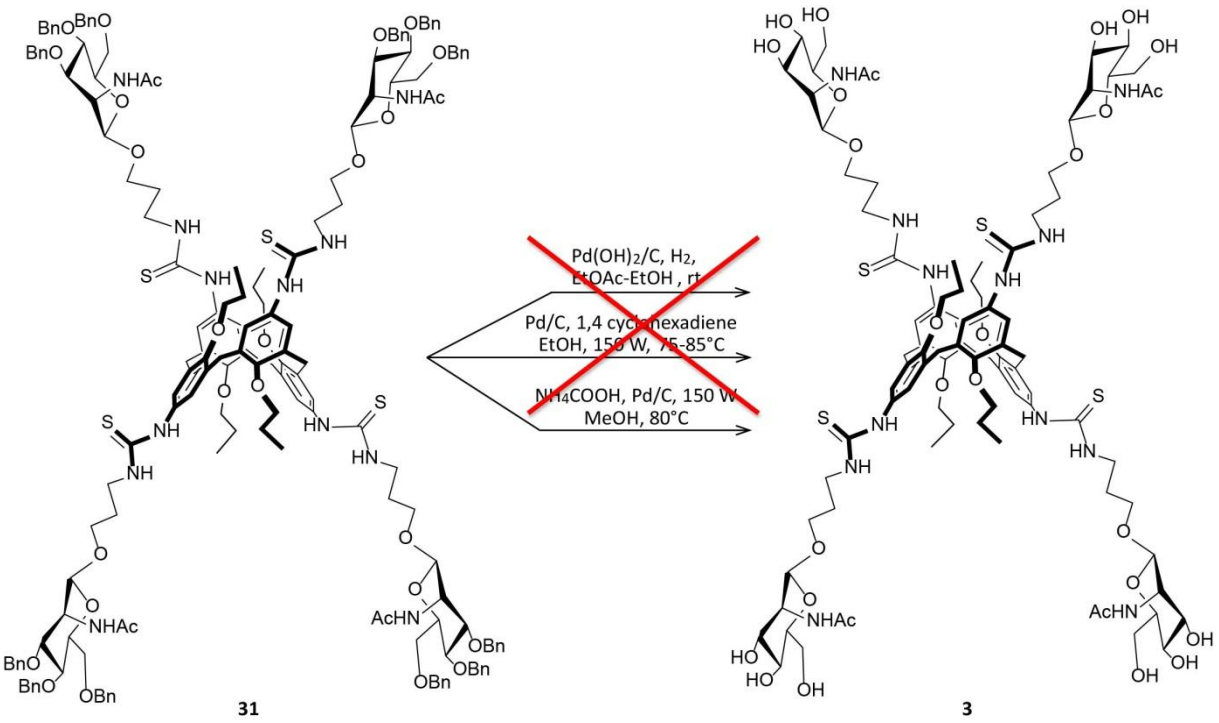
After that the sugar **29** was conjugated to the calix[4]arene scaffolds **10b** and **10c**, in cone and 1,3 alternate conformation, respectively. Coupling reaction (**Scheme 8**) between isothiocyanate calix[4]arenes and carbohydrate led to the formation of glycocalix[4]arenes **30** and **31** with yield of 91% and 50% respectively. The "click" reaction determined the formation of the thiourea spacer, with a purification necessary only for the removal of the sugar excess used to maximize the yields. The benzyl groups can be removed by hydrogenolysis but the reaction required long times (12 Bn groups totally) to afford

compounds **30-31** and so a different catalyst was used to enhance the reaction speed: Pd $(OH)_2/C$.[77]



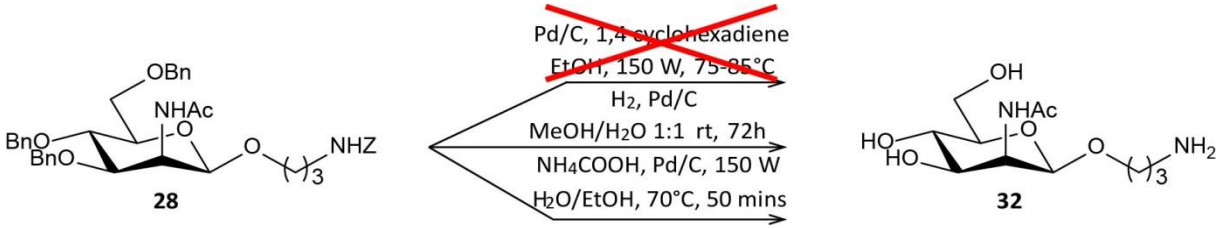
Scheme 8: conjugation reaction using "click" chemistry, between benzylated amino sugar and isothiocyanate calix[4]arenes.

Unfortunately this approach, (**Scheme 9**) tried on the alternate derivatives **31**, did not give the desired product **3**, since degradation of starting material was observed and so an alternative debenzylation method using 1,4-cyclohexadiene as hydrogen source was attempted.[78] The reaction was performed under microwave with the hope that benzyl group removal could take advantage from this technology, enabling access to high temperatures and pressures . Nevertheless, this method too did not give the desired product and so a third alternative way to obtain compound **3** was explored using again the microwave assistance but in the presence of 10% of palladium on charcoal and ammonium formate[79] as hydrogen donor catalytic transfer hydrogenation.[80] Also in this case the tetra functionalized calixarene **3** was not achieved, since even after 2 days the starting calixarene **31** was the only compounds already present. Even partially deprotected intermediates were not found.



Scheme 9: Synthetic scheme of the different methods tried to remove benzyl groups from calixarene **3**

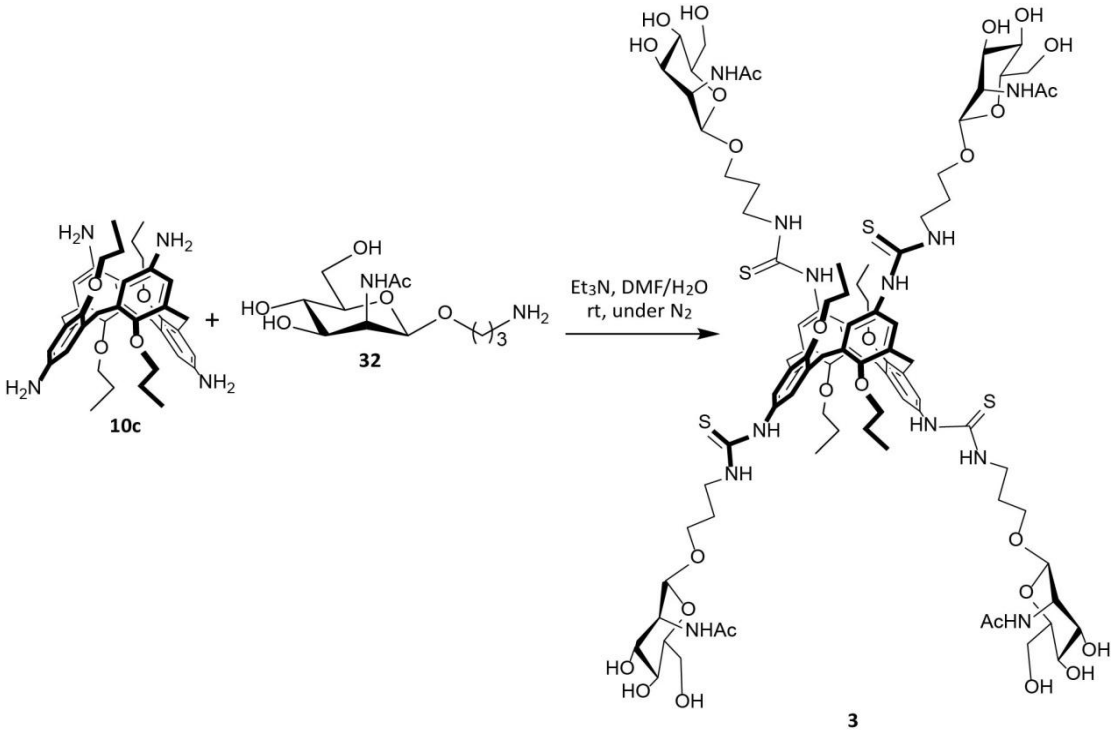
After all these failures, it was planned to go directly to the deprotected clusters through conjugation of the deprotected mannosamine derivative to the calixarene isothiocyanates. Then the removal of the benzyl groups from sugar **28** was attempted to obtain **32.**



Scheme 10: different methodologies for the debenzylation of compound **32**

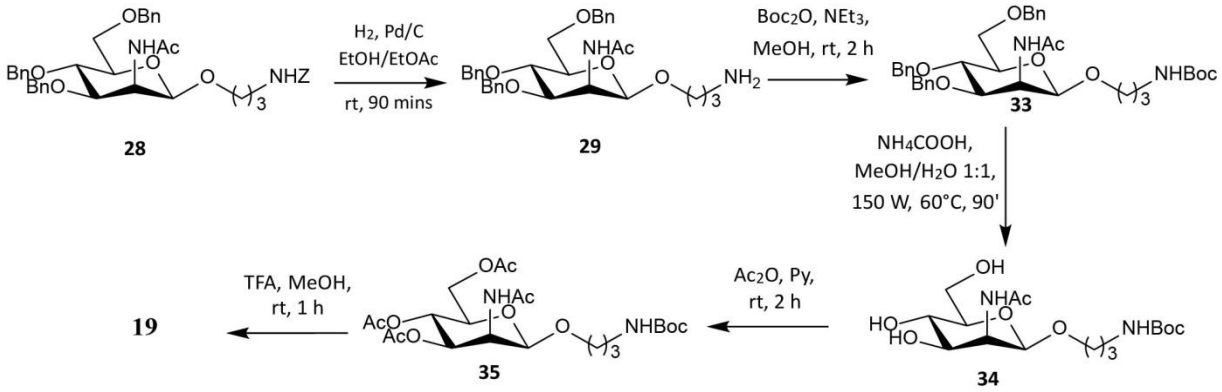
Also in this case several methodologies were tried to arrive to **32** as reported in **Scheme 10**. Removal of benzyl groups using hydrogenolysis[81] with H_2 in presence of Pd (10% on C) gave rise to the fully deprotected sugar in quantitative yield, but the time required to complete the reaction resulted very long (72h). To save time other two procedures were tested, using again MW assisted catalytic transfer hydrogenation in the presence of 1,4-cyclohexadiene or ammonium formate as hydrogen donors. With the first one the obtained crude was constituted by a mixture containing the fully deprotected compound together with a series

of partially deprotected derivatives. On the contrary, employing ammonium formate the desired compound was obtained even if the purification from the salt resulted very difficult and not complete. However, the so reached sugar **32** was used as such for the connection to the calixarene scaffold. A test reaction on the 1,3 alternate isomer **10c** was tried (**Scheme 11**) but only a mixture of compounds with all the different degree of functionalization, mono-, di-, tri and tetra glycosylation, was isolated. Even if the tetra glycocalix[4]arene **3** was obtained, the low yield and the complexity in the purification step suggested to discard also this method looking for a further approach.



Scheme 11: conjugation reaction using fully deprotected sugar **32**

We thus decided to go back to the idea of conjugating the saccharide when still protected also because this could facilitate the purification procedures in the case the desired fully glycosylated calixarenes were obtained in a mixture with partially functionalized intermediates. Already clarified, as described above, that the glycocalixarenes with benzyl protecting groups could not be deprotected, we were forced to plan a longer synthetic pathway including the restore of the acetyl groups on the sugar (**Scheme 12**). It had been verified in the past in fact that the removal of this protecting group from glycocalixarenes was an easy task. A sequence of protection and deprotection reactions starting from **28** was carried out to finally gave rise to compound **19** with a overall yield of 50%.

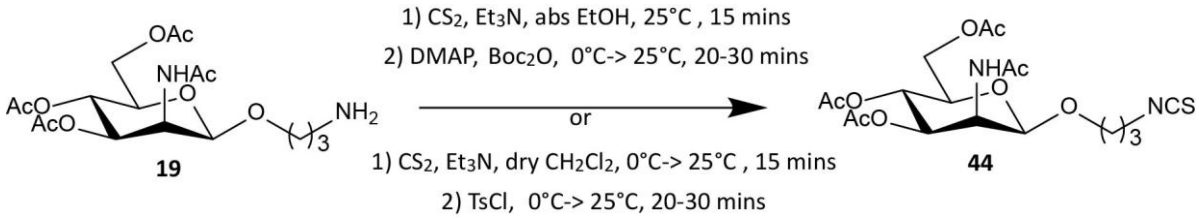


Scheme 12: synthesis of compound **19** by a series of deprotection/protection steps

Going a little more in detail, after the selective removal of N-Cbz from **28** by hydrogenolysis under H_2 atmosphere catalysed by Pd/C protection of the amino group on the three carbon atom chain in **29**, with Boc group orthogonal to benzyl moieties, allowed the formation of compound **33**, that underwent a catalytic hydrogen transfer under MW to obtain **34**, fully deprotected on the -OH groups, without any purification. The excess of NH_4COOH used in the benzyl removal and not completely eliminated after work up, was not a problem, since it did not interfere with the subsequent step. The acetylation in acetic anhydride and pyridine afforded sugar **35** in high yield with a easy purification on silica gel column to eliminate all the impurities present from the previous reactions. This purification was considered necessary since we preferred to be in the condition of obtaining **19** in a highly pure form to be used as such for the conjugation reaction avoiding undesired lost of material of degradation. The last step of this long pathway was then the elimination of Boc protecting group to reach compound **19** in a quantitative yield. On the whole, from glucose, this is without doubts the longest route (15 steps) but finally the desired saccharidic unit **19** was obtained with an 8% overall yield.

Since the synthesis of the hexa-isothiocyanatecalix[6]arene had resulted initially unsuccessfully, as reported in the previous section, to obtain the hexavalent cluster it was also planned to insert the isothiocyanate function on the sugar and use the amino derivative of the macrocycle. To this purpose several methodologies were tried to prepare derivative **44**. Firstly the use of thiophosgene and $CaCO_3$ (not reported in **Scheme 13**) allowed the

formation of only two by-products: the corresponding carbamic acid and the disaccharide with a thiourea linker, as seen by NMR and ESI-MS analysis. This last derivative was reasonably originated by the reaction between the desired product **44** that was forming and the amino sugar **19** still present in the reaction mixture. This procedure was left and isothiocyanation using carbon disulfide and di-*tert*-butyl dicarbonate (Boc_2O)[58] or tosyl chloride (TsCl)[82] as desulfurylation reagents was followed (**Scheme 13**). The mechanism suggested[59] for this reaction involves the formation of a dithiocarbamate, stabilized by the addition of Et_3N. Then Boc_2O (or TsCl) was added to react and form an unstable "mixed ditiocarbamate/carbonate" adduct, then the presence of DMAP quickly decomposes this intermediate to form the product. In both cases the reaction gave **44** but in very low yield (15-18%) and so this strategy was not so convenient. Fortunately, in the meantime, isothiocyanate calix[6]arene **10d** was achieved and then the problem was overcome.



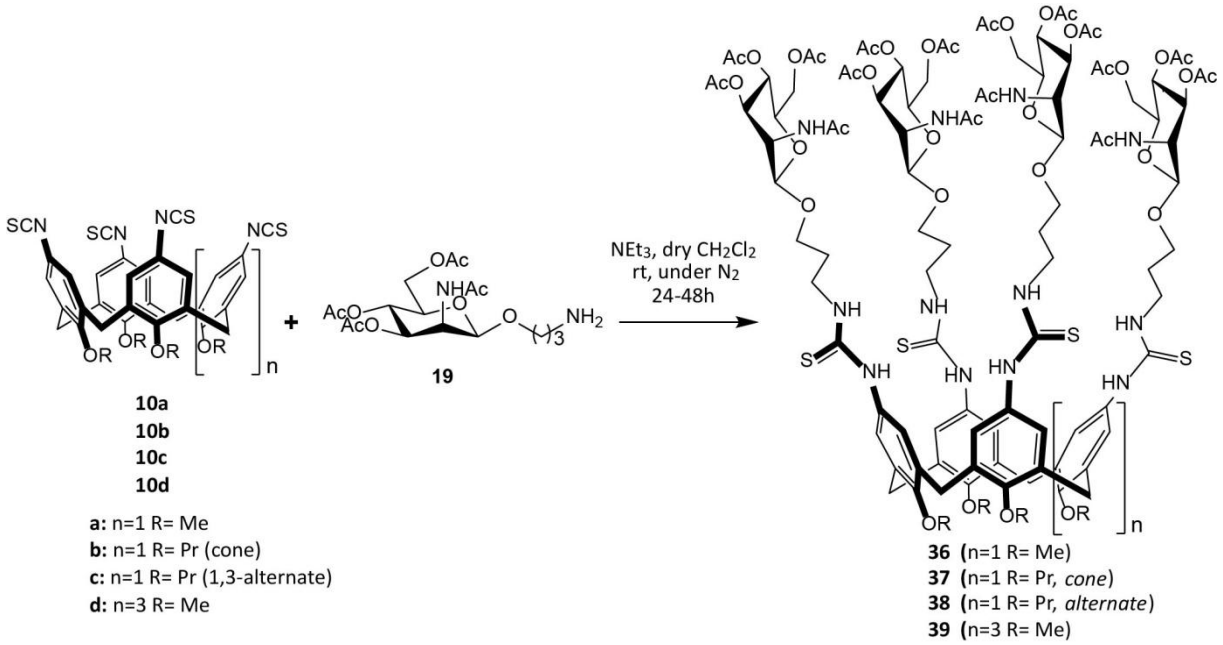
Scheme 13: synthesis of the isothiocyanate sugar **44**

3.2.3 Synthesis of glycocalixarenes 1-4

At the end the calix[n]arene scaffolds **10a-d** and the sugar **19** could undergo the coupling reaction, using the "click" method between the isothiocyanate group on the upper rim of the macrocycle, obtained by isothiocyanation of **9a-d** with thiophosgene, and the amino group at the extremity of the chain in the saccharide.

The reaction was carried out for all the calixarenes in dry DCM and with Et_3N as base used to be sure of the complete deprotonation of the amino groups (**Scheme 14**). Four different glycocalix[n]arene were synthesized: 4N-ManAc[4]Me **36,** *cone*-4N-ManAc[4]Pr **37,** *alt*-4N-ManAc[4]Pr **38**, and 6N-ManAc[6]Me **39**. Despite its belonging to the click chemistry class, this conjugation reaction gave our compounds in moderate yields (25-71 %) and the purification via flash chromatography was necessary for the removal of sugar excess and in some cases of partially functionalized intermediates. Identity and purity of the desired

clusters were confirmed by ESI-MS analysis and NMR spectroscopy and they were fully characterized, since all are new compounds, not already present in literature.



Scheme 14: synthesis of the thiourea-bridged glycocalix[n]arenes **36-39**

Two particular behaviours observed by NMR spectroscopy are noteworthy and are discussed here below. The first one concerns the fully acetylated conformationally mobile 4N-ManAc[4]Me **36**, whose ^{1}H NMR spectra are reported in figure **7**.

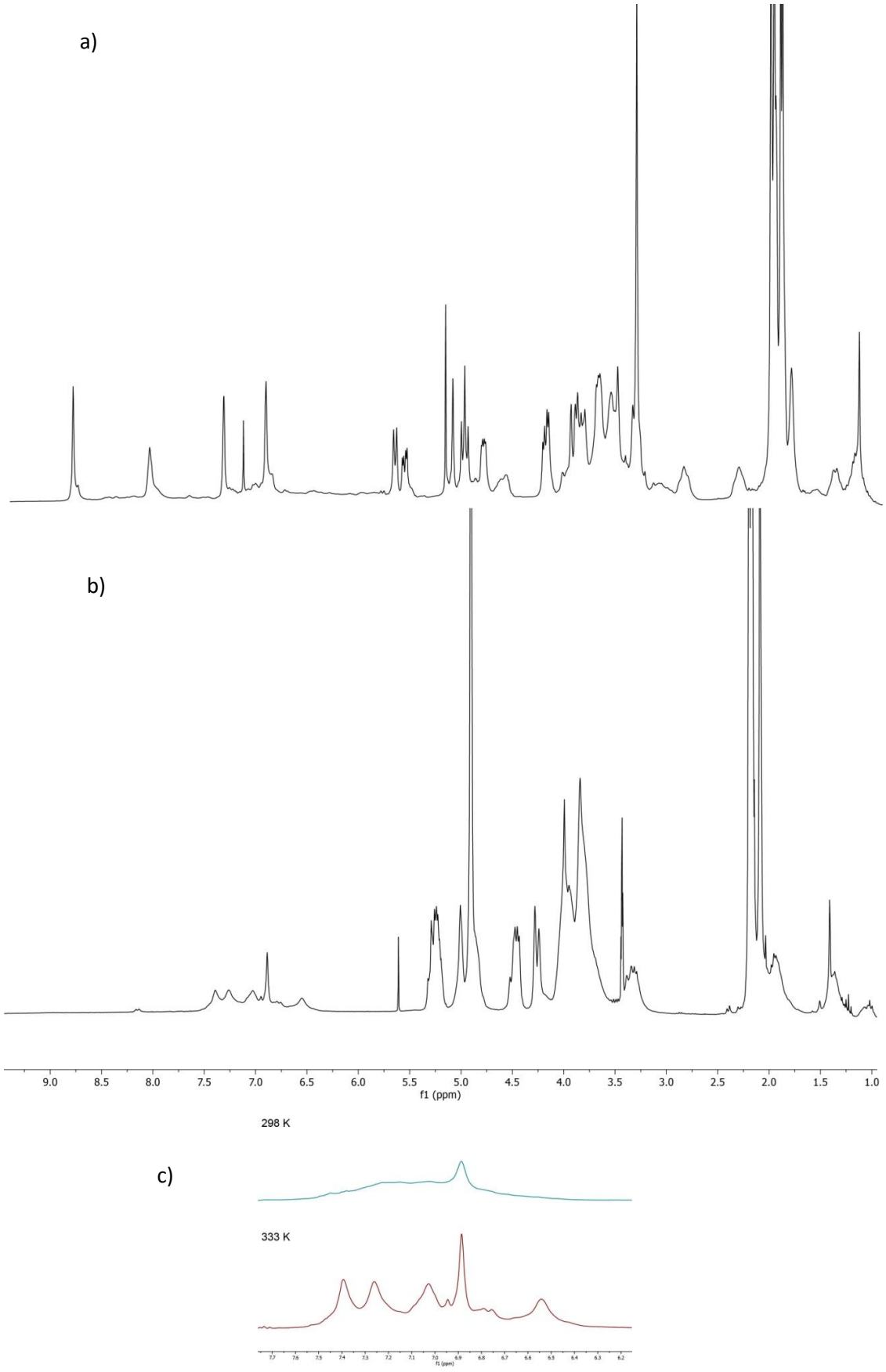


Fig. 7: ^{1}H NMR spectra (400 MHz, temperature) of compound **36** a) in $CDCl_3$ and b) in MeOD. In c) the zoom relative to the aromatic region for a better comparison.

In the spectrum in $CDCl_3$ (Fig. **7a**) it is possible to recognize at 8.91 and 8.16 ppm the proton signals of N*H*CS from thiourea spacer, that confirmed the occurred reaction. However, the peculiarity in the behaviour of this compound is evidenced by the two signals relative to the aromatic protons at 7.42 and 7.04 ppm. From the COSY experiment we could verify that they are faintly correlated even if they appeared as two singlets. This means that they belong to the same aromatic ring with, actually, a *meta* coupling constant. The first conclusion was that the glycocluster could be mainly in a *cone* conformation, accordingly to the spectrum of **37** (see appendix, figure **2A**), the isomer indeed blocked in the cone geometry. On the other hand, it was missed in the spectrum a clear presence of the two doublets relative to the methylene bridges of the macrocycle typical of the cone geometry. Moreover, some signals attributable to the aminopropylic chain at 2.93 and 2.39 appeared more shielded compared with their resonances in the spectrum of the sugar not yet conjugated to the macrocycle. This could be due to the 1,3-*alternate* or partially *alternate* conformation, because the chain protons of a sugar units at the upper rim could be affected by the ring current of one or two proximal aromatic units in the reversed orientation. An hypothesis to explain the strong difference in chemical shift between the two protons of the same aromatic unit was that in $CDCl_3$ the occurrence of hydrogen bonding, involving for example the thiourea units, blocks the substituents in a specific conformation, determining a marked differentiation in their chemical environment. A proof of this was collected acquiring a spectrum of the same compound in a polar solvent, as MeOD (Figure **7b**), to break hydrogen bonds. The situation observed in this medium, in fact, changed substantially, resulting similar to that of calixarene **10a** present in solution in two conformations showing on the whole five signals for the aromatic protons, one due to the *cone* and four to the *partial cone* conformation. In particular the aromatic hydrogens of the cone conformer gave in these conditions of higher polarity only a signal supporting the idea that their splitting in $CDCl_3$ was due to hydrogen bonding. Increasing the temperature of the sample in CD_3OD up to 60 °C, the aromatic signals resulted broader and closer compared to the spectrum at room temperature, but they did not reach the coalescence.

The second situation discussed more in detail is relative to the derivative **38**, the *alt*-4N-ManAc[4]Pr, blocked in the 1,3-alternate geometry. Its ^1H NMR spectra in $CDCl_3$ resulted particular and with unexpected elements that, however, once explained contributed to

better clarify also the behaviour of the tetra-methoxycalixarene **36**. To this end, for example, it was possible to observe (Fig. **8a**) in the aromatic region that the situation was the same as that of derivative 4N-ManAc[4]Me **36**, with two peaks, at 7.38 and 7.05 ppm, each one corresponding to 4 protons, with a small J value, caused by a *meta* coupling. More importantly, the peaks relative to the aminopropylic chain are indeed upfield shifted as observed for **36**.

But the most unexpected situation is that observed for the protons of the methylene bridges. It is well known in fact that in a calix[4]arene in the 1,3-alternate conformation with the same substituents on the four aromatic units these protons usually give rise to a singlet.[47,83] However in this case two singlets at 3.98 and 3.64 ppm (red circle in **figure 8a**) both attributed to the protons of the methylene bridge were present. Also the ^{13}C-NMR spectrum pointed out the same feature since there were 2 signals, at 38.3 and 36.6 ppm, both corresponding to the carbon atom of the methylene bridges. These peculiarities suggested that the two hydrogens of each methylene are equivalent but the methylene bridges are equivalent only two by two. This situation can originate only from the substituents at the upper rim because of the presence of both chiral centres and donor/acceptor groups of hydrogen bonding. In particular the second reason resulted relevant since by recording a ^{1}H-NMR in MeOD, solvent able to break these interactions, the protons of the methylene bridges gave rise only to one singlet at around 3.7 ppm. In addition, in this solvent, the aromatic protons resulted more similar, with their signals closer and with chemical shift values characteristic of calixarenes functionalized with thiourea groups. At this step it resulted clear, for analogy, that the conformation kept by 4N-ManAc[4]Me **36** in $CDCl_3$ solution (figure **7a**) was the *1,3-alternate* one. Further studies via NOESY experiments and molecular modelling (see next paragraph) were subsequently carried out to understand and visualize the arrangement of the substituents in compound **38** that could justify the splitting of the signals relative to the methylene bridges mentioned above.

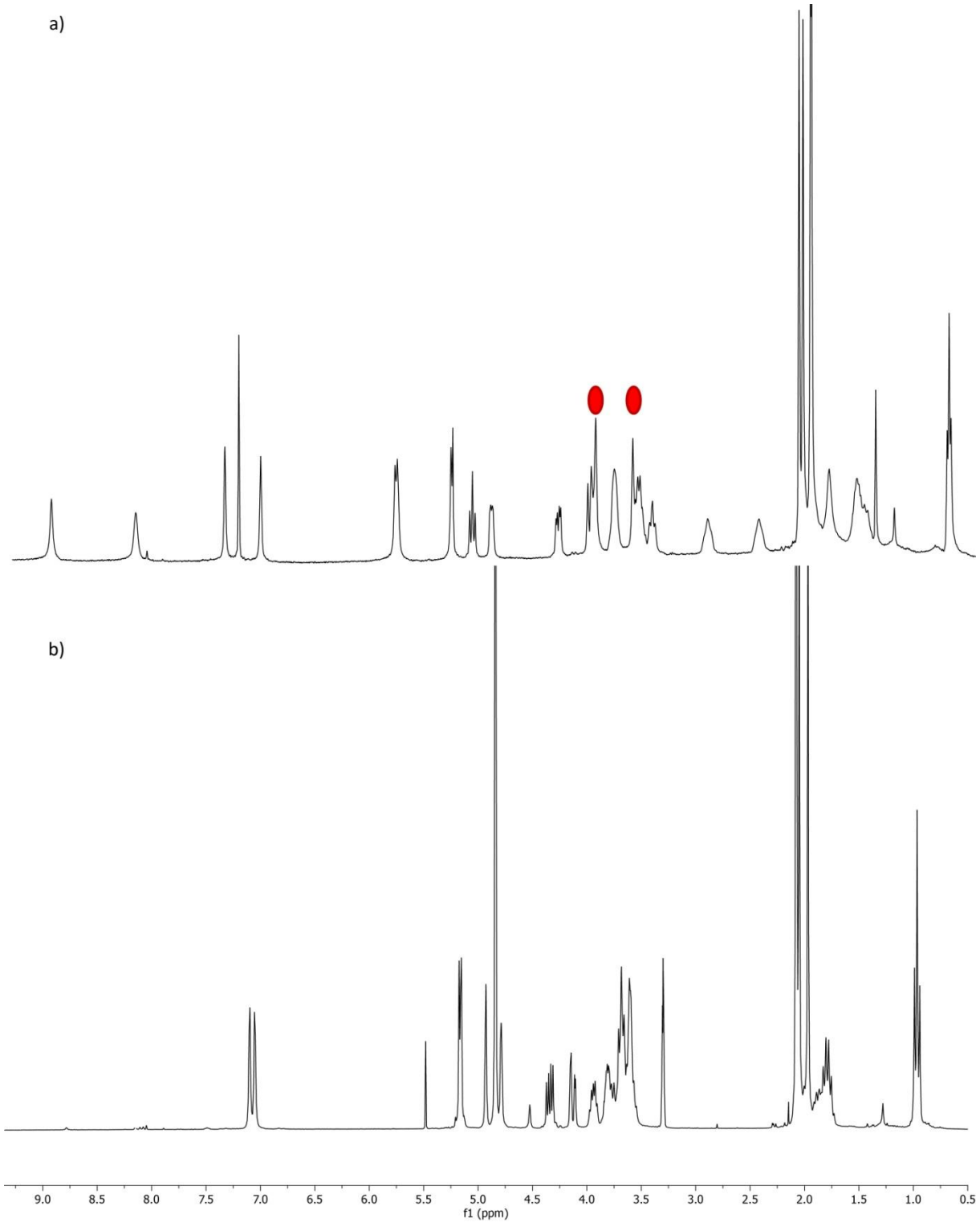
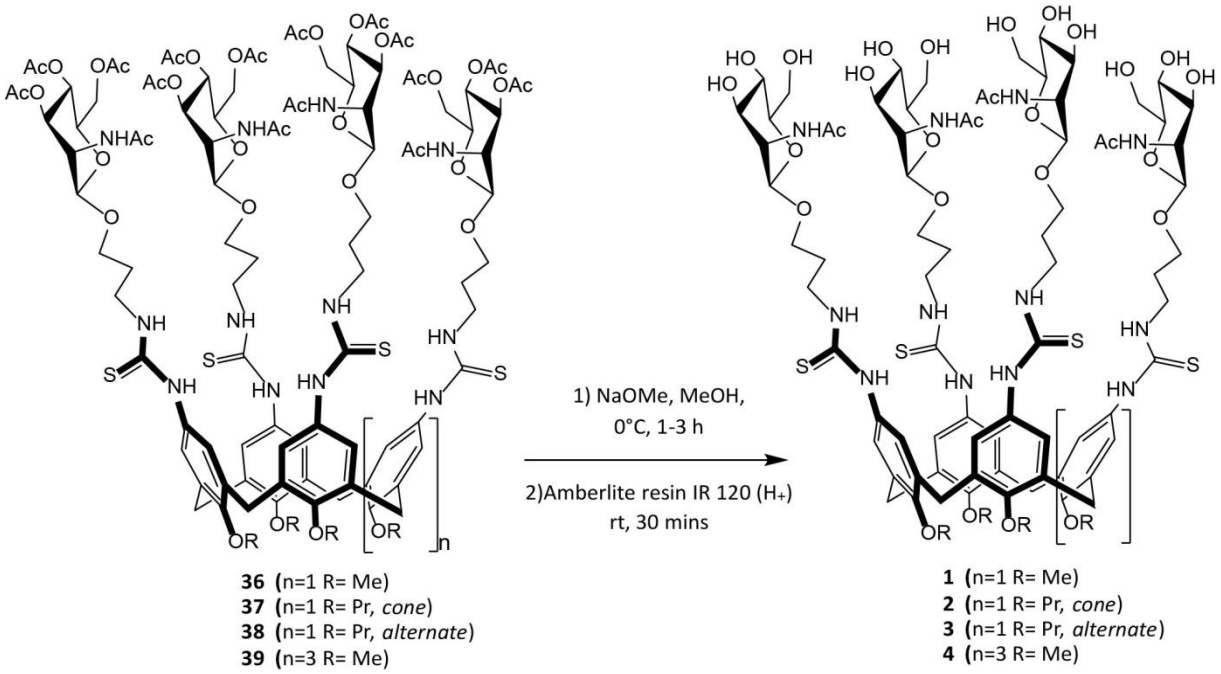


Fig. 8: ^{1}H NMR spectrum (400 MHz,25°C) of compound **38** a) in $CDCl_3$ and b) in MeOD

Eventually, in order to obtain the target compounds **1**-**4**, deacetylation reaction using Zemplen method[84] was performed on clusters **36**-**39**.



36 (n=1 R= Me)
37 (n=1 R= Pr, *cone*)
38 (n=1 R= Pr, *alternate*)
39 (n=3 R= Me)

1 (n=1 R= Me)
2 (n=1 R= Pr, *cone*)
3 (n=1 R= Pr, *alternate*)
4 (n=3 R= Me)

Scheme 15 : Deacetylation reaction to obtain compounds **1-4**

It was carried out at 0 °C, not room temperature, since in literature evidence was reported of epimerization on thioureidomannosides,[85] for long reaction times. In fact this anomerization process was not correlated to the MeONa concentration, but it was found strongly dependent on mainly reaction temperature and, secondarily, on reaction time. Actually in our case the thiourea moiety was not directly linked to the anomeric carbon as in the mannoside studied in that research, but we preferred to work really carefully to prevent any possibility of degradation and/or loss of product after the hard work done. The deprotection reaction, quenched by addition of Amberlite IR-120(H^+) resin, afforded target glycocalixarenes **1-4** (**Scheme 15**), without epimerization and with complete removal of acetyl groups in 1-3h, as determined by ESI-MS and ^{1}H-NMR. Not quantitative but high yields were reached, due to the efforts done to purify the compounds before sending them for biological tests.

All fully deprotected derivatives were characterized by NMR and mass analysis. 4N-ManAc[4]Me **1**, *cone*-4N-ManAc[4]Pr **2**, *alternate*-4N-ManAc[4]Pr **3** resulted well soluble in methanol, while 6N-ManAc[6]Me **4** did not show a complete solubility. Acquiring the **1** ^{1}H-NMR spectrum in MeOD the situation appeared quite confused. Therefore D_2O was used as solvent, since it is well known that conformationally mobile calix[4]arene amphiphiles in water usually tend to adopt a 1,3-alternate structure, which minimizes the lipophilic surfaces

exposed to the solvent. However at room temperature the spectrum did not clarify the conformational situation still showing broad signals. Increasing the temperature up to 80 °C a well-defined spectrum was obtained (Figure **9**). It was possible to identify the H_1 and H_2 signals respectively at 4.69 and 4.43 ppm and the methylene bridge signal demonstrated a 1,3-*alternate* conformation instead of the *cone* and *partial cone* found in the protected precursor **36**.

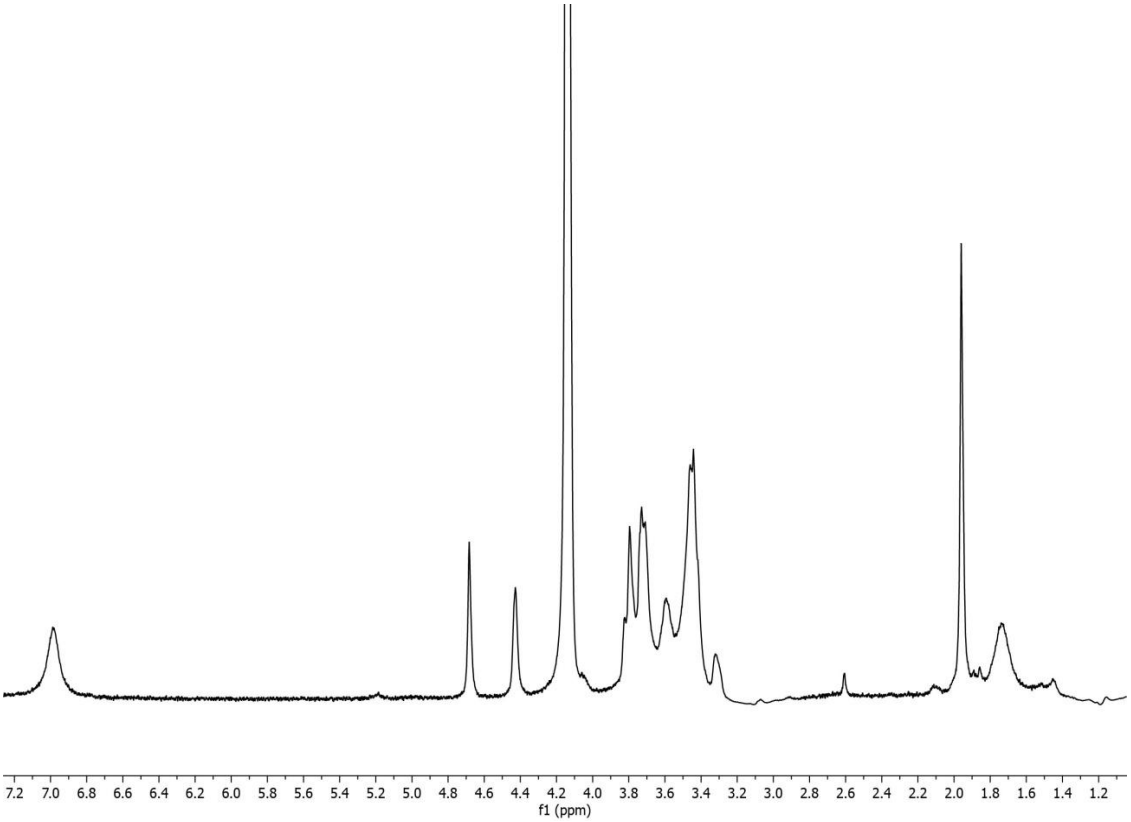


Fig. 9: NMR spectra (400 MHz, D_2O) of compound **1** at 80 °C

As observed for **1**, also 6N-ManAc[6]Me **4** gave in D_2O at room temperature ^1H NMR spectra (see in appendix) with broad signals. For both compounds, this could be due to a slowed mobility of the aromatic units through the macrocyclic annulus and of the substituents in general. On the other hand, in the past it was observed[86] that rather unexpectedly methoxy-glucocalix[6]- and [8]arenes tend to self-assemble in aqueous solution despite the absence of a well defined amphiphilicity as in both **1** and **4**. Then also for these two glycocalixarenes self-aggregation could occur determining or contributing to the broadening of the NMR signals. Increasing the temperature up to 60 °and 80 °C, the situation changed and the

signals started to become sharper. This could be a consequence of the disaggregation of the assembly, or to a poor and slowed mobility in solution of this derivative.

For the *cone*-4N-ManAc[4]Pr **2** a confirmation of the cone structure was given by the signals of the methylene bridge, as 2 duplets, each one counting for 4 protons, at 4.47 and 3.2 ppm while for the *alt*-4N-ManAc[4]Pr **3** a singlet, typical of this conformation, was observed in the region between 3.63-3.50 ppm (see figure **4A** in the appendix).

3.2.4 Computational studies and NOESY experiments for protected glycocalix[4]arenes 36 and 38

In order to explain more deeply the unexpected pattern of NMR signals for the 1,3-alternate macrocyclic derivative **38** , further studies were required. Particularly computational studies on peracetylated glycocalixarenes **36** and **38** were carried out to understand the structure taken in solution. These investigations were performed in the group of Prof. Lucio Toma at the Università degli studi di Pavia.

In figure **10**, the obtained 3D structure for compound **36** in the 1,3-alternate geometry is reported. The same results were obtained for **38**.

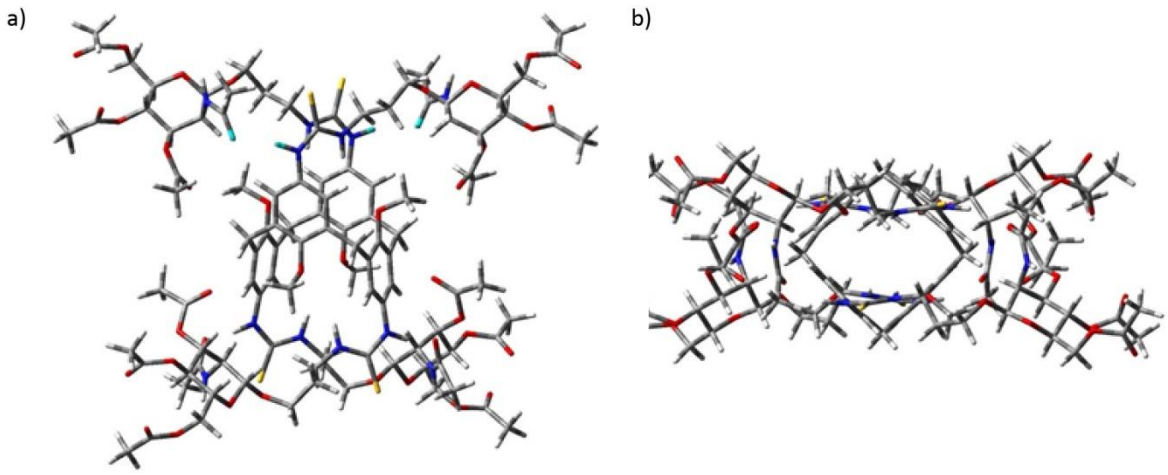


Fig. 10: 3D plots of derivative **36** in a) side view and b) top-view obtained thanks to Prof. Lucio Toma. The NH and CO groups involved in hydrogen bonding are evidenced in cyan.

The structure clearly evidences as there are two different types of methylene bridges and that, at the same time, the two hydrogens linked to the same carbon bridge are surrounded by the same neighbourhood resulting equivalent. This is fully in agreement with the two

singlets observed in the ^{1}H-NMR spectrum. The whole arrangement of the molecule appears determined by four hydrogen bonds, each one between the acetamidic carbonyl group of mannosamine in position 2 and the thioureido NH group linked to the aromatic ring, that stabilize all the structure. The distance between the two groups involved in hydrogen bonding, evidenced in cyan in the structure, (figure **10a**) is 2.04 Å.

Furthermore the specific sugar spatial orientation and the linker folding lead to a NOE distance some hydrogens pairs. For example in the found structure the distance between mannosamine H-2 and aromatic hydrogen in ortho position respect to the thiourea unit is around 2.8 Å, whereas the distance between sugar H-1 and the central CH_2 in the linker are at ≈ 2.5 Å.

In order to confirm the correlations observed in these computational studies, NOESY experiments were carried out on the 1,3-alternate peracetylated glycocalixarene **38**.

Three experiments were completed, changing the mixing time value (d8) from 0.2 to 0.6 ms but no significant variations were found in the three collected bidimensional spectra. Analysing the bidimensional map reported in figure **11**, obtained with d8 = 0.2 ms, we could in particular observe that each NH proton of thiourea correlates with only one of the two signals relative to the aromatic protons. This means that thiourea is in the EZ conformation and in fact calculations indicated a higher stability for this geometry respect to the others, confirmed by the lowest potential energy (10 kcal/mol less respect to the other rotamers).

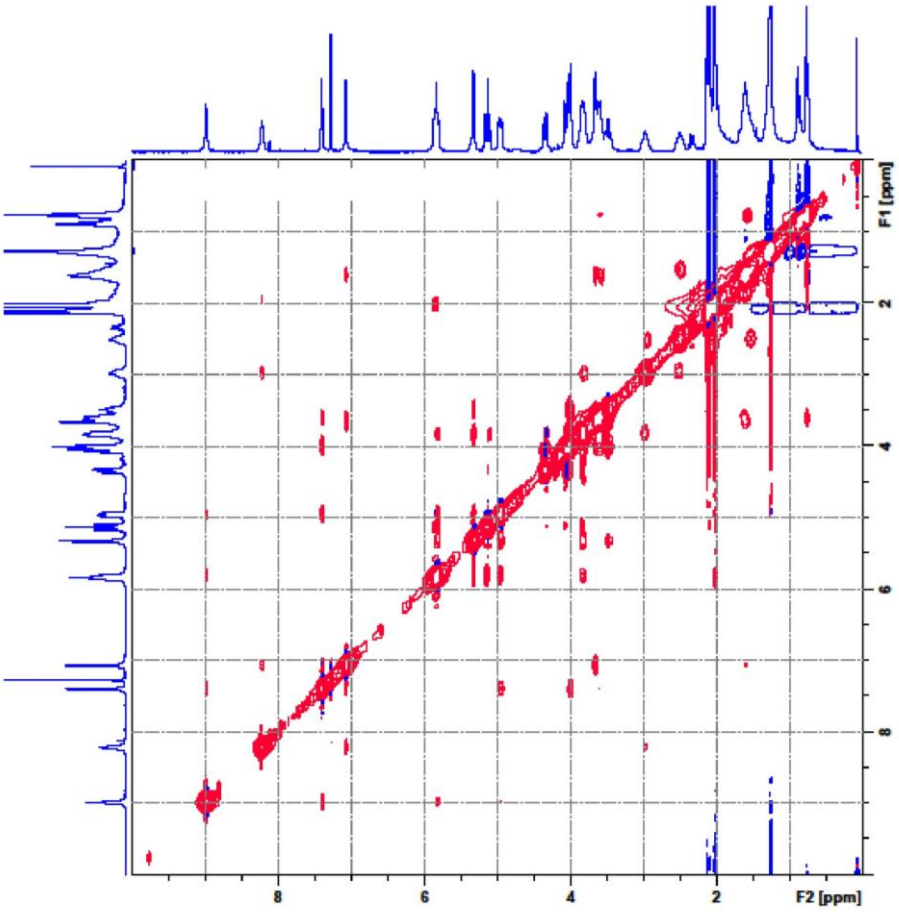


Fig. 11: NOESY spectrum on glycocalixarene **38** (300 MHz, $CDCl_3$)

Then we could find a correlation between mannosamine H-2 and the same aromatic proton already correlating with the hydrogen of the thioureido NH directly attached to the aromatic ring of the calixarene scaffold. This is a key point to confirm what found with computational studies, where the distance between H-2 and $H_{aromatic}$ was around 2.8 Å, due to a retreat of the propylic spacer. It is also interesting to underline that the sugar H-2 correlates with the aromatic proton of the opposite phenyl ring.

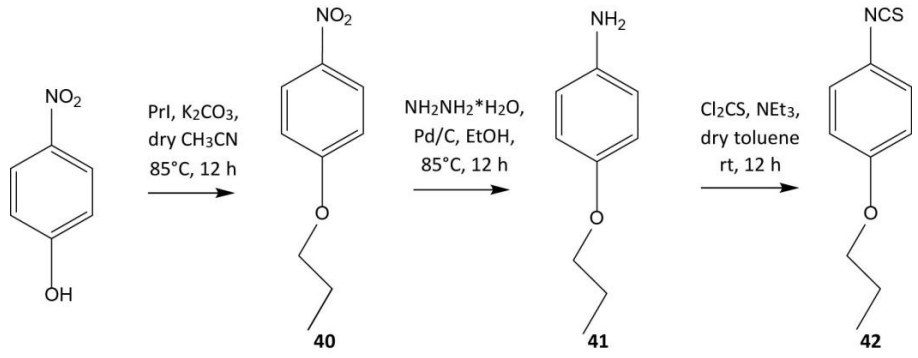
A third noteworthy correlation is between the same thioureido NH and the signal at 5.80 ppm where H-3 and NHAc are overlapped. Actually, in the modelled structure, this thioureido NH is too distant from both, then other investigations in term of modelling will be necessary to justify the specific interaction.

Unfortunately, for the correlation observed in the structure between H-1 and the central CH_2 in the linker (≈ 2.5 Å), there was not evidence in the NOESY spectrum, while cross peaks were present between H-1 and the CH_2 nearby to the anomeric oxygen and with the CH_2

adjacent the thiourea or H_5 (overlapping signals even in this case). However these correlations are expected since H_1 and H_5 are both axial, pointing in the α face.

3.2.5 Synthesis of the monomeric glycosylated system

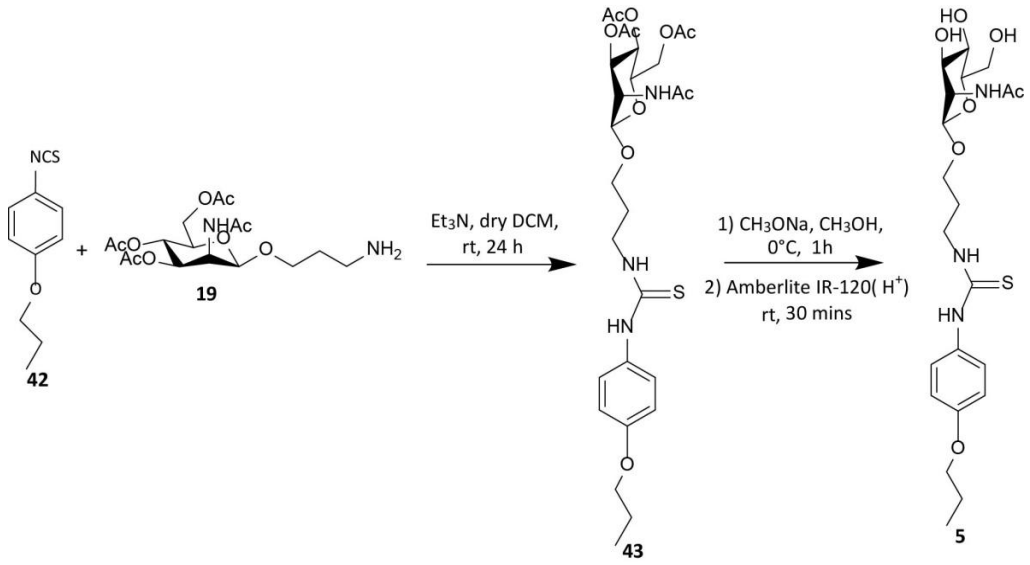
Meanwhile a monovalent acyclic system (compound **5**) was also synthesized in order to verify if the macrocyclic pre-organized structure of calixarene can enhance the binding affinity toward the antibody. This is constituted by a single aromatic unit that will be bearing only one N-acetyl-mannosamine unit, as shown in Figure **6**. The synthesis was carried out using the same approach for the calixarenes (see below **Scheme 16**), starting from the commercially available p-nitro phenol. The alkylation of the hydroxyl group with propyl iodide, with K_2CO_3 as base, gave compound **40**, without further purification and in quantitative yield. The next step was the reduction of the nitro group to the amino group using hydrazine, Pd/C (10%) in ethanol. Also in this case **41** did not need purifications and the yield were high (84%). The achievement of compound **41** was confirmed by ^{1}H-NMR spectra that shown the shift of the aromatic signals at higher field, with the appearance of the NH_2 broad singolet at 3.37 ppm. The isothiocyanate group was introduced using thiophosgene with Et_3N in dry toluene[87] and the product was obtained in a yield of 32%. In this case the compound was purified by a flash chromatography, giving **42** as a yellow oil.



Scheme 16: preparation of the acyclic isothiocyanate monomer **42**

After the preparation of the acyclic monomeric scaffold, the conjugation reaction was carried out similarly to that on the calixarene scaffold.(**Scheme 17**). A click reaction between isothiocyanate group on the aromatic ring and the amino group on the sugar allowed the formation of derivative **43**, with a thiourea spacer. The reaction did not give significant by-

products and a purification step was necessary to remove only the excess of sugar **19**. The presence of the product was confirmed by NMR spectroscopy, ESI-MS and it was fully characterized (figure **7A** in the appendix). The last step was the deprotection of the acetyl protecting groups of the sugar. The reaction was carried out using Zemplèn method[84] (sodium methoxide in methanol) again at 0°C, to avoid epimerization.[85] Completed removal of acetyl groups was observed after 1 hour and the reaction was stopped adding Amberlite resin IR-120(H^+) to have a transesterification process that allow the formation of the desired product **5** in high yield (78%). The ^{1}H-NMR is reported in the appendix (figure **8A**).



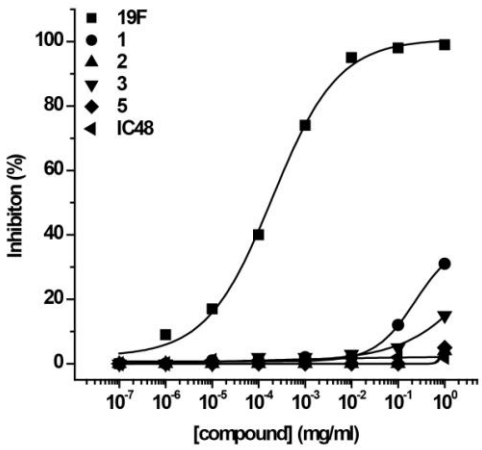
Scheme 17: synthetic scheme of the formation of monoglycoside **40**

3.2.6 Biological studies

Compounds were studied by the group of Grazia Lombardi at the Università del Piemonte Orientale.

Biological properties of synthesized compounds have been investigated in order to compare them with the native polysaccharide 19F (as positive control). Classical competitive ELISA assays were performed, evaluating the capability of increasing concentrations of every multivalent molecule to inhibit the interaction between the polysaccharide 19F, linked to the plates, and its specific anti-19F antibody present in the human system. The graph in figure **12** shows the preliminary data obtained. As expected, the native polysaccharide exhibited higher affinity (IC_{50} value of 1.9 x 10^{-4} mg/ml) and even a higher efficacy (100% of inhibition at 10^{-2} mg/ml) compared to the prepared derivatives **1-5**.

Concerning the calixarene derivatives, the mobile calixarene **1** behaved as better inhibitor compared to the other ones. This could be due to the higher mobility of this derivative that in the presence of the antibody can adopt the best conformation to fit the binding sites of the receptor. Unfortunately so far we could not test the other mobile glycocalixarene **4** which could confirm this hypothesis. Also the *alternate*-4N-ManAc[4]Pr **3** evidenced some interaction ability towards anti-19F antibody but with no high affinity and also lower efficacy (4% as maximum inhibition). Interestingly, the lowest activity was shown by the monomeric model **5**, the cone derivative **2** and a galactosylcalix[4]arene (**IC48**) conformationally mobile as **1**, used as model to verify that the sugar was responsible of the interaction between **1** and anti-19F. The absence of activity in the monomer demonstrated that glycoside cluster effect is playing a role in the studied process. The absence of activity for the cone **2** derivative suggests that the hindrance due to the simultaneous presence of four sugar units at the same part of the space is detrimental for the interaction. This also explains why the 1,3-alternate **3**, less crowded, showed some activity. The absence of activity for **IC48** indicates that the inhibition of the binding between anti-19 F and the polysaccharide obtained with **1** is due to a specific interaction between the N-acetylmannosamine units linked to the calixarene scaffold and not to aspecific contacts with other parts of the macrocyclic structure. These elements indicate that, although not very high the inhibition activity found, compound **1** acts through interesting mechanism that justify a development of the work, in particular with the native trisaccharide.



Compound	IC_{50} (mg/ml)	Maximum Inhibition[a] (%)
19F	1.9×10^{-4}	100 ± 2
1	2.3×10^{-1}	31
2	-	4
3	-	15
5	-	5
IC48	-	2

Fig. 12: Concentration-response curves of saccharides on the inhibition of the binding between the 19F polysaccharide, coated onto the plates, and the anti-19F human polyclonal antibody were evaluated by a competitive ELISA method. In the table on the right are reported the IC_{50} values and the inhibition potency for each studied calixarene.

Moreover, a further important consideration can be done about the results found through these inhibition tests. The difference in efficiency between the natural polysaccharide and the best calixarene derivative **1** is really high but we have to consider that it is constituted by a high number of the trisaccharide repeating units (reported above in figure 2.4). On the contrary derivative **1** is exposing only 4 units and these are simply constituted of N-acetyl-mannosamine, one of the three monosaccharides included in the natural trisaccharide. In light of this fact, the IC_{50} value obtained with **1** must be seen as a very interesting result, suggesting that the planned functionalization of the calixarene scaffold with even only four units of trisaccharide could greatly enhance the activation of the immune response.

3.3 Conclusions

Since carbohydrates play an important role in the activation of the immune response, the research aimed at finding molecules more efficient in the inhibition of the infections. Particularly in the field of vaccinology, synthetic carbohydrate-based vaccines have several advantages and so the scientific community have focalized the attention on those, for the creation of more active systems against viruses, bacteria and also tumours. Among bacteria *Streptococcus pneumoniae* serotype 19F is one of the most virulent and thus the aim of this chapter was the preparation of different glycocalix[n]arenes (n=4,6 and 8) to verify to study their affects on the immunological potential of the saccharidic antigen. The calixarene scaffolds were functionalized at the upper rim with thiourea linker and N-acetylmannosamine as saccharidic unit, bearing a three carbon chain, since this sugar has the pivotal role in the activation of the defence mechanism. There were prepared macrocycles with different valency and stereochemistry in order to see if multivalency and if the diverse position of the ligating units in the space can play a role in the enhancement of the immune response. Calix[6]arene and a mobile calix[4]arene were prepared taking advantage of their higher mobility, due to the occurrence of methyl groups at the lower rim.

In this way they eventually could adapt their conformation to optimize the binding process in terms of binding units employed for this recognition phenomenon.

The synthesized glycocalixarenes bearing the antigenic unit, were studied at the University of Piemonte Orientale in Novara by Dr. Silvia Fallarini to compare their inhibition potency with the natural polysaccharide against the binding of the anti-19F human polyclonal antibody.

Preliminary data shown that the more efficient glycocalixarene was **1**. This behavior suggests us to say this conformationally mobile calix[4]arene has the possibility to explore the antibody surface and dispose the N-acetylmannosamine units in a better way to fit that. However the calix[6]arene **4** was sent after all these compounds studied and we are waiting for the results to see if the highest valency could enhance more the inhibition potency or not.

3.4 Experimental part

General Information. All moisture sensitive reactions were carried out under nitrogen or argon (if specifically required)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 aluminium) 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), Pancaldi solution (molybdatophosphorus acid and Ce(IV)sulphate in 4% sulphuric acid). Reverse phase TLC were performed by using silica gel 60 RP-18 F-254 on aluminium sheets. Merck silica gel 60 (70-230 mesh) was used for flash chromatography and for preparative TLC plates. Sigma Aldrich C18 reverse phase silica gel was used for flash chromatography. Pharmacia Sephadex G-25 (medium) was used for size-exclusion chromatography. ^{1}H NMR and ^{13}C NMR spectra were recorded on Bruker AV300 and Bruker AV400 spectrometers (observation of ^{1}H 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, δ_{TMS} = 0), as internal standard. All ^{13}C NMR spectra were performed with proton decoupling. For ^{1}H NMR spectra recorded in D_2O at values higher that 25°C the correction of chemical shifts was performed using the expression $\delta = 5.060 - 0.0122 \times T(°C) + (2.11 \times 10^{-5}) \times T^2(°C)$ to determine the resonance frequency of water protons.[88] Electrospray ionization (ESI) mass analyses were performed with a Waters spectrometer in both positive and negative mode with MeOH/CH_3CN as solvents. Melting points were determined on an Electrothermal apparatus in closed capillaries. Microwave reactions were performed using CEM Discovery System reactor.

Nomenclature for the saccharidic units

For the saccharidic compounds the numeration was given based on the following figure, where the aminopropylic chain of the anomeric carbon has a progressive number respect to the principal sugar structure, to easily explain NMR spectra:

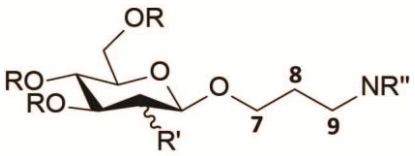


Fig 2.12: schematic numeration of carbon chain for the NMR data

5,11,17,23,29,35-hexa-tert-butyl-37,38,39,40,41,42-hexahydroxycalix[6]-arene (6d)

In a three neck round bottom flask, equipped with mechanical stirrer and a Dean-Stark trap and condenser, under nitrogen, were added p-tert-butylphenol (100g, 0.66 mol), formaldehyde solution 37% (137 ml, 4.210 mol) and KOH pellets (15.26 g, 0.27 mol). The reaction was stirred and heated at 80°C for 2 hours. When the mixture colour started to change into a lemon yellow, xylene (1 L, 8.29 mol) was added to dissolve the semisolid mass and the temperature was increased up. The reaction was monitored by TLC(petroleum ether/DCM 1:1) and after 3 hours the brown solid was filtered on Buchner funnel. the crude was dissolved in chloroform (1.5 L, 18.72 mol) and a HCl 1M solution was added (0.8 L), leaving under stirring for 15 minutes. The green organic phase was separated and washed with chloroform (3x200 ml) and after the combined yellow phases were pooled together and dried over anhydrous Na_2SO_4, filtered and the solvent evaporated under reduced

pressure. The volume was reduced down to 1L and warm acetone (1L, 40°C) was added, leaving the mixture to cool down. The solid was filtered on Buchner funnel, to obtain **6d** as a white solid. Yield = 83%. **^{1}H-NMR** ($CDCl_3$, 400 MHz): δ = 10.52 (s, 6H, O*H*), 7.14 (s, 12H, Ar*H*), 3.94 (s, 12H, ArCH_2Ar), 1.26 (s, 54 H, C(CH_3)$_3$). The other physico-chemical properties are the same of those reported in literature.[56]

5,11,17,23-Tetra- tert-butyl-25,26,27,28-tetra-methoxycalix[4]arene (7a)

In a 2-neck round bottom flask to a solution of **6** (5g, 7.71 mmol) in 80 ml DMF dry, under N_2, cooled down to 0 °C. NaH 60% (3.08 g, 77.71 mmol) was added and the reaction stirred for 30 minutes. At the end CH_3I (5.8 ml, 92.46 mmol) was added and the mixture was kept a room temperature for 2 days and monitored by TLC (eluent: hexane/EtOAc 20:1); when finished it was quenched adding H_2O (150 ml), HCl 1M (150 ml) and the solid was filtered with Buchner funnel. The off-white solid was then purified with cold crystallization in DCM/MeOH to get compound **7a** as a white solid in 84% yield. **^{1}H-NMR**: (400 MHz, $CDCl_3$) δ (ppm): 6.65 (br s, 8H, Ar*H*); 4.25-2.90 (m, 20H, ArCH_2Ar and OCH_3); 1.25 (br s, 36H, *t*-Bu).
The product has the same spectroscopic characteristics reported in literature.[89]

Cone-5,11,17,23-Tetra-tert-butyl-25,26,27,28-tetrapropoxycalix[4]-arene (7b)

In a 2-neck round bottom flask to a solution of **6** (4g, 6.16 mmol) in 80 ml DMF dry, under N_2, cooled down to 0°C. NaH 60% (3.9g, 98.62 mmol) was added and the reaction stirred for 30 minutes. At the end PrI (7.2 ml, 74 mmol) was added and the mixture was kept a room temperature for 2 days and monitored by TLC (eluent: hexane/EtOAc 20:1); when finished it was quenched adding H_2O (200 ml), HCl 1M (200 ml) and the solid was filtered with Buchner funnel. The off-white solid was then purified with cold crystallization in DCM/MeOH to get compound **7a** as a white solid in 23% yield. **^{1}H-NMR**: (400 MHz, $CDCl_3$) δ (ppm): 6.80 (s, 8H, Ar*H*); 4.44 (d, 4H, *J* = 12.4 Hz, ArC*H*Ar ax); 3.84 (t, 8H, *J* = 7.2 Hz, OCH_2CH$_2$CH$_3$); 3.14 (d, 4H, *J* = 12.4 Hz, ArC*H*Ar eq); 2.05 (m, 8H, CH$_2$CH_2CH$_3$); 1.02 (t, 12H, *J* = 7.2 Hz, CH$_2$CH$_2$CH_3).
The product has the same spectroscopic characteristics reported in literature.[90]

***1,3-alternate*-5,11,17,23-Tetra-*tert*-butyl-25,26,27,28-tetrapropoxycalix[4]arene (7c)** [83]

In a two round bottom flask, under nitrogen, tetra tert-butyl-hydroxycalix[4]arene **6** (3.2 g, 4.9 mmol) was dissolved in dry CH_3CN (40 ml) at room temperature. $CsCO_3$ (16 g, 49 mmol) was added and the reaction was kept under stirring for 15 minutes. After the addition of 1-iodopropane (4.89 ml, 49 mmol) the temperature was raised up to 85°C for 48 h. The reaction was monitored by TLC (eluent: hexane/EtOAc 9:1) and quenched with HCl 1M (50 ml)and extracted with DCM. The organic phase was washed with H_2O (50 ml) until neutral pH. The organic extract was dried over anhydrous Na_2SO_4, filtered and concentrated under reduced pressure, giving a white solid. The residue was crystallized using petroleum ether. After filtration on Buchner filter **7c** was obtained as white solid in 37% yield. **^{1}H-NMR**: (400 MHz, $CDCl_3$) δ (ppm): 6.98 (s,8H, ArH), 3.83 (s, 8H, ArC*H*$_2$Ar), 3.32 (t, 8H, *J* = 7.6 Hz,O*CH*$_2$), 1.40-1.30 (m, 36H, t-Bu), 1.20-1.00 (m, 8H, OCH_2CH_2), 0.90-0.75 (m, 12H, $OCH_2CH_2CH_3$).

The product shows the same physical-chemical properties reported in literature.[57]

5,11,17,23,29,35-Hexa-*tert*-butyl-37,38,39,40,41,42-hexamethoxycalix[6]arene (7d)

In a two round bottom flask **6** (5 g, 6.14 mmol) was dissolved in dry DMF (200 ml) under N_2. The reaction was cooled down in a ice bath and NaH 60% (2.95 g, 73.62 mmol) was added, leaving the stirring for 30 minutes. After that time CH_3I (4.58 ml, 73.62 mmol) was added with DMF (100 ml) and the temperature was increased up to 90°C refluxing for 48 h. When by TLC (eluent hexane/EtOAc 8:2) the only presence of the product was verified, the quenching was done adding HCl 1M (300 ml). The organic phase was washed with brine until neutral pH, then dried with Na_2SO_4 and evaporated under reduced pressure. After crystallization in MeOH, the product **7d** was obtained as yellowish sold (yield: 62%).

^{1}H-NMR (400 MHz, $CDCl_3$): δ (ppm) 7.03 (s, 12H, H meta), 3.97 (s, 12H, Ar*CH*$_2$Ar), 3.00 (s, 18H, O*CH*$_3$), 1.16 (s, 54H, C(*CH*$_3$)$_3$).

The spectroscopic data found are in agreement with those reported in literature.[91]

General procedure for the ipso nitration of calix[n]arenes 7a-e

Calixarene (1 mmol) and $NaNO_3$ (10 mmol for each *tert*-butyl group) were put into a round-bottom flask and then CF_3COOH (10 mmol for each *tert*-butyl group) was added dropwise. The mixture was allowed to stir at rt overnight. The reaction was stopped by addition of water (200 ml) and extracted with CH_2Cl_2 (2x100 ml). The combined organic layers were

washed with water (150 ml) and dried over Na_2SO_4. The solvent was removed under reduced pressure.

5,11,17,23-Tetranitro-25,26,27,28-tetramethoxy-calix[4]arene (8a)

The pure product was obtain by crystallization with CH_2Cl_2 / MeOH as a pale yellow powder (77%). Two conformers were present in the NMR spectrum: cone, (10%) and partial cone (90%). **^{1}H-NMR** (300 MHz, $CDCl_3$) δ (ppm): 8.22 (s, 1.8H, Ar*H*, partial cone); 8.13 (s, 1.8*H*, Ar*H*, partial cone); 7.85 (d, 1.8H, *J* = 2.7 Hz, Ar*H*, partial cone); 7.70 (s, 0.8H, Ar*H*, cone); 7.26 (d, 1.8H, *J* = 2.7 Hz, Ar*H*, partial cone); 4.43 (d, 0.4H, *J* = 13.2 Hz, ArC*H*Ar ax., cone); 4.09 (d, 1.8H, *J* = 14.1 Hz, Ar*CH*Ar ax., partial cone); 3.91 (s, 3.6H, ArCH_2Ar, partial cone); 3.85 (s, 1.2H, OCH_3, cone); 3.83 (s, 2.7H, OCH_3, partial cone); 3.82 (s, 5.4H, OCH_3, partial cone); 3.47 (d, 0.8H, *J* = 13.2 Hz, ArC*H*Ar eq., cone); 3.36 (d, 1.8H, *J* = 14.1 Hz, ArC*H*Ar eq., partial cone); 3.04 (s, 2.7H, OCH_3, partial cone). The spectroscopic data found are in agreement with those reported in literature.[47]

***Cone*-5,11,17,23-Tetranitro-25,26,27,28-tetrapropoxy-calix[4]arene (8b)**

The residue was crystallized with CH_2Cl_2/MeOH and the solid was filtered on Buchner funnel to get a yellow-orange solid (96% yield). **^{1}H-NMR** (300 MHz, $CDCl_3$) δ (ppm): 7.60 (s, 8H, Ar*H*); 4.55 (d, 4H, *J* = 10.5 Hz, ArC*H*Ar ax.); 3.98 (t, 8H, *J* = 5.7 Hz, OCH_2CH$_2$CH$_3$); 3.43 (d, 4H, *J* = 10.5 Hz, ArC*H*Ar eq); 1.93 (m, 8H, OCH$_2$CH_2CH$_3$); 1.05 (t, 12H, *J* = 5.4 Hz, OCH$_2$CH$_2$CH_3).
The product has the same physical-chemical properties reported in literature.[47]

***1,3-alternate*-5,11,17,23-Tetranitro-25,26,27,28-tetrapropoxy -calix[4]arene (8c)**

The residue was purified with crystallization in CH_2Cl_2/MeOH and the solid was filtered obtaining a yellowish solid (93% yield). **^{1}H-NMR** (400 MHz, $CDCl_3$): δ(ppm) 7.95 (s, 8H, Ar*H*), 3.79 (t, 8H, *J* = 7.2 Hz, OCH$_2$), 3.73 (s, 8H, ArCH_2Ar), 1.89-1.84 (m, 8H, OCH$_2$CH_2), 1.04 (t, 12H, *J* = 7.2 Hz, OCH$_2$CH$_2$CH_3).
The product has the same physical-chemical properties reported in literature.[47]

5,11,17,23,29,35-hexanitro-37,38,39,40,41,42-hexamethoxycalix[6]arene (8d)

The residue was purified with crystallization in CH_2Cl_2/MeOH (10%) and the solid was filtered on Buchner funnel to obtain a yellow solid (34% yield). **^{1}H-NMR** (400 MHz, $CDCl_3$): δ = 7.71 (s, 12H, Ar*H*), 4.15 (s, 12H, ArC*H*$_2$Ar), 3.83 (s, 18H, OC*H*$_3$).

The spectroscopic data found are in agreement with those reported in literature.[92]

5,11,17,23,29,35,41,47-Octanitro-49,50,51,52,53,54,55,56-octamethoxyicalix[8]arene (8e)

The residue was purified with trituration in DCM/MeOH 9/1 to obtain compound **8e** in 34% yield. **^{1}H-NMR** (300 MHz, $CDCl_3$) δ (ppm): 7.77 (s, 16H, Ar*H*); 4.16 (s, 16H, ArC*H2*Ar); 3.80 (s, 24H, OCH_3). The analytical data corresponded to the literature.[92]

General procedure for the for the reduction of nitrocalix[n]arenes 8a-e

To a suspension of nitrocalix[n]arene (0.1 mmol) either in ethanol (**8a-d**) or in 2-methoxyethanol (**8e**), NH_2NH_2*H_2O (0.5 mmol for each nitro group) and Pd/C (10%) (catalytic amount) were added. The reaction mixture was refluxed at 85°C and stirred overnight. At the end the catalyst was filtered off through a paper filter.

5,11,17,23-Tetraamino-25,26,27,28-tetramethoxy-calix[4]arene (9a)

The produc was obtained as a white solid (90% yield). In solution two conformation were still present: cone(10% ca.) and partial cone (90% ca.). **^{1}H-NMR** (400 MHz, MeOD) δ (ppm): 7.34 (s, 1.6H, Ar*H*, partial cone); 7.11 (s, 1.6H, Ar*H*, partial cone); 6.98 (s, 1.6H, Ar*H*, partial cone); 6.73 (s, 0.8H, Ar*H*, cone); 6.36 (s, 1.6H, Ar*H*, partial cone); 4.42 (d, 0.8H, *J* = 14.0 Hz, ArC*H*Ar ax., cone); 4.10 (d, *J* =14.0 Hz, 1.6H, Ar*CH*Ar ax., partial cone); 3.84 (s, 2.4H, ArC*H*$_2$Ar, partial cone); 3.79 (s, 2.4H, OC*H*$_3$, cone); 3.77 (s, 2.4H, OC*H*$_3$, partial cone); 3.75 (s, 2.4H, OC*H*$_3$, partial cone); 3.30 (d, 0.8H, *J* = 14.0 Hz, ArC*H*Ar eq., cone); 3.23 (d, 3.2H, *J* = 14.0 Hz, ArC*H*Ar eq., partial cone); 3.13 (s, 2.4H, OC*H*$_3$, partial cone).

The analitycal data are in agreement with those found in literature.[47]

***cone*-5,11,17,23-Tetraamino-25,26,27,28-tetrapropoxy-calix[4]arene (9b)**

The filter was washed with DCM/EtOH and the solvent removed under vacuum, to obtain a white solid in 92% yield. **^{1}H-NMR** (300 MHz, $CDCl_3$) δ (ppm): 6.05 (s, 8H, Ar*H*); 4.31 (d, 4H, *J* = 13.2 Hz, ArC*H*Ar ax.); 3.72 (t, 8H, *J* = 7.3 Hz, OC*H*$_2$$CH_2CH_3$); 3.15 (br s, 8H, N*H*$_2$); 2.91 (d, *J* =

13.2 Hz, 4H, ArC*H*Ar eq.); 1.86 (m, 8H, $OCH_2CH_2CH_3$); 0.95 (t, 12H, *J* = 7.4 Hz, $OCH_2CH_2CH_3$).
The spectroscopic data found are in agreement with those reported in literature.[93]

***1,3-Alternate*-5,11,17,23-Tetraamino-25,26,27,28-tetrapropoxy-calix[4]arene (9c)**

After washing the filter with EtOH and the solvent was removed under reduced pressure. The pure product was obtained as beige solid (95%). **^{1}H-NMR** (400 MHz,$CDCl_3$): δ(ppm) 6.44 (s, 8H, Ar*H*), 3.51 (t, 8H, *J* = 7.2 Hz, OCH_2), 3.41 (s, 8H, $ArCH_2Ar$), 3.16 (s, 8H, NH_2), 1.79-1.70 (m, 8H, OCH_2CH_2), 1.03 (t, *J* = 7.2 Hz, 12H, $OCH_2CH_2CH_3$).
The analitycal data are in agreement with those found in literature.[47]

5,11,17,23,29,35-Hexaamino-37,38,39,40,41,42-hexamethoxycalix[6]arene*6 HCl (9d)

The paper filter was washed with CH_2Cl_2 and HCl 1M and the solvents were removed under reduced pressure. The residue was dissolved in a small amount of CH_2Cl_2 and the pure product was obtained after trituration with hexane as a yellowish powder (65%). **^{1}H-NMR** ($CDCl_3$, 300 MHz): δ (pmm) 6.23 (s, 12H, Ar*H*), 3.82 (s, 12H, $ArCH_2Ar$), 3.42 (s, 18H, OCH_3), (s, 18H, NH_3^+). The product shows the same characteristics present in literature.[92]

5,11,17,23,29,35,41,47-Octaamino-37,38,39,40,41,42,43,44-octamethoxycalix[8]arene*8 HCl (9e)

The paper filter was washed with CH_2Cl_2 and HCl 1M and the solvents were removed under reduced pressure. The pure product was obtained after trituration with hexane as a brown powder (90% yield). **^{1}H-NMR** (300 MHz, $CDCl_3$) δ (ppm): 6.87 (s, 16H, Ar*H*); 3.97 (s, 16H, $ArCH_2Ar$); 3.46 (s, 24H, OCH_3).
The spectroscopic data found are in agreement with those reported in literature.[92]

General procedure for the isothiocyanation of aminocalix[n]arenes 9a-c[94]

In a two round bottom flask aminocalixarene (1 mmol) was dissolved in toluene dry. Thiophosgene (3 equivalents for each amino group) and Et_3N (6 equivalents for each amino group) were added and the reaction was stirred at rt for 12-48 h. The reaction was monitored by TLC and was quenched with H_2O, extracted with DCM. The organic phase was

washed with HCl 1 M (1 time), brine (2 times) and the solvent was removed under reduced pressure. The residue was purified with flash chromatography.

5,11,17,23-Tetraisothiocyanate-25,26,27,28-tetramethoxy-calix[4]arene (10a)

The mixture was purified under flash chromatography (hexane/DCM 7:3) to yiel the pure product as a yellowish solid (23%). In solution two conformers were present: cone (20%ca.) and partial cone (80%ca.). **^{1}H-NMR** (300 MHz, $CDCl_3$) δ (ppm): 7.14 (s, 1.6H, Ar*H*, partial cone); 7.09 (s, 1.6H, Ar*H*, partial cone); 6.81 (s, 1.6H, Ar*H*, partial cone); 6.65 (s, 1.6H, Ar*H*, cone); 6.29 (s, 1.6H, Ar*H*, partial cone); 4.26 (d, 0.8H, *J* = 13.5 Hz, ArC*H*Ar ax., cone); 3.97 (d, 1.6H, *J* =14.0 Hz, Ar*CH*Ar ax., partial cone); 3.78 (s, 2.4H, OC*H*$_3$, cone); 3.74 (s, 2.4H, OC*H*$_3$, partial cone); 3.68 (s, 4.8H, OC*H*$_3$, partial cone); 3.56 (s, 3.2H, ArC*H*$_2$Ar, partial cone); 3.13 (d, 0.8H, *J* = 13.5 Hz, ArC*H*Ar eq, cone); 3.06 (d, 1.6H, *J* = 14.0 Hz, ArC*H*Ar eq, partial cone); 3.04 (s, 2.4H, OC*H*$_3$, partial cone). **^{13}C NMR** (75MHz, $CDCl_3$) δ (ppm): 127.7, 126.5, 126.2, 125.7, 125.6 (*CAr*, both cone and partial cone); 61.9, 61.4, 60.2, 59.8 (O*C*H_3); 35.0 (Ar*C*H_2Ar, partial cone); 30.3, 30.1 (Ar*C*H_2Ar, both cone and partial cone). **ESI-MS(+):** m/z 731 [100%, $(M+Na)^+$].

M.p. : > 180 °C dec

***Cone*-5,11,17,23-Tetraisohiocyanate-25,26,27,28-tetrapropoxy-calix[4]arene (10b)**

The crude was purified by flash chromatography (eluent: cyclohexane/DCM 4:1) to yield **10b** as a white solid (40%). **^{1}H-NMR** (400 MHz, $CDCl_3$) δ (ppm): 6.55 (s, 8H, Ar*H*); 4.35 (d, 4H, *J* = 13.6 Hz, Ar*CH*Ar ax.); 3.79 (t, 8H, *J* = 7.6 Hz, OC*H*2CH2CH3); 3.08 (d, 4H, *J* = 13.6 Hz, Ar*CH*Ar eq.); 1.85 (m, 8H, $OCH_2CH_2CH_3$); 0.96 (t, 12H, *J* = 7.6 Hz, $OCH_2CH_2CH_3$). The product shows the same physical-chemical properties reported in literature. [95,96]

***1,3-Alternate*-5,11,17,23-Tetraisothiocyanate-25,26,27,28-tetrapropoxy-calix[4]arene (10c)**

The crude was purified by flash chromatography (eluent: cyclohexane/DCM 4:1) to yield the pure product as a off-white solid (33%). **^{1}H-NMR** (400 MHz, $CDCl_3$) δ (ppm): 6.90 (s, 8H, Ar*H*), 3.65 (t, 8H, J=7.2, OC*H*$_2$), 3.47 (s, 8H, ArC*H*$_2$Ar), 1.87-178 (m, 8H, OCH_2CH_2), 1.09 (t, J=7.2, 12H, $OCH_2CH_2CH_3$). **^{13}C-NMR** (100 MHz, $CDCl_3$): δ(ppm) 155.3 (C ipso), 134.0 (N*C*S), 127.1 (Ar), 125.0 (C para), 74.7 (O*C*H_2),35.1 (Ar*C*H_2Ar), 23.9 (*C*H_2CH_3), 10.7 (*C*H_3). **ESI-MS(+):** m/z 731 [100%, $(M+Na)^+$].

M.p. : 220.8-221.3 °C

5,11,17,23,29,35-Hexaisothiocyanate-37,38,39,40,41,42-hexamethoxycalix[6]arene*6 HCl (10d)

In a round-bottom flask **9d** (0,150 g, 0.185 mmol) was dissolved in DCM. After Cl_2CS (254 µl, 3.33 mmol), $BaCO_3$ (0,657 g, 3.33 mmol) and H_2O were added with the remaining amount of DCM. The reaction was left at rt for 48 h and when finished (TLC eluent: hexane/EtOAc 8:2) the mixture was diluted with DCM/ H_2O. After the extraction the organic phase was removed under reduced pressure to give a crude that was purified by flash chromatography (eluent: hexane/EtOAc 8:2) and crystallized with CH_3CN to yield the pure product as a white solid (20%). **^{1}H-NMR** ($CDCl_3$, 300 MHz): δ (ppm) δ = 6.74 (s, 12H, Ar*H*), 3.89 (s, 12H, Ar*CH*$_2$Ar), 3.54 (s, 18H, OC*H*$_3$). **^{13}C-NMR** (75 MHz, $CDCl_3$): δ(ppm) 155.3 (C ipso); 135.2 (C-orto); 133.9 (N*C*S); 126.6 (C-para); 126.2 (C-meta); 61.0 (O*C*H_3); 30.2 (Ar*C*H_2Ar). **ESI-MS(+):** m/z 1085.2 [80%, $(M+Na)^+$], m/z 1101.2 [100%, $(M+K)^+$].

M.p.: dec > 200°C

1,2,3,4,6-penta-*O*-acetyl-β-D-glucopyranoside (11)

In a two neck round bottom flask sodium acetate (95.61 g, 1.17 mol) was suspended in acetic anhydride (390 ml, 4.13 mol) and beate at 130°C. At that temperature D-glucose (30.0 g, 0.167 mol) was added batch wise for 3 h, monitoring the reaction via TLC (eluent: hexane/EtOAc 1:1). When it was finished, it was poured into a flask containing ice and water (5 times the initial volume of acetic anhydride). The mixture was stirred at room temperature overnight. The solid was recovered by filtration on Buchner funnel and then dissolved in DCM and extracted with water. The organic phase was evaporated under reduced pressure to give rise to **11** in 65% yield as a white solid and a mixture of α,β anomers. **^{1}H-NMR** ($CDCl_3$) δ (ppm): 6.32 (d, 0.09 H, $J_{1\text{-}2}$ = 3.6 Hz, H_1 α); 5.70 (d, 1H, $J_{1\text{-}2}$ = 8.1 Hz, H_1 β); 5.46 (t, 0.09 H, J = 10.0 Hz, H_3 α); 5.29-5.21 (m, 1H, H_3 β); 5.15-5.08 (m, 2H, H_2 and H_4); 4.29 (dd, 1H, $J_{5\text{-}6a}$ = 4.4, $J_{6a\text{-}6b}$= 12.4 Hz, H_{6a}); 4.11 (dd, 1H, $J_{5\text{-}6b}$ = 4.4, $J_{6a\text{-}6b}$= 12.4 Hz, H_{6b}); 3.85-3.80 (m, 1H, H_5); 2.11 (s, 3H, OAc), 2.09 (s, 3H, OAc); 2.03 (s, 6H, 2xOAc), 2.01 (s, 3H, OAc).

The product shows the same spectroscopic characteristics reported in literature. [63]

2,3,4,6-Tetra-*O*-acetyl-α-D-glucopyranosyl bromide (12)

In a round bottom flask peracetylated glucose **11** (42.06 g, 0.11 mol) was dissolved in 60 ml of dry CH_2Cl_2 and a 33% solution of HBr in acetic acid (43 ml, 0.75 mol)) was added. The reaction was stirred at room temperature for 3 h and checked via TLC (eluent: hexane/EtOAc 1:1). When completed , the mixture was diluted in CH_2Cl_2 (30 ml) and transferred in a Becker containing 300 ml of ice and water. $NaHCO_3$ was added drop wise under stirring and after that the organic phase was washed with a saturated solution of $NaHCO_3$ (3x 90 ml) until neutral pH and H_2O (1x90 ml). The organic phase was dried over anhydrous Na_2CO_3, filtered off and the solvent removed under reduced pressure to give compound **12** as a white solid. (yield = 84%) **^{1}H-NMR** (400 MHz, $CDCl_3$) δ (ppm): 6.61 (d, 1H, *J* = 4.0 Hz, H_1); 5.56 (t, 1H, *J* = 9.6 Hz, H_3); 5.16 (t, 1H, *J* = 9.6 Hz, H_4); 4.84 (dd, 9.6 Hz, $J_{1\text{-}2}$ = 4.4, $J_{2\text{-}3}$= 12.4 Hz = 4.0, 1H, H_2); 4.35-4.25 (m, 2H, H_5, H_{6b}); 4.11 (dd, 1H, *J* = 2.3, 10.5 Hz, H_{6a}); 2.09 (s, 3H, OAc), 2.07 (s, 3H, OAc), 2.05(s, 3H, OAc), 2.03 (s, 3H, OAc).

The spectroscopic data are the same of those found in literature.[97]

2,3,4,6-Tetra-*O*-acetyl-1,5-anhydro-D-arabino-hex-1-enitol (13)

In a two neck round bottom flask, under nitrogen, **12** (4.9 g, 11.92 mmol) and NaI (6.79 g, 45.28 mmol) was added in dry acetone (60ml). The reaction was stirred at rt for 20 minutes and then Et_2NH (13.6 ml) was added drop wise. The mixture was kept at room temperature under stirring for 16 h and checked via TLC (eluent: hexane/EtOAc 1:1). When finished the precipitate present in the crude was removed on a paper filter, washing with acetone. The solvent was removed under reduced pressure and then dissolved in DCM (50 ml) and extracted with H_2O (30 ml). The organic phase was washed with HCl 1M (30 ml) and again H_2O (4x30 ml) and then the combined phases were evaporated under vacuum. The crude was purified using flash chromatography (eluent: hexane/EtOAc 1:1) and then with a crystallization in hexane to obtain compound **13** in 60% yield as a white solid. **^{1}H-NMR** (400 MHz, $CDCl_3$): δ (ppm) 6.62 (s, 1H, H_1), 5.57 (d, 1H, *J* = 4 Hz, H_3), 5.04 (t, 1H, *J* = 4.8 Hz, H_4), 4.22 (m, 2H, H_5, H_{6b}), 4.06 (d, 1H, $J_{1\text{-}2}$ = 3.2 Hz, $J_{6a\text{-}6b}$ = 11.2 Hz, H_{6a}), 1.85, 1.84, 1.83, 1.82 (4s, 12H, 4xCO*C*H_3). The physical-chemical data are the same of those reported in literature.[64]

3,4,6-Tri-O-acetyl-1,5-anhydro-D-glucose oxime (14)

To a solution of hydroxyglucal ester **13** (0.08 g, 0.24 mmol) in pyridine (5 ml) was added

NH_2OH*HCl (0.06 g, 0.85 mmol). The mixture was stirred for 16 h at room temperature and monitored by TLC (eluent: hexane/EtOAc 1:1). At the end the crude was diluted with ice/H_2O (10 ml) and extracted with CH_2Cl_2 (2x15 ml). The combined organic phases were washed with H_2O (15 ml) and subsequently with HCl 1M (15 ml), H_2O/ice (3x15 ml).[64] Removal of the solvent in vacuo left a crystalline white residue, which underwent to trituration in diethyl ether to yield **14** (89%) as amorphous white solid. **^{1}H-NMR** (300 MHz, $CDCl_3$): δ (ppm) 10.69 (br s, 1H, N*H*OH), 5.43 (d, 1H, *J* = 6.9 Hz, H_3); 4.98 (t, 1H, *J* = 6.9 Hz, H_4); 4.94 (s, 1H, H_1 eq.); 4.14 (m, 2H, $H_{6a,b}$); 3.98 (d, 1H, $J_{1\text{-}2}$ = 15 Hz, H_1); 3.60 (m, 1H, H_2); 1.92-1.85 (s, 12H, COCH_3). The product shows the same analytical data reported in literature.[72]

3,4,6-Tri-O-acetyl-1,5-anhydro-D-glucose acetyloxime (15)

To a solution of **14** (0.07 g, 0.24 mmol) in 3 ml of pyridine, acetic anhydride (0.24 g, 2.35 mmol) was added and the reaction was kept a room temperature for 4 h and checked by TLC (eluent: hexane/EtOAc 1:1). The mixture was diluted with CH_2Cl_2 (10 ml) and ice/H_2O (15 ml) was added. The organic phase was extracted with H_2O (15 ml) and was subsequently with HCl 1M (15 ml), H_2O/ice (15 ml), $NaHCO_3$ (15 ml) and H_2O (15 ml). The solvent was removed under reduced pressure and the crude was crystallized in EtOH to obtain pure **15** as a white solid in quantitative yield. **^{1}H-NMR** (300 MHz, $CDCl_3$): δ (ppm) 5.61 (d, 1H, *J* = 6.9 Hz, H_3), 5.18 (t, 1H, *J* = 7.8 Hz, H_4); 5.3 (d, 1H, $J_{1eq\text{-}2}$ = 8.1 Hz, H_{1eq}); 4.28-4.20 (m, 3H, H_2, $H_{6a,b}$); 4.14 (d, 1H, $J_{1ax\text{-}2}$ = 16 Hz, H_{1ax}); 3.69 (m, 1H, H_5); 2.16, 2.12, 2.09, 2.07 (s, 12H, 3xCOCH_3, NOAc). The product presents the same physical-chemical data shown in literature.[72]

3,4,6-Tri-O-acetyl-2-acetyloxime-α-D-*arabino*-hexopyranosyl bromide (16)

In a round bottom flask compound **15** (0.07 g, 0.19 mmol) was added in CH_2Cl_2 (5 ml) and subsequently $KBrO_3$ (0.19 g, 1.16 mmol) and $Na_2S_2O_4$ (0.20 g, 1.16 mmol) in acqueous solution (3 ml of each). The reaction was stirred for 24 h at room temperature. When finished (eluent: hexane/EtOAc 1:1) the mixture was dilute with CH_2Cl_2 (5 ml) and a 1 M acqueous solution of $Na_2S_2O_5$ (3 ml) was added, well shaken, and separated. The organic phase was further washed in a saturated solution of $NaHCO_3$ (2x5 ml) and water (5 ml), then dried over $MgSO_4$. After filtration the solvent was removed in *vacuo* and the residue was purified on flash chromatography (eluent: hexane/EtOAc 1:1) and yielded as a yellow solid

(50%). **^{1}H-NMR** (400 MHz, $CDCl_3$): δ (ppm) 7.25 (s, 1H, H_1), 6.19 (d, 1H, *J* = 8.8 Hz, H_3); 5.35 (t, 1H, *J* = 10 Hz, H_4); 4.43-4.31 (m, 3H, H_2, $H_{6a,b}$); 4.14 (d, 1H, *J* = 1.6 Hz, H_5); 2.15, 2.12, 2.10, 2.08 (s, 12H, 3xCOC*H*$_3$, NOAc).

The product presents the same spectroscopic data shown in literature.[66]

3-Azido-1-propanol (17) [98]

In a round bottom flask water NaN_3 (5.6 g, 86.13 mmol), tetra-butylammonium hydroxide (0.8 ml, 0.34 mmol) and 2-chloro-1-propanol (4.5 ml, 53.83 mmol) were added. The reaction was stirred at 80°C for 72 h. When the reaction is finished (TLC eluent: hexane/Et_2O 1:1), the mixture was diluted with H_2O (5ml), extracted with Et_2O (3x15 ml) and then dried over $MgSO_4$. The solvent was removed under vacuo to give the product as a colourless oil in 70% yield. %). **^{1}H-NMR** (400 MHz, $CDCl_3$) δ (ppm): 3.75 (t, 2H, *J* = 6 Hz, CH_2CH_2OH); 3.45 (t, 2H, *J* = 6 Hz, $CH_2CH_2N_3$); 3.18 (s, 1H, OH); 1.8 (q, 2H, *J* = 6 Hz, $CH_2CH_2CH_2$). The analytical data are in agreement with those found in literature.[99]

3,4,6-Tri-*O*-acetyl-α-D-arabino-hexopyranosil-2-ulosyl bromide (20)

To a solution of hydroxyglycal ester **13** (0.2 g, 0.61 mmol) in CH_2Cl_2 (5 ml) was added molecular sieves 3Å and abs. MeOH (30 µl, 0.73 mmol) and the mixture was stirred for 30 minutes at rt. NBS (0.13 g, 0.73 mmol) was added with continuous stirring for additional 20-30 minutes (TLC monitoring). The mixture was then diluted with CH_2Cl_2 (5 ml), washed with 10% $Na_2S_2O_3$ solution (5 ml), water (2x5 ml) and the organic phase was removed under reduced pressure. The resulting oil was triturated in Et_2O and then crystallized in hexane. The pure product was obtain as a colourless oil in 59%yield. **^{1}H-NMR** (400 MHz, $CDCl_3$):δ (ppm) 6.38 (s, 1H, H_1), 5.99 (d, 1H, *J* = 13.6 Hz, H_3), 5.45 (t, 1H, *J* = 13.6 Hz, H_4), 4.52-4.48 (m, 1H, H_5), 4.39 (dd, 1H, $J_{6a\text{-}5}$ = 4 Hz, $J_{6a\text{-}6b}$ = 12.8 Hz, H_{6a}), 4.16 (dd, 1H, $J_{6b\text{-}5}$ = 2 Hz, $J_{6b\text{-}6a}$ = 12.8 Hz, H_{6b}).

The product presents the same physical-chemical characteristics reported in literature.[72]

3,4,6-Tri-O-acetyl-1,2-O-(1-methoxyethylidene)-α-D-glucopyranose (22)

To a solution of α-acetobromoglucose **13** (7.83 g, 19.05 mmol) in dry CH_2Cl_2 (12ml) were added sym-collidine (3.1 ml, 23.82 mmol), TBAB (3.07 g, 9.53 mmol) and dry MeOH (4.1 ml, 100.98 mmol). After being stirred at 45°C for 24 h (TLC eluent: hexane/EtOAc 6:4), the

reaction mixture was diluted with CH_2Cl_2 (10ml) and washed with water (2x25 ml) and brine (25 ml). The organic layer was dried ($NaSO_4$), evaporated under reduced pressure, and the residue was not purified by column chromatography on silica gel. The product **22** was obtained as a pale yellow oil (99%). **^{1}H-NMR** (400 MHz, $CDCl_3$) δ (ppm): 5.72 (d, 1H, J_{1-2} = 4.8 Hz, H_1); 5.18 (t, 1H, *J* = 3.2 Hz, H_3); 4.90 (dd, 1H, *J* = 2.4, 9.6 Hz, H_4); 4.32 (dd, 1H, J_{1-2} = 2.6, J_{2-3} = 5.2, Hz, H_2); 4.19-4.18 (m, 2H, $H_{6a,b}$); 4.10-3.90 (m, 1H, H_5); 3.28 (s, 3H, OCH_3); 2.11 (s, 3H, OAc); 2.09 (s, 3H, OAc); 2.07 (s, 3H, OAc); 1.69 (s, 3H, CH_3). The product presents the same physical-chemical characteristics reported in literature.[100]

3,4,6-Tri-O-benzyl-1,2-O-(1-methoxyethylidene)-α-D-glucopyranose (24)

A solution of **22** (7.57 g, 20.90 mmol) in dry methanol (10 ml)was treated with anhydrous sodium methoxid (pH 8-9) at rt for 1 h. The reaction mixture was then concentrated under reduced pressure to give compound **23**. To a solution of the residue in dry DMF was added portionwise at 0°C sodium hydride 60% (3.715 g, 92.88 mmol). After complete liberation of hydrogen, benzyl bromide (6.4 ml, 53.41 mmol) was added drop wise and the reaction mixture was stirred at rt for overnight. The excess of reagents was destroyed by careful treatment with methanol and then the reaction mixture was concentrated at the minimal volume under reduced pressure and the residue was diluted with Et_2O (100 ml) and extracted with water (2x130 ml).The combined extracts were successively dried ($NaSO_4$), and evaporated under reduced pressure. The residue was purified by column chromatography on silica gel (eluent: hexane/EtOAc/Et_3N 85:15:1 v/v/v) as eluent to give **24** as a yellow oil in 70% yield. **^{1}H-NMR** (400 MHz, $CDCl_3$) δ (ppm): 7.17-7.35 (m, 15H, Ar*H*); 5.77 (d, 1H, J_{1-2} = 5.2 Hz, H_1); 4.70 (d, 1H, *J* = 12.0 Hz, C*H*HPh); 4.72-4.48 (m, 4H, 2xCH_2Ph); 4.48 (d, 1H, *J* = 12.3 Hz, CH*H*Ph); 4.41 (td, 1H, J_{1-2} = 4.4, J_{2-3} = 9.2 Hz, H_2); 3.86 (t, 1H, *J* = 4 Hz, H_3); 3.80 (dt, 1H, *J* = 3.2 and 9.6 Hz, H_5); 3.70 (dd, 1H, J_{4-5} = 3.2, J_{4-3} = 9.6 Hz, H_4); 3.62 (d, 2H, *J* J_{6-5} = 3.6 Hz, $H_{6a,b}$); 3.28 (s, 3H, OCH_3); 1.65 (s,3H, CH_3). The product presents the same physical-chemical characteristics reported in literature.[100]

3-(Benzyloxycarbonylamino)-1-propanol (25)

To a solution of 3-amino-1-propanol (30 g 399.4 mmol) and NaOH 1M (400 ml) at 0°C, benxylchloroformiate (67 ml, 439.4 mmol). After 10 minutes the reaction was kept at rt and

under stirring for 1 h. The solution changed from colourless to white and then CH_2Cl_2 was added and the mixture was stirred for other 3 h, monitoring via TLC (eluent: hexane/EtOAc4:6). The organic phase was extracted, dried over $NaSO_4$, filtered off and evaporated under vacuum. The crude was purified via flash chromatography (eluent: hexane/EtOAc4:6) and the product was obtained as a crystalline white solid (yield= 70%). **^{1}H-NMR** (400 MHz, MeOD) δ (ppm): 7.34-7.28 (m, 5H, Ar*H*); 5.06 (s, 2H, C*H*$_2$Ph); 3.58 (t, 2H, *J* = 6 Hz, CH_2CH_2NH); 3.20 (t, 2H, *J* = 6 Hz, OCH_2CH_2); 1.70 (q, 2H, *J* = 6 Hz, $CH_2CH_2CH_2$).
The product presents the same physical-chemical characteristics reported in literature.[101,102]

N-(Benzyloxycarbonyl)aminopropyl-3,4,6-tri-O-benzyl-β-D-glucopyranoside (26)

Trimethylsilyl triflate (648 µl, 3.58 mmol) was added to a solution of **24** (20.0 g, 95.5 mmol) and benzyl N-(3-hydroxypropyl)carbamate **25** (20.0 g, 95.5 mmol) in CH_2Cl_2 (50 mL), cooled at 0°C. The reaction was monitored by TLC (hexane/EtOAc 7:3) and after 2.5 h the mixture was neutralised with Et_3N and diluted with CH_2Cl_2. The organic layer was washed with H_2O and dried over Na_2SO_4. Removal of the solvent afforded a mixture of 2-hydroxy and 2-O-acetylated derivatives; this was dissolved in methanol (20 mL) and solid MeONa was added. Monitoring of the reaction by TLC (hexane/EtOAc 7:3) indicated complete formation of **26** in 4 h. To neutralize the solution Amberlite IR-120 resin (H^+) was added and after the removal of the resin on paper filter, the crude was purified by silica gel chromatography (hexane/AcOEt 6:4) to give **26** as a yellow oil (63%). **^{1}H-NMR** (300 MHz, $CDCl_3$) δ (ppm): 7.37-7.11 (m, 10H, Ar*H*); 5.34 (br s, 1H, N*H*Z); 5.07 (s, 2H, C*H*$_2$Z); 4.92 (d, 1H, *J* = 11.1 Hz, CH*H*Ph); 4.83 (d, 1H, *J* = 8.7 Hz, CH*H*Ph); 4.80 (d, 1H, *J* = 8.1 Hz, CH*H*Ph); 4.55-4.46 (m, 3H, 3xCH*H*Ph); 4.22 (d, 1H,$J_{1\text{-}2}$= 7.2 Hz, H_1); 3.94-3.89 (m, 1H,OC*H*HCH_2); 3.70-3.59 (m, 3H, $H_{6a,b}$, OCH*H*CH_2); 3.58-3.47 (m, 5H, H_2, H_3, H_4, H_5 and CH_2CHHNH); 3.29-3.18 (m, 1H, CH_2CH*H*NH); 2.78 (s, 1H, O*H*); 1.81-1.74 (m, 2H, $CH_2CH_2CH_2$). The spectroscopic data found are in agreement with those reported in literature.[81]

N-(Benzyloxycarbonyl)aminopropyl 3,4,6-tri-O-benzyl-β-D-mannopyranoside (27)

In a two round bottom flask, to a solution of **27** (3.13 g, 4.89 mmol), pyridine (1.60 ml, 19.58 mmol) was added under argon in dry CH_2Cl_2 (25 ml). The mixture was cooled to 0°C and after 15 min triflic anhydride (2.50 ml, 14.67 mmol) was added dropwise. The reaction was monitored by TLC (hexane/EtOAc 6:4). After 30 min, the reaction mixture was diluted with

CH_2Cl_2 (10 ml) and the organic layer was washed with aq. HCl solution (5%) and then with satd. aq. $NaHCO_3$ solution until neutralization, dried over Na_2SO_4 and concentrated under reduced pressure. The 2-O-triflate intermediate was dissolved under nitrogen in dry toluene (70 ml) and then NBu_4N_3 (3.18 g, 11.17 mmol) was quickly added. The solution was stirred at 55°C for 24 h and monitored by TLC (hexane/EtOAc 6:4). When finished, the reaction mixture was concentrated and the crude residue was purified by silica gel chromatography (hexane/ EtOAc 6:4) to give compound **27** as a yellow oil (55%). **^{1}H-NMR** (300 MHz, $CDCl_3$) δ (ppm): 7.36-7.11 (m, 20H, Ar*H*); 5.44 (br s, 1H, N*H*Z); 5.09 (d, 1H, *J* = 12.3 Hz, CH*H*NHZ); 5.03 (d, 1H, *J* = 12.3 Hz, C*H*HNHZ); 4.81 (d, 1H, *J* = 10.8 Hz, CH*H*Ph); 4.72 (d, 1H, *J* = 12.0 Hz, CH*H*Ph); 4.65 (d, 1H, *J* = 11.7 Hz, CH*H*Ph); 4.53 (d, 1H, *J* = 12.0 Hz, CH*H*Ph); 4.49 (d, 1H, *J* = 12.1 Hz, CH*H*Ph); 4.45 (d, 1H, *J* = 12.0 Hz, CH*H*Ph); 4.41 (d, 1H, $J_{1\text{-}2}$ = 5.2 Hz, H_1); 3.93-3.83 (m, 2H, H_2 and H_{7a}); 3.74-3.56 (m, 5H, H_3, H_4, $H_{6a,b}$, H_{7b} and CH_2CH*H*O); 3.42-3.24 (m, 3H, H_5 and $H_{9a,b}$); 1.81-1.72 (m, 2H, $H_{8a,b}$). The spectroscopic data found are in agreement with those reported in literature.[81]

N-(Benzyloxycarbonyl)aminopropyl 2-acetamido-3,4,6-tri-O-benzyl-2-deoxy-β-D-mannopyranoside (28)

In a two round bottom flask, under nitrogen, **27** (1.32 g, 1.97 mmol) was dissolved in dry MeOH. $NiCl_2$x $6H_2O$ (47 mg, 0.20 mmol) and after $NaBH_4$ (0.15 g, 3.95 mmol) portionwise were added. The reaction became brown and was left at 0°C for 30 minutes, checked via TLC (eluent: hexane/acetone 1:1).the crude was filtered over a Celite pad (washing with MeOH) and the solvent was removed under vacuum The crude amine was dissolved in MeOH (12 ml), acetic anhydride (1.08 ml, 32.07 mmol) was added, and the solution was stirred at rt for 1 h and then concentrated to dryness. The residue was purified by silica gel chromatography (hexane/ acetone 1:1) to afford **28** as a yellowish oil (80%). **^{1}H-NMR** (400 MHz, $CDCl_3$) δ (ppm): 7.33-7.15 (m, 20H, Ar*H*); 5.69 (d, 1H, *J* = 9.8 Hz, N*H*Ac); 5.43 (br s, 1H, N*H*Z); 5.08 (s, 2H, CH_2Z); 4.87-4.80 (m, 3H, CH_2Ph and H_2); 4.55-4.43 (m, 5H, 2xCH_2Ph and H_1); 3.76-3.85 (m, 2H, OC*H*HCH_2); 3.69-3.45 (m, 5H, H_3, H_4, H_5 and $H_{6a,b}$); 3.44-3.33 (m, 1H, OCH*H*CH_2); 3.35-3.20 (m, 2H, CH_2CH_2NH); 1.97 (s, 3H, CH_3CO); 1.78-1.71 (m, 2H, CH_2C$H_2$$CH_2$). The spectroscopic data found are in agreement with those reported in literature.[81]

Aminopropyl 2-acetamido-3,4,6-tri-O-benzyl-2-deoxy-β-D-mannopyranoside (29)

28 (0.54 g, 0.79 mmol) was dissolved in EtOAc/EtOH 1:9 (10 ml) and then Pd/C (10%) was added in catalytic amount. The mixture was reacted at rt at the Parr apparatus at 1.5 bar of H_2 at rt for 90 minutes. The reaction was controlled by ESI-MS and TLC (eluent: hexane/acetone 1:1 and DCM/MeOH/Et_3N 9:1:1%). The catalyst was filtered off and the solvent removed under reduce pressure to give compound **29** as a yellow-orange oil (86%).

^{1}H-NMR (300 MHz, $CDCl_3$) δ (ppm): 7.37-7.16 (m, 15H, Ar*H*); 6.20 (d, 1H, *J* = 9.9 Hz, N*H*Ac); 4.88-4.82 (m, 4H, H_1, H_2 and 2xCH*H*Ph); 4.59-4.43 (m, 5H, H_4 and 4xC*H*HPh); 3.91-3.85 (m, 1H, OC*H*HCH$_2$); 3.77-3.49 (m, 4H, H_3, H_5, $H_{6a,b}$); 3.44-3.41 (m, 1H, OCH*H*CH$_2$); 2.78-2.73 (m, 2H, CH$_2$C*H*$_2$NH$_2$); 2.03 (s, 3H, C*H*$_3$CO); 1.74-1.70 (m, 2H, CH$_2$C*H*$_2$CH$_2$). **^{13}C-NMR**: (75 MHz, $CDCl_3$) δ (ppm): 171.9 (*C*OCH$_3$); 138.2, 137.9, 137.7 (Ar); 128.4, 128.33, 128.31, 127.9, 127.85, 127.8, 127.7 (Ar*C*H); 99.5 (C_1); 75.0 (C_4); 74.8, 74.1 (*C*H$_2$Ph); 73.3 (C_3); 71.0 (*C*H$_2$Ph); 68.9 (C_6); 67.6 (C_7); 49.2 (C_2); 38.5 (C_9); 27.9 (C_8); 23.6 (CO*C*H$_3$).

ESI-MS (+): m/z 549 [100%; $(M+H)^+$].

General procedure for the conjugation reaction (derivatives 30 and 31)

Isothiocyanate calix[4]arene (1 eq) was dissolved in CH_2Cl_2 dry. NEt_3 (30 eq) and sugar **29** (5 eq) were added and the reaction was stirred at rt for 24-48 h. When the reaction was completed, checked via ESI-MS and TLC (eluent: DCM/MeOH 20:1), the crude was then evaporated under reduced pressure and purified by flash chromatography to afford pure perbenzylated glycocalix[4]arenes.

***Cone*-5,11,17,23-tetrakis[(aminopropyl-2-acetamido-3,4,6-Tri-*O*-benzyl-2-deoxy-β-D-mannopiranosyl)thioureido]-25,26,27,28-tetrapropoxycalix[4]arene (30)**

The pure glycocalix[4]arene was obtained after purification on silica gel flash column (eluent: DCM/MeOH 20:1) in 50% yield as a white solid. **^{1}H-NMR** (400 MHz, $CDCl_3$): δ(ppm) 8.57 (br s, 4H, N*H*CS); 7.35-7.24 (m, 60H, Ar benzyl); 7.13 (s, 8H, Ar*H*); 6.60 (br s, 4H, N*HCS*); 6.21 (br s, 4H, N*H*Ac); 4.92 (d, 4H, J=4.0 Hz, H_2); 4.85 (d, 4H, J = 10.8 Hz, C*H*HPh); 4.78 (d, 4H, J = 11.2 Hz, CH*H*Ph);4.57 (d, 4H, J = 12.0 Hz, C*H*HPh); 4.48-4.42 (m, 16H, H_1, C*H*$_2$Ph, CH*H*Ph); 4.36 (d, 4H, J = 13.2 Hz, ArC*H*Ar eq); 3.97-3.54 (m, 44H, H_3, H_4, H_5, $H_{6a,b}$, OC*H*$_2$CH$_2$CH$_3$, OC*H*$_2$CH$_2$CH$_2$NH, NHC*H*$_2$); 3.06 (d, 4H, J = 13.6 Hz, ArC*H*$_2$Ar ax); 2.06 (s, 12H, NHCOC*H*$_3$); 1.89-1.83 (m, 16H, OCH$_2$C*H*$_2$, CH$_2$C*H*$_2$CH$_2$); 0.95 (br t, 12 H, OCH$_2$CH$_2$C*H*$_3$). **^{13}C-NMR** (100 MHz,

$CDCl_3$): δ(ppm) 180.5(*C*S); 172.1 (*C*OCH_3); 154.4 (Ar ipso); 138.2, 137.9, 137.7, 132.3 (*C*Ar); 128.5, 128.3, 128.2, 127.8, 127.7, 123.9 (Ar*C*H); 99.9 (C_1); 79.9 (C_5); 76.9-71.0 (3x*C*H_2Ph, C_3, C_4, C_6); 68.9 (C_7, O*C*$H_2CH_2CH_3$) 49.7 (C_2); 43.4 (C_9); 29.7 (Ar*C*H_2Ar); 28.9 (C_8); 23.7 (*C*H_3CO); 23.2 (CH_2*C*H_2CH_3); 10.2 (CH_2CH_2*C*H_3). **ESI-MS(+)**:m/z 1531 $[M+2Na]^{2+}$

***1,3-alternate*-5,11,17,23-tetrakis[(aminopropil-2-acetamido-3,4,6-Tri-*O*-benzyl-2-deoxy-β-D-mannopiranosyl)thioureido]-25,26,27,28-tetrapropoxycalix[4]arene (31)**

The pure product was afforded by purification on silica gel column (eluent: DCM/MeOH 20:1) in 91% yield as a white solid. **^{1}H-NMR** (400 MHz, $CDCl_3$): δ(ppm) 8.71 (s, 4H, N*H*CS); 7.35-7.23 (m, 60H, Ar benzyl); 7.16 (s, 8H, Ar*H*); 6.86 (s, 4H, N*HCS*); 5.86 (d, 4H, J=9.9 Hz, N*H*Ac); 5.12 (d, 4H, J=4.0 Hz, H_2); 4.90-4.87 (m, 4H, H_3); 4.85-4.60 (m, 24H, OC*H*$_2$Ph); 4.50-4.47(m, 4H, H_1); 4.01-3.41 (m, 48H,ArC*H*$_2$Ar, $H_{6a,b}$, H_5, H_4, OC*H*$_2CH_2CH_3$, OC*H*$_2CH_2CH_2$NH, NHC*H*$_2$); 2.12 (s, 12H, NHCOC*H*$_3$); 1.93-1.90 (m, 8H, CH_2C*H*$_2CH_2$); 1.83-1.79 (m, 8H, OCH_2C*H*$_2$); 0.94 (t, 12H, J=8.0 Hz, OCH_2CH_2C*H*$_3$). **^{13}C-NMR** (100 MHz, $CDCl_3$): δ(ppm) 181.0(*C*S); 171.8 (*C*OCH_3); 153.7 (Ar ipso); 138.2, 137.9 ,137.8, 134.1, 132.6 (*C*Ar); 128.5 128.3, 128.2, 128.0, 127.9, 127.8, 127.7, 126.5 (Ar*C*H); 99.9 (C_1); 80.5 (C_5); 77.0-68.6 (3x*C*H_2Ph, C_3, C_4, C_6, C_7, O*C*$H_2CH_2CH_3$); 49.7 (C_2); 43.4 (C_9); 36.1 (Ar*C*H_2Ar); 29.0 (C_8); 23.7 (*C*H_3CO); 23.6 (CH_2*C*H_2CH_3); 10.3 (CH_2CH_2*C*H_3).**ESI-MS(+)**:m/z 1531 $[M+2Na]^{2+}$

M.p: n.a.

Aminopropyl 2-acetamido-2-deoxy-β-D-mannopyranoside (32)

Hydrogenolysis: 28 (95 mg, 0.14 mmol) was dissolved in MeOH/H_2O 1:1 (5 ml) and then Pd/C (10%) was added in catalytic amount. The mixture was reacted at rt at the Parr apparatus at 1.5 bar of H_2 at rt for 72h (adding a second portion of catalyst after 24h). The reaction was controlled by ESI-MS and TLC (eluent: hexane/acetone 1:1 and EtOAc/i-PrOH/MeOH/ H 7:3:0.5) and when it was completed the catalyst was filtered off. The solvent was removed under reduced pressure to afford compound **32** as a white foam in quantitative yield.

Hydrogen transfer:[103] **28** was dissolved in MeOH/H_2O 1:1 (5 ml) and then Pd/C (10%) was added in catalytic amount. NH_4COOH was added to the solution and the reaction was heated at 60°C for 1-3 h under MW irradiation (150 W). After checking the complete removal of

benzyl groups via TLC (eluent: hexane/acetone 1:1 and EtOAc/i-PrOH/MeOH/ H 7:3:0.5), the catalyst was filtered off and the solvent was eliminated under reduced pressure to lead compound **32** in 84% yield.

^{1}H-NMR (300 MHz, D_2O): δ(ppm) 4.71 (s, 1H, H_1); 4.42 (d, 1H, $J_{1\text{-}2}$ = Hz, H_2); 3.99-3.82 (m, 3H, H_3, H_{6a}, H_{7a}); 3.77-3.70 (m, 2H, H_{6b}, H_{7b}); 3.46 (t, 1H, J = 9.6 Hz, H_4); 3.35 (m, 1H, H_5); 3.05 (m, 2H, $CH_2CH_2CH_2NH_2$); 2.01 (s, 3H, $COCH_3$); 1.88(m,2H, $CH_2CH_2CH_2$). The spectroscopic data found are in agreement with those reported in literature.[81]

N-(Boc)aminopropyl 2-acetamido-3,4,6-tri-O-benzyl-2-deoxy-β-D-mannopyranoside (33)

Compound **29** (0.34 g, 0.63 mmol) was dissolved in MeOH (8 ml), then Et_3N (348 µl, 2.5 mmol) and Boc_2O (0.68 g, 3.13 mmol) were subsequently added. The reaction was checked via TLC (DCM/MeOH/Et_3N 9:0.5:0.1) and when it was finished after 2h, the mixture was evaporated under vacuum. Many methanol washings were done in order to remove the residual Boc_2O. The pure product was obtained as a pale oil in quantitative yield. **^{1}H-NMR** (400 MHz, $CDCl_3$) δ (ppm): 7.25-7.03 (m, 15H, Ar*H*); 6.18 (d, 1H, *J* = 6.0 Hz, N*H*Ac); 5.26 (br s, 1H, N*H*Boc); 4.71 (m, 3H, H_2 and CH_2Ph); 4.46-4.32 (m, 5H, H_1 and 2xCH_2Ph); 3,72-3.67 (m, 1H, H_{7a}); 3.60 (m, 3H, H_5, $H_{6a,b}$); 3.54-3.49 (m, 1H, H_3); 3.47-3.40 (m, 1H, H_{7b}); 3.34-3.31 (m, 1H, H_4); 3.11-3.05 (m, 2H, CH_2CH_2NHBoc); 1.80 (s, 3H, $COCH_3$); 1.61-1.58 (m, 2H, $CH_2CH_2CH_2$); 1.30 (s, 9H, $C(CH_3)_3$). **^{13}C-NMR** (100 MHz, $CDCl_3$) δ (ppm): 170.8 ($COCH_3$); 156.1 ($CO(CH_3)_3$); 137.8 (*C*Ar); 128.4, 128.3, 127.8, 127.7, 127.6 (Ar*C*H); 99.5 (C_1); 85.0 ($C(CH_3)$; 80.3 (C_5); 74.9 (C_4); 74.7 (CH_2Ph); 74.0 (C_3); 73.3 (CH_2Ph); 71.0 (CH_2Ph); 68.8 (C_6); 67.1 (C_7); 49.2 (C_2); 37.6 (C_9); 29.7 (C_8); 27.2 ($C(CH_3)_3$); 23.3 ($COCH_3$). **ESI-MS (+):** m/z 671 [100%; $(M+Na)^+$]; m/z 571 [60%; $(M+Na\text{-}Boc)^+$].

N-(Boc)aminopropyl-2-acetamido-2-deoxy-β-D-mannopyranoside (34)

In a MW tube compound **33** (0.60 g, 0.93 mmol) was dissolved in MeOH/H_2O 1:1 (5 ml) and then Pd/C (10%) was added in catalytic amount. NH_4COOH (0.23 g, 3.7 mmol) was added and the mixture was heated at 60°C by MW irradiation (150 W) for a time between 1 and 3 h, monitoring the reaction by ESI-MS and TLC (EtOAc/MeOH 8:2). When it was completed the catalyst was removed by filtration on paper filter and the solvent was removed under reduced pressure to afford the fully O-deprotected sugar **34** in 84% yield as a white foam, without purification steps. **^{1}H-NMR** (400 MHz, MeOD) δ (ppm): 7.84 (d, 1H, *J* = 9.2 Hz,

N*H*Ac); 4.65 (s, 1H, H_1); 4.47 (d, 1H, J = 3.2 Hz, H_2); 3.88-3.87 (m, 3H, H6a$_{,b}$, H_{7a}); 3.68 (dd, 1H, $J_{2\text{-}3}$ = 4, $J_{3\text{-}4}$ = 9.6 Hz, H_3); 3.60-3.58 (m, 1H, H_{7b}); 3.53 (t, 1H, *J* = 9.6 Hz, H_4); 3.28-3.22 (m, 1H, H_5); 3.13 (t, 2H, *J* = 6.4 Hz, CH_2CH_2NHBoc); 2.05 (s, 3H, $COCH_3$); 1.73 (t, 2H, *J* = 6.0 Hz, $CH_2CH_2CH_2$); 1.45 (s, 9H, $C(CH_3)_3$). **^{13}C-NMR** (100 MHz, $CDCl_3$) δ (ppm): 173.4 ($COCH_3$); 157.1 ($CO(CH_3)_3$); 99.4 (C_1); 78.5 ($C(CH_3)$; 76.9 (C_5); 73.0 (C_3); 66.9 (C_7); 66.3 (C_4); 60.5 (C_6); 53.5 (C_2); 36.9 (C_9); 29.5 (C_8); 27.4 ($C(CH_3)_3$); 21.4 ($COCH_3$). **ESI-MS (+)**: m/z 401 [100%; $(M+Na)^+$]; m/z 301 [100%; $(M+Na\text{-}Boc)^+$].

N-(Boc)aminopropyl 2-acetamido-3,4,6-tri-O-acetyl-2-deoxy-β-D-mannopyranoside (35)

Pyridine (6 ml) was added to **34** (0.18 g, 0.47 mmol) and the acetic anhydride was added (740 µl, 7.91 mmol). The reaction was stirred at room temperature for 1 h and monitored by ESI-MS and TLC (eluent: hexane/EtOAc 2:8). The mixture was then evaporated under vacuum to give a yellow oil. It was purified by a silica gel pad (eluent: hexane/EtOAc 2:8) because in the further step there was not allowed the purification using this stationary phase. Pure **35** was achieved in quantitative yield as a white foam. **^{1}H-NMR** (300 MHz, $CDCl_3$) δ (ppm): 6.05 (d, 1H, *J* = 9.0 Hz, N*H*Ac); 5.02 (t, 1H, *J* = 9.0 Hz, H_4); 4.92 (dd, 1H, $J_{2\text{-}3}$ = 6.0 and $J_{3\text{-}4}$ =12.0 Hz, H_3); 4.82 (t, 1H, *J* = 6.0 Hz, N*H*Boc); 4.65-4.62 (m, 2H, H_1, H_2); 4.22 (dd, 1H, $J_{5\text{-}6a}$ = 6.0, $J_{6a\text{-}6b}$ = 15.0 Hz, H_{6a}); 4.07 (dd, 1H, $J_{5\text{-}6b}$ = 6.0, $J_{6a\text{-}6b}$ = 15.0 Hz, H_{6b}); 4.01-3.99 (m, 1H, H_{7a}); 3.72-3.62 (m, 1H, H_5); 3.55-3.48 (m, 1H, H_{7b}); 3.13 (d, 2H, *J* = 6.0 Hz, CH_2CH_2NHBoc); 2.09 (s, 3H, OAc); 2.06 (s, 3H, OAc); 2.04 (s, 3H, OAc); 2.01 (s, 3H, $NHCOCH_3$); 1.70 (q, 2H, *J* = 6.0 Hz, $CH_2CH_2CH_2$); 1.42 (s, 9H, $C(CH_3)_3$). **^{13}C-NMR** (75 MHz, $CDCl_3$) δ (ppm): 171.4 ($NHCOCH_3$); 170.5, 170.2, 169.7 ($COCH_3$); 156.1 ($CO(CH_3)_3$); 98.6 (C_1); 78.9 ($C(CH_3)$; 72.3 (C_5); 71.3 (C_3); 67.0 (C_7); 66.2 (C_4); 62.6 (C_6); 49.9 (C_2); 37.3 (C_9); 29.5 (C_8); 28.3 ($C(CH_3)_3$); 22.9-20.6 ($COCH_3$). **ESI-MS (+)**: m/z 527 [100%; $(M+Na)^+$]; m/z 427 [100%; $(M+Na\text{-}Boc)^+$].

Aminopropyl-2-acetamido-3,4,6-tri-O-acetyl-2-deoxy-β-D-mannopyranoside (19)

The Boc-amino protected derivative **35** (0.24 g, 0.47 mmol) was dissolved in dry DCM (15 ml) and then trifluoroacetic acid(1.27 ml, 16.45 mmol) was added dropwise. The reaction was monitored by TLC (eluent: DCM/MeOH 85:15) and was quenched after 1 h under stirring. Since the pH was acid, Et_3N was added to neutralize the solution. The solvent was evaporated under reduced pressure and washings with MeOH were done to easily eliminate

Et_3N. The amino sugar **19** was obtained as a yellow-orange oil in quantitative yield, without further purifications. **^{1}H-NMR** (300 MHz, $CDCl_3$) δ (ppm): 6.85 (d, 1H, *J* = 13.5 Hz, N*H*Ac); 5.02 (t, 1H, *J* = 9.6 Hz, H_4); 4.88 (dd, 1H, $J_{2\text{-}3}$ = 4.2, $J_{3\text{-}4}$ = 9.6 Hz, H_3); 4.68-4.64 (m, 2H, H_1, H_2); 4.14 (dd, 1H, $J_{6a\text{-}5}$ = 6.0, $J_{6a\text{-}6b}$ = 12.4 Hz, H_{6a}); 4.03 (dd, 1H, $J_{5\text{-}6b}$ = 2.4, $J_{6a\text{-}6b}$ = 12.0 Hz, H_{6b}); 3.90-3.82 (m, 1H, H_{7a}); 3.62-3.59 (m, 2H, H_5, H_{7b}); 3.16-3.07 (m, 2H, $CH_2CH_2NH_2$); 1.98 (s, 3H, OAc); 1.95 (s, 3H, OAc); 1.94 (s, 3H, OAc); 1.93 (s, 3H, NHCOCH_3); 1.92-1.80 (m, 2H, $CH_2CH_2CH_2$). **^{13}C-NMR** (75 MHz, $CDCl_3$) δ (ppm): 161.2 (*C*OCH_3); 98.2 (C_1); 72.3 (C_5); 71.3 (C_3); 66.6 (C_7); 66.0 (C_4); 62.4 (C_6); 49.5 (C_2); 37.4 (C_9); 26.3 (C_8); 22.4-20.4 (CO*C*H_3). **ESI-MS (+)**: m/z 427 [100%; $(M+Na)^+$]; m/z 405 [30%; $(M+H)^+$].

General procedure for the conjugation reaction between acetylated glycosylamino 19 and isothiocyanate calix[n]arene derivatives 10a-d[88,95] and the monomer 38

In a two neck round bottom flask isothiocyanate calix[4]- or calix[6]arene (1 eq) was dissolved in 3 ml of dry CH_2Cl_2 under N_2 atmosphere. Then for each isothiocyanate group of the calixarene 1.25 equivalents of glycosyl amino **19** 1 equivalent of Et_3N were added. The mixture was allowed to react at room temperature for 24-48 h (with addition of half equivalents of sugar and base after 1 night). When it was completed (checked by ESI-MS and TLC), the solvent was removed under reduced pressure. The crude was in all the cases purified by flask column chromatography on silica gel.

5,11,17,23-tetrakis[(aminopropil-2-acetamido-3,4,6-Tri-*O*-acetyl-2-deoxy-β-D-mannopiranosyl)thioureido]-25,26,27,28-tetramethoxycalix[4]arene (36)

The pure compound was obtained with purification via flask column chromatography (eluent: DCM/MeOH 20:1) as a brownish solid. (51%). **^{1}H-NMR** (300 MHz, $CDCl_3$) δ (ppm): 8.91 (br s, 4H, N*H*CS); 8.16 (br s, 4H, CSN*H*CH_2); 7.44 (s, 4H, Ar*H*); 7.02 (s, 4H, Ar*H*); 5.76 (d, 4H, *J* = 8.9 Hz, N*H*Ac); 5.67 (dd, 4H, $J_{2\text{-}3}$ = 3.9 , $J_{3\text{-}4}$ = 9.9 Hz, H_3); 5.19 (s, 4H, H_1); 5.08 (t, 4H, *J* = 10.0 Hz, H_4); 4.92-4.88 (m, 4H, H_2); 4.27-4.25 (m, 8H, H_{6a} and ArCH*H*Ar eq); 4.03-3.90 (m, 4H, H_{6b}); 3.79-3.77 (m, 12H, H_5, H_{7a}, H_{9a}); 3.76-3.75 (m, 4H, H_{7b}); 3.39 (s, 12H, OCH_3); 3.27-3.25 (m, 4H, ArC*H*HAr ax); 2.94-2.90 (m, 4H, H_{9b}); 2.36-2.39 (m, 4H, H_{8a}); 2.08 (s, 12H, 3xOAc); 2.03 (s, 12H, 3xOAc); 1.98 (s, 12H, 3xOAc); 1.96 (s, 12H, 3xNHCOCH_3); 1.45-1.42 (m, 4H, H_{8b}). **^{13}C-NMR** (75 MHz, $CDCl_3$) δ (ppm): 181.7 (*C*S); 171.7, 171,5, 170.6, 169.9 (*C*OCH_3); 154.2 (Ar ipso); 135.2 (Ar ortho); 134.0 (Ar para); 127.5 (Ar*C*H); 99.7 (C_1); 76.6 (C_3); 73.2 (C_7);

67.4 (Ar*C*H_2Ar); 66.1 (C_4); 62.6 (C_6); 58.3 (O*C*H_3); 51.2 (C_2); 42.4 (C_9); 35.1 (C_5); 29.5 (C_8); 23.4-20.8 (CO*C*H_3). **ESI-MS (+):** m/z 1186 [100%; $(M+2Na)^{2+}$]. **M.p.:** 136.6-138.2

***Cone*-5,11,17,23-tetrakis[(aminopropil-2-acetamido-3,4,6-Tri-*O*-acetyl-2-deoxy-β-D-mannopiranosyl)thioureido]-25,26,27,28-tetrapropoxycalix[4]arene (37)**

The crude was purified via flash column chromatography (eluent: DCM/MeOH 20:1) giving cone derivative in 37% yield as a white solid. **^{1}H-NMR** (400 MHz, $CDCl_3$) δ (ppm): 8.36 (br s, 4H, N*H*CS); 6.65 (br s, 8H, Ar*H*), 6.38 (br s, 8H, CSN*H*CH_2, N*H*Ac); 5.15-5.09 (m, 8H, H_3, H_4); 4.78 (d, 4H, $J_{2\text{-}1}$ = 9.2 Hz, H_2); 4.74 (s, 4H, H_1); 4.41 (d, 4H, *J* = 13.2 Hz, ArC*H*HAr ax); 4.27 (dd, 4H, $J_{6a\text{-}5}$ = 5.6, $J_{6a\text{-}6b}$ = 12.0 Hz, H_{6a}); 4.13 (dd, 4H, $J_{5\text{-}6b}$ = 2.4, $J_{6a\text{-}6b}$ = 12.0 Hz, H_{6b}); 3.97-3.96 (m, 4H, H_{7a}); 3.84 (br s, 12H, OC$H_2$$CH_2$, H_{9b}); 3.70-3.69 (m, 4 H, H_5); 3.57 (br s, 4 H, H_{7b}); 3.49-3.48 (m, 4H, H_{9a}); 3.14 (d, 4H, *J* = 13.2 Hz ArC*H*HAr eq); 2.09 (s, 12H, 3xOAc); 2.07 (s, 12H, 3xOAc); 2.06 (s, 12H, 3xOAc); 2.05 (s, 12H, 3xOAc); 2.03 (s, 12H, 3xOAc); 1.96-1.90 (m, 16H, $ArCH_2CH_2CH_3$, $CH_2CH_2CH_2$). **^{13}C-NMR** (100 MHz, $CDCl_3$) δ (ppm): 180.4 (*C*S); 172.1 (NH*C*OCH_3), 170.7, 170.5, 169.8 (3X *C*OCH_3); 154.5 (Ar ipso); 136.2 (Ar ortho); 132.1 (Ar para); 124.8 (Ar*C*H); 99.1 (C_1); 72.4 (C_5); 71.4 (C_4); 68.7 (C_7); 66. (C_3); 62.5 (C_6); 50.3 (C_2); 43.0 (C_9); 30.9 (Ar*C*H_2Ar); 29.9 (C_8); 28.9 (ArOCH_2*C*H_2)23.4-20.7 (CO*C*H_3), 10.2 (ArOCH_2CH_2*C*H_3). **ESI-MS (+):** m/z 1242.7 [100%; $(M+2Na)^{2+}$]. **M.p.:** > 130°C dec

***1,3-alternate*-5,11,17,23-tetrakis[(aminopropil-2-acetamido-3,4,6-Tri-*O*-acetyl-2-deoxy-β-D-mannopiranosyl)thioureido]-25,26,27,28-tetrapropoxycalix[4]arene (38)**

The pure product was obtained after purification via flash column chromatography (eluent: DCM(MeOH 20:1) as white solid (37% yield). **^{1}H-NMR** (400 MHz, $CDCl_3$) δ (ppm): 8.97 (br s, 4H, N*H*CS); 8.19 (br s, 4H, CSN*H*CH_2); 7.37 (s, 4H, Ar*H*); 7.05 (s, 4H, Ar*H*); 5.80 (d, 8H, *J* = 8.88 Hz, N*H*Ac, H_3); 5.29 (d, 4H, $J_{1\text{-}2}$ = 5.6 Hz, H_1); 5.11 (t, 4H, *J* = 10.0 Hz, H_4); 4.97-4.88 (m, 4H, H_2); 4.31 (dd, 4H, $J_{6a\text{-}5}$ = 5.2, $J_{6a\text{-}6b}$ = 12 Hz, H_{6a}); 4.81-3.83 (m, 12H, H_{6b}, ArC*H*HAr, H_{9a}); 3.87-3.73 (m, 8H, H_5, H_{7a}); 3.67-3.40 (m, 16H, ArCH*H*Ar, H_{9b}, ArOCH_2); 2.94 (s, 4H, H_{7b}); 2.48 (s, 4H, H_{8a}); 2.11 (s, 12H, 3xOAc); 2.08 (s, 12H, 3xOAc); 2.00 (s, 12H, 3xOAc); 1.99 (s, 12H, $ArCH_2CH_2CH_3$); 1.84 (s, 12H, 3xNHCOCH_3); 1.69-1.45 (m, 12H, $OCH_2CH_2CH_3$, H_{8b}). **^{13}C-NMR** (100 MHz, $CDCl_3$) δ (ppm): 181.9 (*C*S); 171.4, 170.6, 170.5, 169.9 (3x*C*OCH_3, NH*C*OCH_3); 153.6 (Ar ipso); 136.5(Ar otho); 133.1 (Ar para); 128.6,128.4 (Ar meta); 99.9 (C_1), 74.9 (ArO*C*H_2);

72.9 (C_3); 72.3 (C_5); 67.7 (C_7); 66.1 (C_4); 62.3 (C_6), 50.9 (C_2); 42.7 (C_9); 38.2 (2xAr*C*H_2Ar); 36.6 (2xAr*C*H_2Ar); 29.4 (C_8); 23.6 (ArOCH_2*C*H_2); 23.4-20.8 (CO*C*H_3); 8.9 (ArOCH_2CH_2*C*H_3). **ESI-MS (+)**: m/z 1219.4 [20%; $(M+2H)^{2+}$], m/z 1242.6 [20%; $(M+2Na)^{2+}$]. **M.p.:** 145.7-147.3 °C

5,11,17,23,29,35-hexakis[(aminopropil-2-acetamido-3,4,6-Tri-*O*-acetyl-2-deoxy-β-D-mannopiranosyl)thioureido]-37,38,39,40,41,42-hexamethoxycalix[6]arene (39)

The product was obtained with purification via flash column chromatography (eluent DCM/MeOH 96:4), followed by preparative layer on silica gel (eluent DCM/MeOH 94:6, run twice). The pure glycocalix[6]arene was afforded as a yellowish solid (25% yield). **^{1}H-NMR** (400 MHz, $CDCl_3$) δ (ppm): 8.66 (br s, 4H, N*H*CS); 6.92 (br s, 6H, Ar*H*); 7.37 (s, 12H, CSN*H*CH_2, N*H*Ac); 5.31-5.30 (m, 6H, H_4); 5.27-5-25 (m, 6H, H_3); 4.77-4-71 (m, 12H, H_1, H_2); 4.29-4.28 (m, 6H, H_{6a}); 4.14-4.13 (m, 6H, H_{6b}); 4.28-3-56 (m, 60H, H_5, OCH_3, ArC*H*$_2$Ar, OC*H*$_2$$CH_2$, OCH_2CH_2C*H*$_2$); 2.10-2.02 (m, 72h, NHAc, 3xOAc); 1.81 (br s, 12H, OCH_2C*H*$_2$$CH_2$). **^{13}C-NMR** (100 MHz, $CDCl_3$) δ (ppm): 180.6 (*C*S); 172.1 .4, 170.7, 170.3, 169.7 (3x*C*OCH_3, NH*C*OCH_3); 154.2 (Ar ipso), 134.8(br signal, Ar otho, Ar para); 125.3 (Ar meta); 99.9 (C_1);72.5 (C_5); 71.3 (C_3); 68.6 (C_7); 66.1 (C_4); 62.6 (C_6); 61.0 (O*C*H_3); 50.4 (C_2); 42.9 (C_9); 29.7 (Ar*C*H_2Ar); 29.0 (C_8): 23.4 (NH*C*OCH_3); 20.9, 20.8, 20.7 (3x*C*OCH_3).

General procedure for the deacetylation reaction of glycocalix[n]arenes 36-39 and monomer 43

Peracetylated glycoclusters were dissolved in MeOH and freshly prepared CH_3ONa (solid) was added until pH 8-9. The mixture was stirred at 0°C for 1-3 h. The progress of the reaction was followed by ESI-MS and/or TLC. After the complete deacetylation the Amberlite resin IR-120 (H^+) was added for quenching and stirred for other 30 minutes till neutral pH. After the resin removal by filtration on paper filter, the solvent was evaporated under vacuum to give the final product. Purification step was required for all the compound to have the maximum purity for the biological tests.

5,11,17,23-tetrakis[(aminopropil-2-acetamido-2-deoxy-β-D-mannopiranosyl)thioureido]-25,26,27,28-tetramethoxycalix[4]arene (1)

The first purification step was on C18 reverse phase silica gel column (eluent: MeOH/H_2O 5.5:4.5), followed by a size exclusion chromatography on Sephadex G-25 as stationary phase

(eluent: MeOH/H_2O 1.5:8.5). The final purification step was on a flash column in normal phase (eluent; i-PrOH/ H_2O/ Et_3N 8:2:0.5) to afford the mobile calix[4]arene **1** in 40% yield. **^{1}H-NMR** (400 MHz, D_2O, 80 °C) δ (ppm): 6.98 (br s, 8H, Ar*H*); 5.03 (br s, 4H, H_1); 4.43 (br s, 4H, H_2); 3.88-3.22 (m, 48H, H_3, H_4, H_5, H_{6a}, H_{6b}, OC*H*$_2$CH$_2$, CH$_2$C*H*$_2$NHCS, OC*H*$_3$); 1.97 (s, 12H, 4xNHCOC*H*$_3$); 1.80-1.74 (m, 8H, CH$_2$C*H*$_2$CH$_2$). **^{13}C-NMR** (100 MHz, MeOD, 55 °C) δ (ppm): 180.6 (*C*S); 173.6 (NH*C*OCH$_3$); 155.7 (Ar ipso); 135.32 (Ar ortho); 132.4 (Ar para); 124.4 (Ar*C*H); 99.5 (C_1); 77.0 (C_5); 77.0 (C_5); 73.0 (C_3); 67.2 (C_4,C_7); 610 (C_6, OCH_3); 53.6 (C_2); 42.1 (C_9, Ar*C*H$_2$Ar); 29.2 (C_8); 22.2 (NHCO*C*H$_3$).

ESI-MS (+): m/z 1843 [100%; $(M+Na)^+$]; ESI$^-$ m/z 1819 [100%; $(M-H)^-$], m/z 1878 [40%; $(M+Na+Cl)^-$]. **M. p.:** > 140°C dec

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

The pure product **2** was achieved after a purification on C18 reverse phase silica gel column (eluent: MeOH/H_2O 8:2) in 78% yield as a white solid. **^{1}H-NMR** (400 MHz, CD_3OD, 25°C) δ (ppm): 6.71 (br s, 8H Ar*H*); 4.71 (s, 4H, H_1); 4.53 (s, 4H, H_2); 4.47 (d, 4H, *J* = 13.2 Hz, ArCH*H*Ar ax); 3.94-3.86 (m, 44H, $H_{6a,b}$ H_{9a} , ArOC*H*$_2$); 3.71-3.63 (m, 16H, H_3, $H_{7a,b}$, H_{9b}); 3.58 (t, 4H, *J* = 9.6 Hz, H_4); 3.29-3-27 (m, 4H, H_5); 3.18 (d, 4H, *J* = 13.2 Hz, ArCH*H*Ar eq); 2.08 (s, 12H, NHCOC*H*$_3$); 2.02-1.96 (m, 8H, ArO CH$_2$C*H*$_2$); 1.86 (br s, 8H, CH$_2$C*H*$_2$CH$_2$); 1.06 (t, 12H, *J* = 6.8 Hz, CH$_2$CH$_2$*C*H$_3$). **^{13}C-NMR** (100 MHz, CD_3OD, 25°C) δ (ppm): 179 (*C*S); 173.6 (NH*C*OCH$_3$); 154.0 (Aripso); 135.32 (Ar ortho); 132.2 (Ar para); 123.9 (Ar*C*H); 99.5 (C_1); 76.9 (C_5); 76.8 (C_6); 72.8 (C_3); 67.3 (C_7); 66.9 (C_4); 60.7 (ArO*C*H$_2$); 53.6 (C_2); 47.7 (C_9); 42.2 (C_2); 30.5 (Ar*C*H$_2$Ar); 28.7 (C_8); 23.1 (ArOCH$_2$*C*H$_2$); 21.7 (NHCO*C*H$_3$); 9.5 (ArOCH$_2$CH$_2$*C*H$_3$). **ESI-MS (±)**: m/z 1957 [100%; $(M+Na)^+$]; m/z 991 [100%; $(M+2Na)^{2+}$]; m/z 1932 [100% $(M-H)^-$]. **M.p.** : > 178°C dec

***1,3-Alternate*-5,11,17,23-tetrakis[(aminopropil-2-acetamido-2-deoxy-β-D-mannopiranosyl)thioureido]-25,26,27,28-tetrapropoxycalix[4]arene (3)**

Compound **3** was obtained with high purity after purification via C18 reverse phase silica gel chromatography (eluent: MeOH/H_2O 6:4) as a white solid (yield: 83%).

^{1}H-NMR (400 MHz, CD_3OD, 25 °C) δ (ppm): 7.09 (s, 8H, Ar*H*); 4.71 (s, 4H, H_1); 4.56 (d, 4H, *J* = 3.6 Hz, H_2); 4.06-3.94 (m, 4H, H_{9a}); 3.87 (d, 8H, *J* = 3.2 Hz, $H_{6a,b}$); 3.83-3.64 (m, 24H, H_3, $H_{7a,b}$,

H_{9b}, ArOCH_2); 3.63-3.50 (m, 12H, ArCH_2Ar, H_4); 3.40-3.24 (m, 12H, H_5, ArOCH$_2$CH_2); 2.10 (s, 12H, NHCOCH_3); 2.01-1.81 (m, 8H, $H_{8a,b}$); 1.04 (t, *J* = 7.6 Hz, 12H, $CH_2CH_2CH_3$). **^{13}C-NMR** (100 MHz, CD_3OD, 25°C) δ (ppm): 180.7 (*C*S); 173.5 (NH*C*OCH$_3$); 153.9 (Ar ipso); 133.6 (Ar ortho); 131.6 (Ar para); 125.4 (Ar*C*H); 99.52 (C_1), 76.90 (C_5); 74.9 (ArO*C*H$_2$); 72.8 (C_3); 67.2 (C_9); 66.9 (C_4); 60.7 (C_6); 53.6 (C_2), 42.2 (C_7); 35.0 (Ar*C*H$_2$Ar); 28.8 (C_8), 23.6 (ArOCH$_2$*C*H$_2$), 21.6 (NHCO*C*H$_3$), 9.7 (ArOCH$_2$CH$_2$*C*H$_3$). **ESI-MS (+)**: m/z 1957 [100%; $(M+Na)^+$]; m/z 991 [100%; $(M+2Na)^{2+}$]; m/z 1932 [100% $(M-H)^-$]. **M.p.:** > 176°C dec

5,11,17,23,29,35-hexakis[(aminopropil-2-acetamido-2-deoxy-β-D-mannopiranosyl)thioureido]-37,38,39,40,41,42-hexamethoxycalix[6]arene (4)

Purification with via C18 reverse phase silica gel chromatography (eluent: MeOH/H_2O 64:36) afforded to obtain in 54% yield as a off-white solid. **^{1}H-NMR** (400 MHz, D_2O, 80 °C) δ (ppm): 6.95 (s, 6H, Ar*H*); 4.46 (s, 6H, H_1); 4.56 (d, 6H, *J* = 3.9 Hz, H_2); 3.94 (m, 12H, ArC*H*$_2$Ar); 3.88-3.74 (m, 24H, H_3, $H_{6a,b}$, H_{7a}); 3.65-3.61 (m, 6H, H_{7b}); 3.53 (t, 18H, *J* = 8.6 Hz, OC*H*$_3$); 3.35-3-34 (m, 6H, H_5); 2.02 (s, 18H, NHCOC*H*$_3$); 1.81 (t, 12H, *J* = 12.6 , $H_{8a,b}$). **^{13}C-NMR** (100 MHz, D_2O, 80 °C) δ (ppm): 180.7 (*C*S); 175.4 (NH*C*OCH$_3$); 154.8 (Ar ipso); 135.6 (Ar ortho); 134.1 (Ar para); 126.8 (Ar*C*H); 99.9 (C_1), 77.2 (C_5); 75.0 (C_3); 67.7 (C_4, C_7); 61.2 (O*C*H$_3$); 59.9 (C_6); 53.8 (C_2); 42.5 (C_9); 30.7 (C_8); 29.2 (Ar*C*H$_2$Ar); 22.8 (NHCO*C*H$_3$).

ESI-MS (+): m/z 1401.3 [100%; $(M+2K)^+$]; m/z 1383.4 [80%; $(M+2Na)^{2+}$]; m/z 1364.9 [50%; $(M+2H)^+$].

M.p.: 186.2-187.6 °C

1-nitro-4-propoxy-benzene (40)

In a 2 neck round bottom flask under N_2, in p-nitrophenol (0.5 g, 3.59 mmol), iodopropane, (0.7 ml, 7.19 mmol) and K_2CO_3 (0.99 g, 7.19 mmol) were added in dry CH_3CN (30 ml). The mixture was allowed to react, refluxing, for 24 h and monitored by TLC (eluent: hexane/EtOAc 1:1). After the removal of solvent under vacuum, the mixture was dissolved in DCM (50 ml) and washed with HCl 1M (50 ml), H_2O (50 ml) till neutral pH. Then the solvent was evaporated under reduced pressure, giving the pure product without any purification steps, as a yellow oil (93% yield).

^{1}H-NMR (300 MHz, $CDCl_3$): δ (ppm) 8.20-8.15 (m, 2H, ArH); 6.95-6.90 (m, 2H, ArH); 4.00 (t, 2H, *J* = 7.2 Hz, OC*H*$_2$); 1.84 (quin, 2H, *J* = 7.2 Hz, OCH$_2$C*H*$_2$); 1.04 (t, 3H, *J* = 7.2 Hz, C*H*$_3$). The

physical-chemical data shown are the same of those reported in literature.[104]

1-amino-4-propoxy benzene (41)

Under N_2, **36** (0.61 g, 3.34 mmol) was dissolved in EtOH (10 ml) and a catalytic amount of Pd/C (10%) was added. Hydrazine hydrate (3.0 ml, 62.07 mmol). The reaction was stirred overnight at 85°C and when it was finished (TLC eluent: hexane/EtOAc 1:1) the catalyst was removed by filtration. The solvent was removed under reduced pressure and the pure compound **37** was obtained without any purification (yellowish solid, 91% yield). **^{1}H-NMR** (300 MHz, $CDCl_3$) δ (ppm): 6.76-6.72 (m, 2H, Ar*H*); 6.66-6.62 (m, 2H, Ar*H*); 3.84 (t, 2H, *J* = 7.0 Hz, OC*H*$_2$); 3.40 (s, 2H, NH_2); 1.76 (quin, 2H, *J* = 7.0 Hz, OCH_2CH_2); 1.01 (t, 3H, *J* = 7.0 Hz, C*H*$_3$).[104,105]

1-isothiocyanate-4-propoxy-benzene (42)

In a 2 neck round bottom flask, under N_2, **37** (0.2 g, 1.32 mmol) was dissolved in dry toluene (20 ml) and subsequently thiophosgene (35 µl, 0.45 mmol) and Et_3N (1.84 ml, 13.3 mmol) were added. The mixture was stirred at rt for 12h and monitored by TLC (eluent: cyclohexane/EtOAc 7:3). After removal of the solvent at the end of the reaction, the mixture was dissolved in DCM (30ml) and extracted with HCl 1M (25 ml), H_2O (25 ml) and brine (25ml). The organic phase was evaporated under vacuum and the crude was purified on silica gel pad (eluent: cyclohexane/DCM 9:1) to afford **38** as a yellow oil (yield: 32%).

^{1}H-NMR (400 MHz, $CDCl_3$) δ (ppm): 7.12 (d, 2H, *J* = 8.4 Hz, Ar*H*); 6.82 (d, 2H, *J* = 8.4 Hz, Ar*H*); 3.88 (t, 2H, *J* = 8.4 Hz, OC*H*$_2$); 1.78 (quint, 2H, *J* = 7.0 Hz, OCH_2CH_2); 1.02 (t, 3H, *J* = 7.0 Hz, C*H*$_3$).

[(aminopropil-2-acetamido-3,4,6-Tri-*O*-acetyl-2-deoxy-β-D-mannopiranosyl)thioureido]-1-propoxy benzene (43)

The crude was purified by flash column chromatography (eluent: DCM/MeOH 20:1 or hexane/EtOAc/MeOH 6:4:1) in 53% yield as off-white solid. **^{1}H-NMR** (400 MHz, $CDCl_3$) δ (ppm): 8.25 (br s, 1H, N*H*CS); 7.18 (d, 2H, *J* = 8 Hz, Ar*H*); 6.86 (d, 2H, *J* = 8 Hz, Ar*H*); 6.23 (br s, 1H, CH_2NHCS); 6.05 (d, 1H, *J* = 7.6Hz, N*H*Ac); 5.08 (t, 1H, *J* = 9.6 Hz, H_4); 4.99 (dd, 1H, $J_{2\text{-}3}$ = 4,

$J_{3\text{-}4}$ = 10 Hz, H_3); 4.72 (d, 1H, *J* = 5.2 Hz, H_2); 4.63 (s, 1H, H_1); 4.26 (dd, 1H, $J_{5\text{-}6a}$ = 5.2, $J_{6a\text{-}6b}$ = 12.4 Hz, H_{6a}); 4.08 (d, 1H, *J* = 12 Hz, H_{6b}); 3.89-3.86 (m, 3H, H_{7a}, OCH_2); 3.64-3.63 (m, 1H, H_{9a}); 3.62-3.60 (m, 2H, H_5, H_{7b}); 3.53 (m, 1H, H_{9b}); 2.09, 2.06, 2.03, 2.02, 1.97 (s, 12H, 3x$COCH_3$, NHCOCH_3); 1.75 (m, 4H, OCH_2CH_2, $H_{8a,b}$); 0.98 (t, 3H, J=7.2 Hz, CH_3). **^{13}C-NMR** (100 MHz, $CDCl_3$) δ (ppm): 181.7 (*C*S); 171.7 (NH*C*OCH_3); 170.6, 170.1, 169.7 (3X *C*OCH_3); 157.9 (Ar ipso); 129.8 (At ortho); 127.3 (Ar para); 127.3, 125.6 (Ar meta); 99.1 (C_1); 72.5 (C_5); 71.0 (C_3); 69.8 (ArO*C*H_2); 68.5 (C_7); 66.0 (C_4); 62.4 (C_6); 50.3 (C_2); 43.1 (C_9); 29.7 (C_8); 23.4 (CO*C*H_3); 20.8 (ArOCH_2*C*H_2); 20.8-20.7 ($COCH_3$); 10.5 (ArOCH_2CH_2*C*H_3). **ESI-MS (+)**: m/z 620.3 [100%; $(M+Na)^+$].

M.p.: 90.7-91.1 °C

[(aminopropil-2-acetamido-2-deoxy-β-D-mannopiranosyl)thioureido]-1-propoxy benzene (5)

The pure product was obtained by purification on C18 reverse silica gel column (eluent: MeOH/H_2O 1:1) as a off-white solid (yield = 66%).

^{1}H-NMR (400 MHz, MeOD) δ (ppm): 7.22 (d, 2H, *J* = 7.0 Hz, Ar*H*); 6.92 (d, 2H, *J* = 8 Hz, Ar*H*); 4.65 (d, 1H, $J_{1\text{-}2}$ = 1.2 Hz, H_1); 4.48 (d, 1H, *J* = 3.6 Hz, H_2); 3.95 (t, 2H, *J* = 6.4, OCH_2); 3.91 (m, 1H, H_{7a}); 3.86 (m, 1H, H_{6a}); 3.69 (br s, 1H, H_{9a}); 3.66 (dd, 1H, $J_{2\text{-}3}$ = 5.2, $J_{3\text{-}4}$ = 9.6 Hz, H_3); 3.63 (m, 2H, H_{7b}, H_{9b}); 3.53 (t, 1H, H_4); 2.04 (NHCOCH_3); 1.89-1.76 (m, 4H, OCH_2CH_2, $H_{8a,b}$); 1.06 (t, 3H, *J* = 7.2 Hz, CH_3). **^{13}C-NMR** (100 MHz, MeOD) δ (ppm): 181.1 (*C*S); 173.5 (NH*C*OCH_3); 157.7 (Ar ipso); 127.3 (Ar para); 126.7, 114.5 (Ar meta); 99.4 (C_1); 77.0 (C_5); 76.9 (C_5); 76.8 (ArO*C*H_2); 69.4 (C_4, C_7); 60.6 (C_6); 53.5 (C_2); 41.9 (C_9); 28.6 (C_8); 22.3 (OCH_2*C*H_2); 21.4 (NHCO*C*H_3); 9.4 (ArOCH_2CH_2*C*H_3). **ESI-MS (+)**: m/z 494.4 [100% $(M+Na)^+$] **M.p.:** > 76°C dec

3.5 References

(1) González-Fernández, Á.; Faro, J.; Fernández, C. *Vaccine* **2008**, *26* (3), 292–300.

(2) Morelli, L.; Poletti, L.; Lay, L. *European J. Org. Chem.* **2011**, 5723–5777.

(3) Mond, James J. , Lees, Andrew, Clifford, M. S. *Annu. Rev. Immunol.* **2011**, *13*, 655–692.

(4) Thomas, S.; Luxon, B. a. *Expert Rev. Vaccines* **2013**, *12*, 1301–1311.

(5) Torosantucci, A. *J. Exp. Med.* **2005**, *202* (5), 597–606.

(6) Paoletti, L. C.; Kasper, D. L. *Expert Opin. Biol. Ther.* **2003**, *3* (6), 975–984.

(7) Kodali, S.; Vinogradov, E.; Lin, F.; Khoury, N.; Hao, L.; Pavliak, V.; Jones, H.; Laverde, D.; Huebner, J.; Jansen, K. U.; Anderson, A. S.; Donald, R. G. K. *J. Biol. Chem.* **2015**, *290* (32), jbc.M115.655852.

(8) Cremades, N.; Velázquez-Campoy, A.; Martínez-Júlvez, M.; Neira, J. L.; Pérez-Dorado, I.; Hermoso, J.; Jiménez, P.; Lanas, A.; Hoffman, P. S.; Sancho, J. *ACS Chem. Biol.* **2009**, *4* (11), 928–938.

(9) McMichael, A. J.; Borrow, P.; Tomaras, G. D.; Goonetilleke, N.; Haynes, B. F. *Nat Rev Immunol* **2010**, *10* (1), 11–23.

(10) Ouerfelli, O.; Warren, J. D.; Wilson, R. M.; Danishefsky, S. J. *Expert Rev. Vaccines* **2005**, *4* (5), 677–685.

(11) Nativi, C.; Renaudet, O. *ACS Med. Chem. Lett.* **2014**, *5* (11), 1176–1178.

(12) Lucas, A. H.; Rittenhouse-olson, K.; Kronenberg, M.; Apicella, M. A.; Wang, D.; Schreiber, J. R.; Taylor, C. E. *Vaccine* **2010**, *28*, 1121–1131.

(13) Goldblatt, D. *Clin Exp Immunol* **2000**, *119* (119), 1–3.

(14) Ni, J.; Song, H.; Wang, Y.; Stamatos, N. M.; Wang, L.-X. *Bioconjug. Chem.* **2006**, *17* (2), 493–500.

(15) Laferriere, C. A.; Sood, R. K.; Muys, J. D. E.; Michon, F.; Jennings, H. J. *Infect. Immun.* **1998**, *66* (6), 2441–2446.

(16) Robbins,J. B., Schneerson,R. , Anderson P., S. H. *JAMA* **1996**, *276* (5683), 1181.

(17) Peri, F. *Chem. Soc. Rev.* **2013**, *42* (11), 4543–4556.

(18) Bernardi, A.; Jiménez-Barbero, J.; Casnati, A.; De Castro, C.; Darbre, T.; Fieschi, F.; Finne, J.; Funken, H.; Jaeger, K.-E.; Lahmann, M.; Lindhorst, T. K.; Marradi, M.; Messner, P.; Molinaro, A.; Murphy, P. V.; Nativi, C.; Oscarson, S.; Penadés, S.; Peri, F.; Pieters, R. J.; Renaudet, O.; Reymond, J.-L.; Richichi, B.; Rojo, J.; Sansone, F.; Schäffer, C.; Turnbull, W. B.; Velasco-Torrijos, T.; Vidal, S.; Vincent, S.; Wennekes, T.; Zuilhof, H.; Imberty, A. *Chem. Soc. Rev.* **2013**, *42* (11), 4709–4727.

(19) Nitz, M.; Bundle, D. R. *J. Org. Chem.* **2001**, *66*, 8411–8423.

(20) Legnani, L.; Ronchi, S.; Fallarini, S.; Lombardi, G.; Campo, F.; Panza, L.; Lay, L.; Poletti, L.; Toma, L.; Ronchetti, F.; Compostella, F. *Org. Biomol. Chem.* **2009**, *7* (21), 4428–4436.

(21) Goebel, W. F.; Avery, O. T. *J. Exp. Med.* **1929**, *50* (5), 533–550.

(22) *Lancet* **1992**, *340* (8819), 592–594.

(23) Vliegenthart, J. F. G. *FEBS Lett.* **2006**, *580*, 2945–2950.

(24) Brady C et al. *Bioprocess Tech.* **2012**, *10*, 50–55.

(25) Kuberan, B.; Linhardt, R. J. *Curr. Org. Chem.* **2000**, *4*, 653–677.

(26) Lepenies, B.; Yin, J.; Seeberger, P. H. *Curr. Opin. Chem. Biol.* **2010**, *14* (3), 404–411.

(27) Hammerschmidt, S.; Wolff, S.; Hocke, A.; Rosseau, S.; Müller, E.; Rohde, M. *Infect. Immun.* **2005**, *73* (8), 4653–4667.

(28) Roberts, I. S. *Annu. Rev. Microbiol.* **1996**, *50*, 285–315.

(29) Alonsodevelasco, E.; Verheul, A. F.; Verhoef, J.; Snippe, H. *Microbiol. Rev.* **1995**, *59* (4), 591.

(30) Lee, C. *Mol. Immunol.* **1987**, *10*, 1005–1019.

(31) Barry M. Gray, G. M. C. I. and H. C. D. J. *J. Infect. Dis.* **1979**, No. 140, 979–983.

(32) Grabenstein, J. D.; Klugman, K. P. *Clin. Microbiol. Infect.* **2012**, *18* (5), 15–24.

(33) Tai, S. S. *Crit. Rev. Microbiol.* **2006**, *32*, 139–153.

(34) Heidelberger, M.; Goebel, W. F.; Avery, O. T. *J. Exp. Med.* **1925**, *42* (5), 727–745.

(35) Avery, O. T.; Dubos, R. *J. Exp. Med.* **1931**, *54* (1), 51–71.

(36) Austrian, R. *J Infect Dis.* **1977**, *(Supplemen*, S38–S42.

(37) Mbelle, N.; Huebner, R. E.; Wasas, A. D.; Kimura, A.; Chang, I.; Klugman, K. P. *J. Infect. Dis.* **1999**, *180* (4), 1171–1176.

(38) Dagan, R.; Goldblatt, D.; Maleckar, J. R.; Yaïch, M.; Eskola, J. *Infect. Immun.* **2004**, *72* (9), 5383–5391.

(39) Verez-Bencomo, V.; Fernandez-Santana, V.; Hardy, E.; Toledo, M. E.; Rodriguez, M. C.; Heynngnezz, L.; Rodriguez, A.; Baly, A.; Herrera, L.; Izquierdo, M.; Villar, A.; Valdes, Y.; Cosme, K.; Deler, M. L.; Montane, M.; Garcia, E.; Ramos, A.; Aguilar, A.; Medina, E.; Torano, G.; Sosa, I.; Hernandez, I.; Martinez, R.; Muzachio, A.; Carmenates, A.; Costa, L.; Cardoso, F.; Campa, C.; Diaz, M.; Roy, R. *Science (80-.).* **2004**, *305* (5683), 522–525.

(40) Bousquet, E.; Khitri, M.; Lay, L.; Nicotra, F. **1998**, *311*, 171–181.

(41) Bonaccorsi, F.; Catelani, G.; Oscarson, S. *Carbohydr. Res.* **2009**, *344* (12), 1442–1448.

(42) Ronchi, P.; Scarponi, C.; Salvi, M.; Fallarini, S.; Polito, L.; Caneva, E.; Bagnoli, L.; Lay, L.; Chimica, D.; Milano, I.-. *J. Org. Chem.* **2013**, *78*, 5172–5183.

(43) Legnani, L.; Ronchi, S.; Fallarini, S.; Lombardi, G.; Campo, F.; Panza, L.; Lay, L.; Poletti, L.; Toma, L.; Compostella, F. *Org. Biomol. Chem.* **2009**, *7*, 4428–4436.

(44) Paulsen, H., Helpap, B., Lorentzen, J. P. *Carbohydr. Res.* **1988**, *179*, 173–197.

(45) Sugawara, T., Igarashi, K. *Carbohydr. Res.* **1988**, *172* (2), 195–207.

(46) Mammen, M.; Choi, S.-K.; Whitesides, G. M. *Angew. Chemie, Int. Ed. English* **1998**, *37* (20), 2754–2794.

(47) Sansone, F.; Dudic, M.; Donofrio, G.; Rivetti, C.; Baldini, L.; Casnati, A.; Cellai, S.; Ungaro, R. *J. Am. Chem. Soc.* **2006**, *128* (8), 14528–14536.

(48) Consoli, G. M. L.; Cunsolo, F.; Geraci, C.; Sgarlata, V. *Org. Lett.* **2004**, *6* (23), 4163–4166.

(49) Fallarini, S.; Paoletti, T.; Battaglini, C. O.; Ronchi, P.; Lay, L.; Bonomi, R.; Jha, S.; Mancin, F.; Scrimin, P.; Lombardi, G. *Nanoscale* **2013**, *5* (1), 390–400.

(50) Manea, B. F.; Bindoli, C.; Fallarini, S.; Lombardi, G.; Polito, L.; Lay, L.; Bonomi, R.; Mancin, F.; Scrimin, P. *Adv. Mater.* **2008**, *20*, 4348–4352.

(51) Moses, J. E.; Moorhouse, A. D. *Chem. Soc. Rev.* **2007**, *36* (8), 1249–1262.

(52) Torvinen, M.; Neitola, R.; Sansone, F.; Baldini, L.; Ungaro, R.; Casnati, A. *Org. Biomol. Chem.* **2010**, *8*, 906–915.

(53) Consoli, G. M. L.; Cunsolo, F.; Geraci, C.; Mecca, T.; Biomolecolare, C.; Santuario, V.; Ct, I. V. *Tetrahedron Lett.* **2003**, *44*, 7467–7470.

(54) García Fernández, José Manuel, O. M. C. *Adv. Carbohydr. Chem. Biochem.* **1999**, *55*, 35–135.

(55) Benito, J. M.; Gómez-García, M.; Jiménez Blanco, J. L.; Ortiz Mellet, C.; García Fernández, J. M. *J. Org. Chem.* **2001**, *66* (4), 1366–1372.

(56) Gutsche, C. D. *Calixarenes 2001* **2001**, 1–25.

(57) Kenis, P. J. .; Noordman, O. F. .; Schönherr, H.; Kerver, E. G.; Snellink-Ruel, B. H. .; van Hummel, G. J.; Harkema, S.; van der Vorst, C. P. J. .; Hare, J.; Picken, S. J.; Others. *Chem. Eur. J.* **1998**, *4*, 1225–1234.

(58) Kim, T.; Kim, Y.-J.; Han, I.-H.; Lee, D.; Ham, J.; Kang, K. S.; Lee, J. W. *Bioorg. Med. Chem. Lett.* **2015**, *25* (1), 62–66.

(59) Munch, H.; Hansen, J. S.; Pittelkow, M.; Christensen, J. B.; Boas, U. *Tetrahedron Lett.* **2008**, *49* (19), 3117–3119.

(60) Gridley, J. J.; Osborn, H. M. I. *J. Chem. Soc. Perkin Trans. 1* **2000**, No. 10, 1471–1491.

(61) Cumpstey, I. *Org. Biomol. Chem.* **2012**, *10*, 2503.

(62) Zhang, Z.; Liu, J.; Verlinde, C. L. M. J.; Hol, W. G. J.; Fan, E. *J. Org. Chem.* **2004**, *69* (22), 7737–7740.

(63) Michihata, N.; Kaneko, Y.; Kasai, Y.; Tanigawa, K.; Hirokane, T.; Higasa, S.; Yamada, H. *J. Org. Chem.* **2013**, *78*, 4319–4328.

(64) Lichtenthaler, F. W.; Kaji, E.; Weprek, S. *J. Org. Chem.* **1985**, *50* (19), 3505–3515.

(65) Lichtenthaler, F.W. , Jarglis, P. *Tetrahedron Lett.* **198AD**, *21*, 1425–1428.

(66) Lergenmüller, M.; Kläres, U.; Lichtenthaler, F. W. *Carbohydr. Res.* **2009**, *344* (16), 2127–2136.

(67) Lichtenthaler, F. W. *Chem. Rev.* **2011**, *111* (9), 5569–5609.

(68) Ferrier, Robert J., T. P. C. *J. Chem. Soc., Perkin Trans. 1* **1980**, No. 2767-2773.

(69) Czifra, K. *Tetrahedron Lett.* **2002**, *43*, 8849–8852.

(70) Lergenmu, M.; Lichtenthaler, F. W. *Carbohydr. Res.* **2007**, *342*, 2132–2137.

(71) Schneider-adams, T.; Lichtenthaler, F. W. *J. Org. Chem.* **1994**, *59* (14), 6728–6734.

(72) Lichtenthaler, Frieder W., Kläres, U. , Lergenmüller, M., Schwidetzky, S. *Synthesis (Stuttg).* **1992**, 179–184.

(73) Torres-sanchez, M. I.; Zaccaria, C.; Buzzi, B.; Miglio, G.; Lombardi, G.; Polito, L.; Russo, G.; Lay, L. *Chem. Eur. J.* **2007**, *13*, 6623–6635.

(74) Morelli, L.; Cancogni, D.; Tontini, M.; Nilo, A.; Filippini, S.; Costantino, P.; Romano, M. R.; Berti, F.; Adamo, R.; Lay, L. *Beilstein J. Org. Chem.* **2014**, *10*, 2367–2376.

(75) Ii, D. C. J.; Widlanski, T. S. *Org. Lett.* **2004**, *6* (25), 4643–4646.

(76) P.G. Fox, J. Ehretsmann, B. C. E. *J. Catal.* **1971**, *20*, 58–66.

(77) Murali, C.; Shashidhar, M. S.; Gopinath, C. S. *Tetrahedron* **2007**, *63* (19), 4149–4155.

(78) Quinn, J. F.; Razzano, D. A.; Golden, K. C.; Gregg, B. T. *Tetrahedron Lett.* **2008**, *49* (42), 6137–6140.

(79) Daga, M. C.; Taddei, M.; Varchi, G. *Tetrahedron Lett.* **2001**, *42* (31), 5191–5194.

(80) Brieger, G.; Nestrick, T. J. *Chem. Rev.* **1974**, *64* (5), 567–580.

(81) Torres-sanchez, M. I.; Zaccaria, C.; Buzzi, B.; Miglio, G.; Lombardi, G.; Polito, L.; Russo, G.; Lay, L. **2007**, 6623–6635.

(82) Sureshbabu, V. V.; Naik, S. A.; Hemantha, H. P.; Narendra, N.; Das, U.; Guru Row, T. N. *J. Org. Chem.* **2009**, *74* (15), 5260–5266.

(83) Iwamoto, K.; Araki, K.; Shinkai, S. *J. Org. Chem.* **1991**, *29* (26), 4955–4962.

(84) Ren, B.; Wang, M.; Liu, J.; Ge, J.; Zhang, X.; Dong, H. *Green Chem.* **2015**, *17* (3), 1390–1394.

(85) Manuel, J.; Ortiz, C.; Sadalapure, K.; Garcı, M. *Carbohydr. Res.* **1999**, *320*, 37–48.

(86) Sansone, F.; Baldini, L.; Casnati, A.; Ungaro, R. *Supramol. Chem.* **2008**, *20*, 161–168.

(87) van Wageningen, André M. A., Snip, Erwin, Verboom, Willem, Reinhoudt David N., Boerrigte, H. *Liebigs Ann.* **1997**, *11*, 2235–2245.

(88) André, S.; Sansone, F.; Kaltner, H.; Casnati, A.; Kopitz, J.; Gabius, H. J.; Ungaro, R. *ChemBioChem* **2008**, *9* (10), 1649–1661.

(89) Gutsche, C.David Dhawan, Balram Levine, Jeffrey A. Hyun No, Kwang Bauer, L. J. *Tetrahedron Lett.* **1983**, *39* (3), 409–426.

(90) Kelderman, E., Der Haeg, L., Heesink, G. J. T., Verboom, W., Engbersen, J. F. J. *Angew. Chemie - Int. Ed.* **1992**, *39*, 91–92.

(91) Chang, Suk-Kyu and Cho, I. *J. Chem. Soc., Perkin Trans. 1* **1986**, *1*, 211–214.

(92) Dudic, M.; Colombo, A.; Sansone, F.; Casnati, A. *Tetrahedron* **2004**, *60*, 11613–11618.

(93) Timmerman, P.; Weidmann, J.-L.; Jolliffe, K. a.; Prins, L. J.; Reinhoudt, D. N.; Shinkai, S.; Frish, L.; Cohen, Y. *J. Chem. Soc. Perkin Trans. 2* **2000**, 2077–2089.

(94) Gallego-Yerga, L.; Lomazzi, M.; Sansone, F.; Ortiz Mellet, C.; Casnati, A.; García Fernández, J. M. *Chem. Commun. (Camb).* **2014**, *50* (56), 7440–7443.

(95) Sansone, F.; Chierici, E.; Casnati, A.; Ungaro, R. *Org. Biomol. Chem.* **2003**, *11*, 1802–1809.

(96) Liu, S.-Y.; He, Y.-B.; Wu, J.-L.; Wei, L.-H.; Qin, H.-J.; Meng, L.-Z.; Hu, L. *Org. Biomol. Chem.* **2004**, *2* (11), 1582–1586.

(97) Nie, X.; Wang, G. *J. Org. Chem.* **2005**, *70* (22), 8687–8692.

(98) Urien, M.; Erothu, H.; Cloutet, E.; Hiorns, R. C.; Vignau, L.; Cramail, H. *Macromolecules* **2008**, *41* (19), 7033–7040.

(99) Boyer, C.; Liu, J.; Bulmus, V.; Davis, T. P.; Barner-Kowollik, C.; Stenzel, M. H. *Macromolecules* **2008**, *41* (15), 5641–5650.

(100) Mouton, C.; Tillequin, F.; Seguin, E.; Monneret, C. *J. Chem. Soc. Perkin Trans. 1* **1998**, No. 13, 2055–2060.

(101) Miller, D. J.; Bashir-Uddin Surfraz, M.; Akhtar, M.; Gani, D.; Allemann, R. K. *Org. Biomol. Chem.* **2004**, *2*, 671–688.

(102) Geall, A. J.; Blagbrough, I. S. *Tetrahedron* **2000**, *56*, 2449–2460.

(103) Koziol, A.; Lendzion-Paluch, A.; Manikowski, A. *Org. Process Res. Dev.* **2013**, *17* (5), 869–875.

(104) Ho, I.-T.; Chu, J.-H.; Chung, W.-S. *European J. Org. Chem.* **2011**, *2011* (8), 1472–1481.

(105) Selva, M.; Tundo, P.; Perosa, A. *J. Org. Chem.* **2002**, *67* (26), 9238–9247.

3.6 APPENDIX: NMR spectra

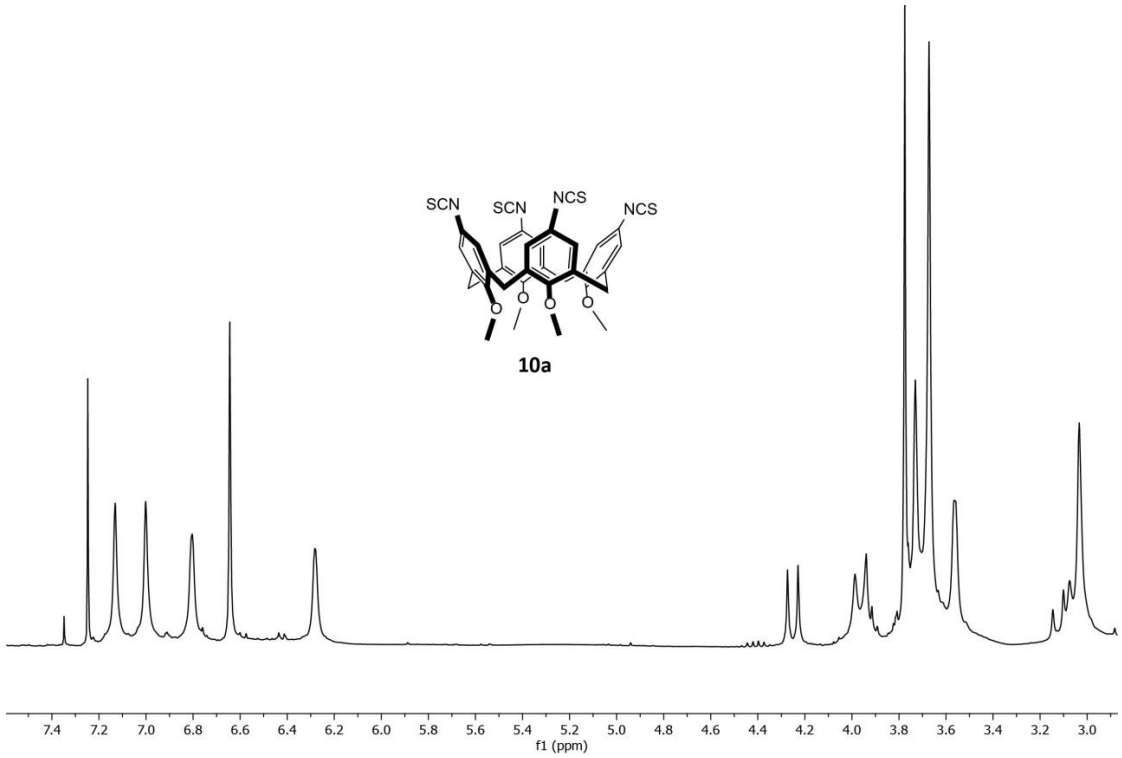


Fig 1A : NMR spectrum (300 MHz, $CDCl_3$) of **10a**

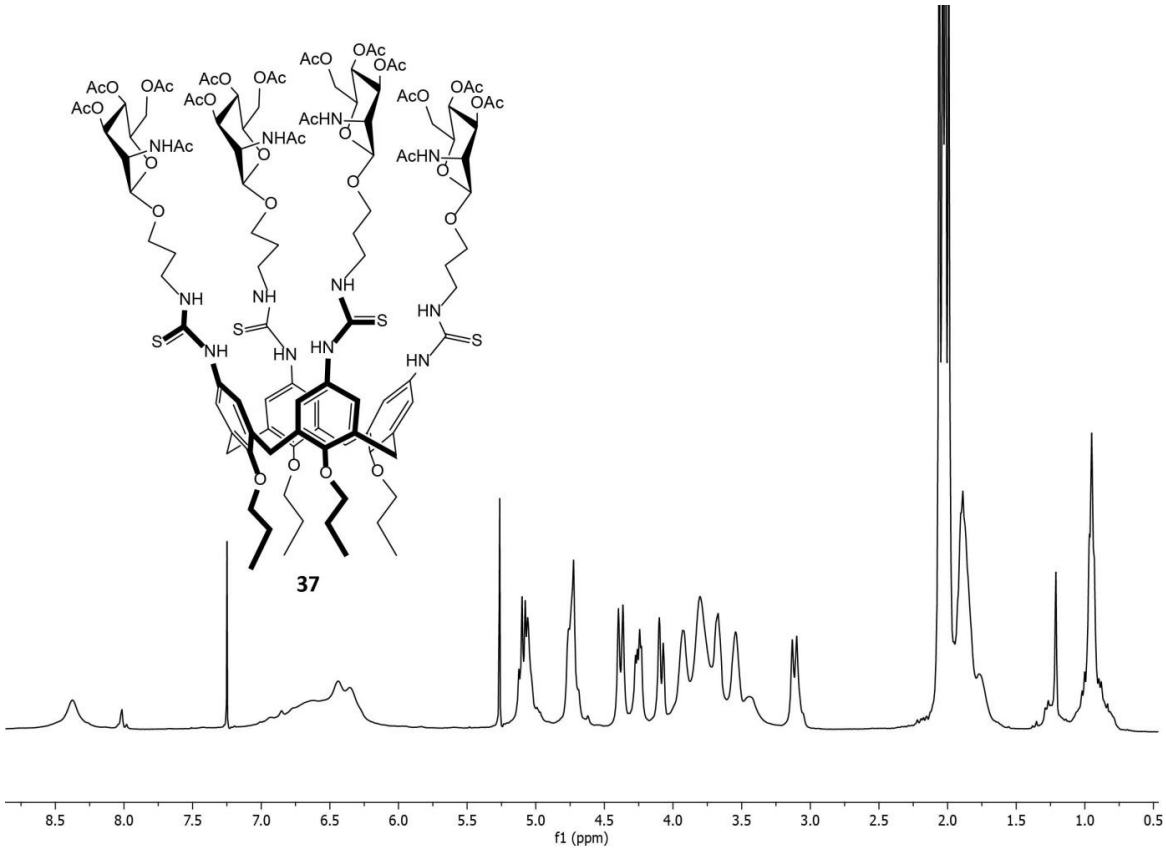


Fig 2A : NMR spectrum (400 MHz, $CDCl_3$) of **37**

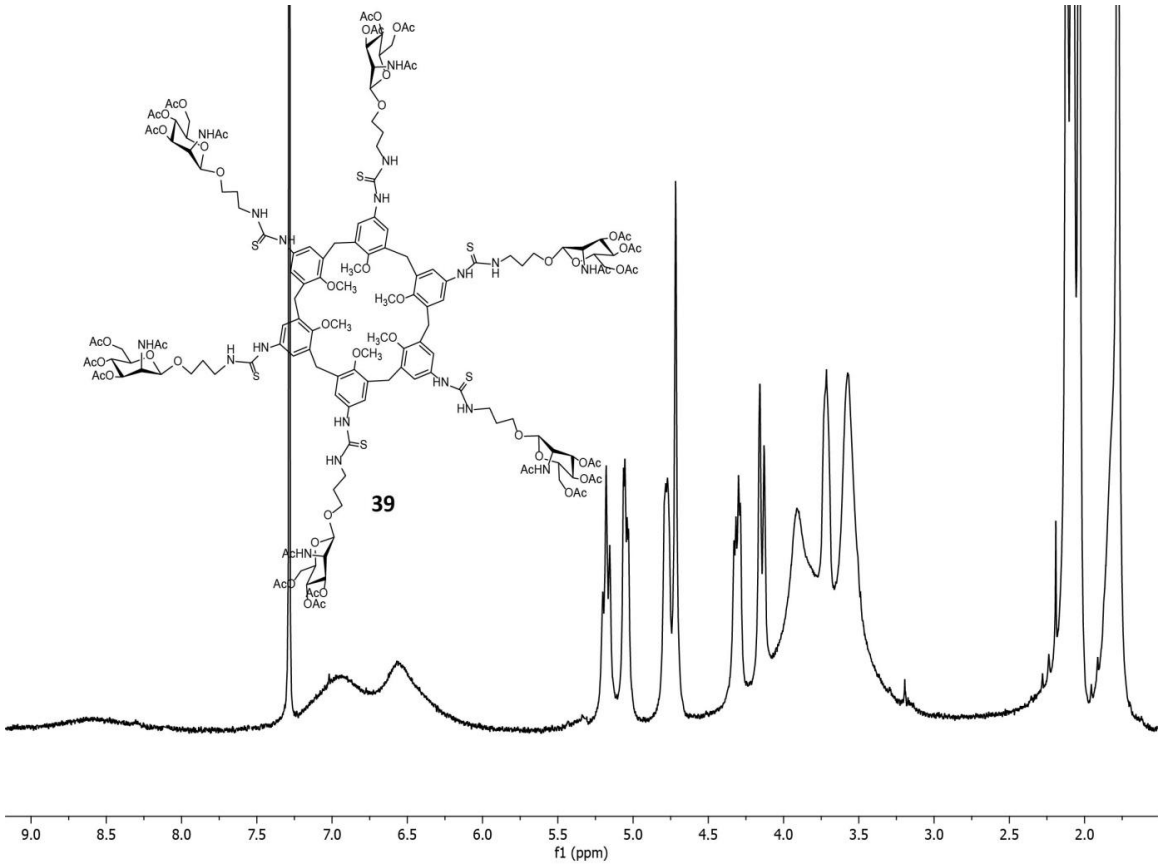


Fig 3A : NMR spectrum (400 MHz, $CDCl_3$, 80°C) of **39**

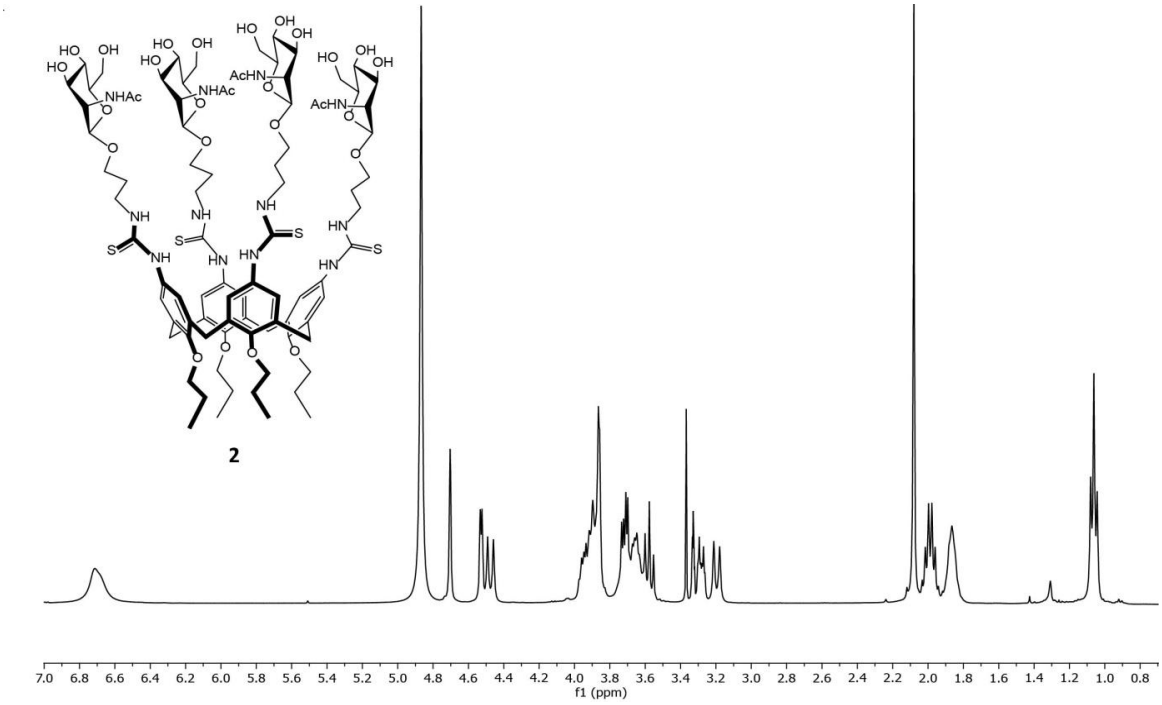


Fig 4A : NMR spectrum (400 MHz, MeOD) of **2**

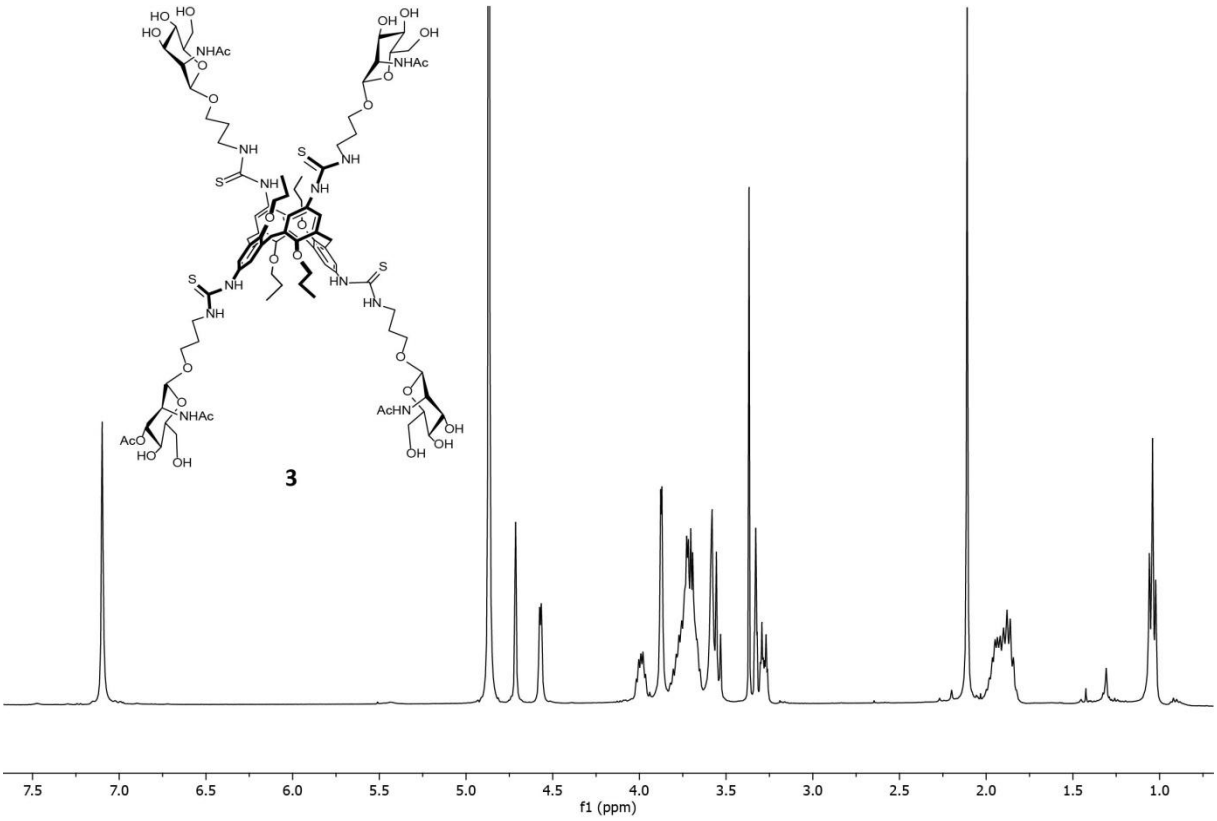


Fig 5A : NMR spectrum (400 MHz, MeOD) of **3**

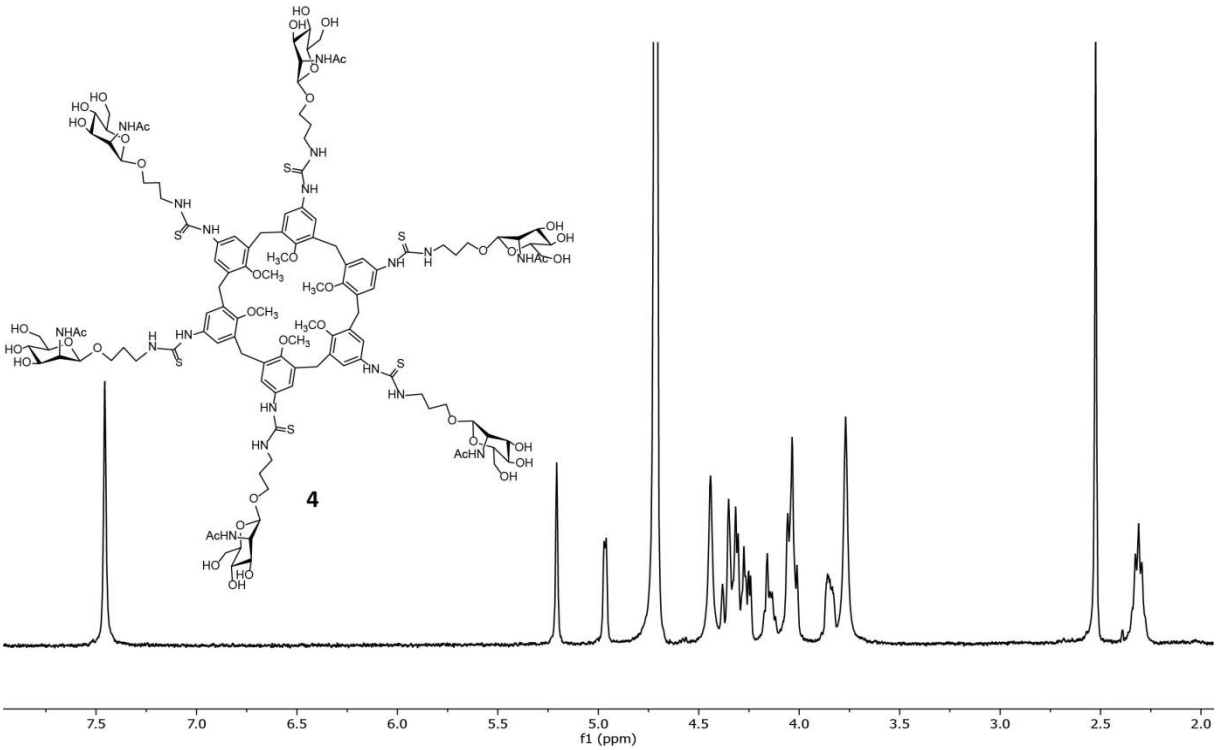


Fig. 6A : NMR spectrum (400 MHz, D_2O, 80°C) of **4**

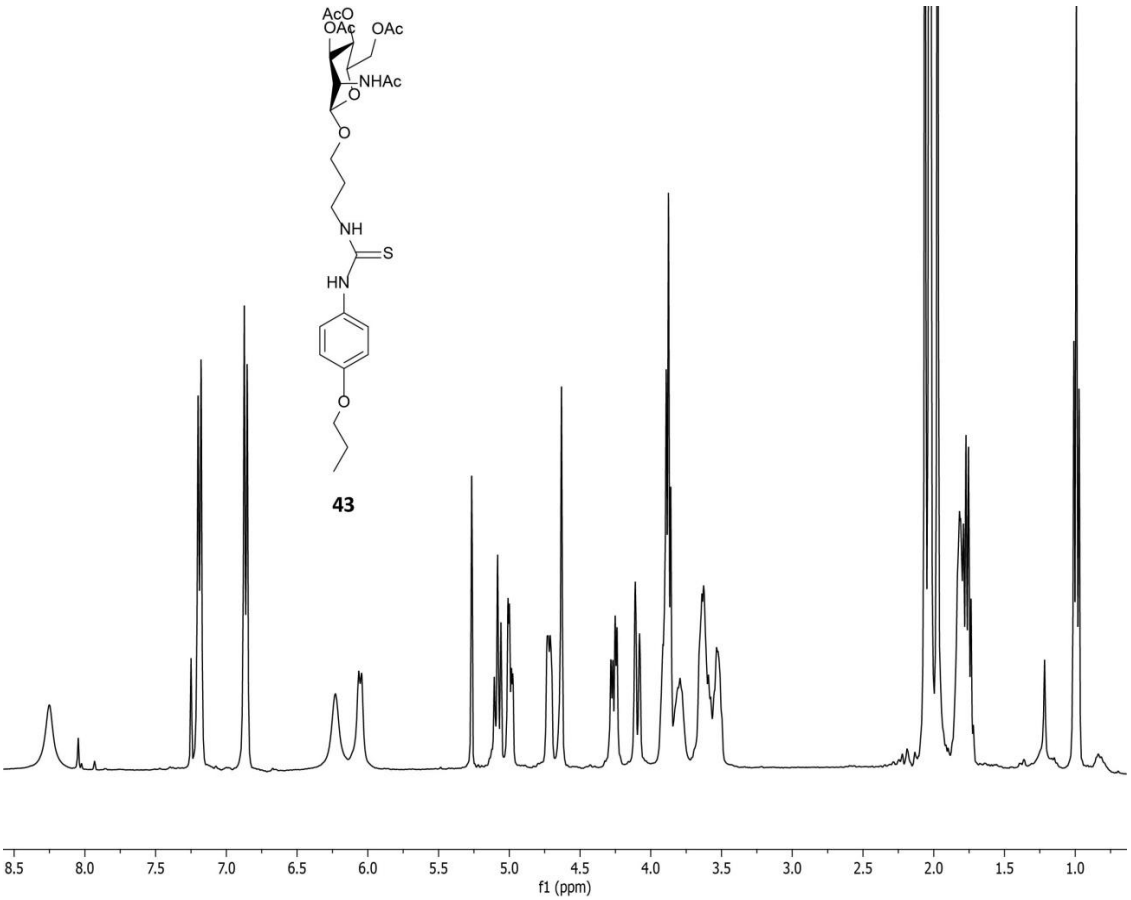


Fig 7A : NMR spectrum (400 MHz, $CDCl_3$) of **43**

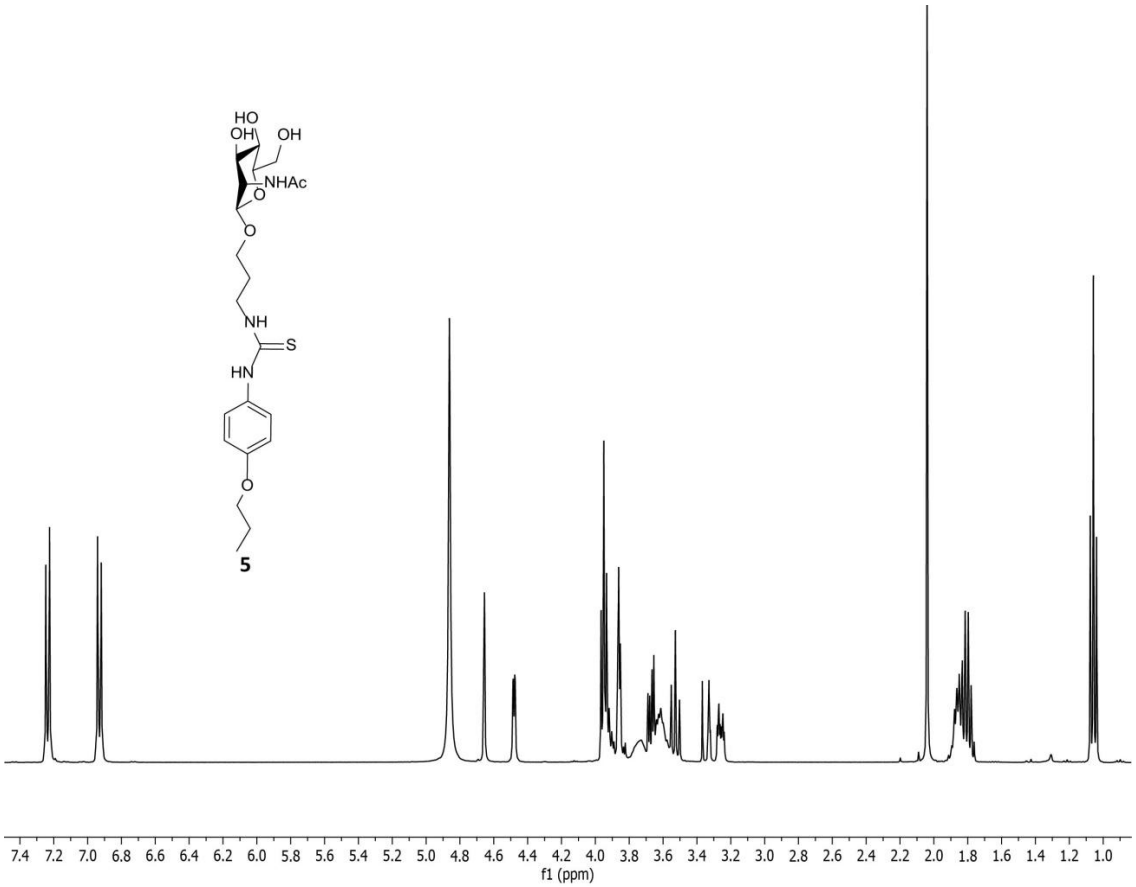


Fig 8A : NMR spectrum (400 MHz, MeOD) of **5**

Chapter 4

Structural studies of protein-ligand interactions

4.1 Introduction

4.1.1 The ^{1}H-^{15}N HSQC experiment and the chemical shift perturbation to study protein-ligand interactions

The possibility to understand how proteins interact with their biological counterparts and with synthetic molecules designed as potential ligands is very important since these interactions are involved in or can modulate many biological processes and therefore can help in the development of new drugs.

NMR spectroscopy[1] can be used to understand these binding phenomena. This technique is a powerful tool to study protein-protein[2] and protein-ligand interactions[3,4] which can yield structural information about the binding site as well as thermodynamic data. Nowadays there are several experiments that allow to study these phenomena[5], based on changes in relaxation or diffusion behaviour of the ligand, NOE effects between protein and ligand. Also chemical shifts can be used to follow the binding process. Particularly, using 2D ^{1}H-^{15}N HSQC (heteronuclear single quantum correlation) spectroscopy, it is possible to obtain information about binding by comparing the bidimensional maps of the ^{15}N-enriched protein in the absence and presence of ligands. For this type of experiment a ^{15}N labelled protein sample is required and the unlabelled ligand is added to the protein sample and for every point of titration the 2D spectrum is recorded. If there is interaction, the presence of the ligand alters the chemical environment around the protein, giving chemical shift changes (chemical shift perturbation[6], CSP). Even for very weak complexes these CPSs can be discernible.

In the ^{1}H-^{15}N HSQC spectrum the present signals (cross-peaks) are related to the H-N coupling of each amide bond in the protein backbone and one signal per amino acid, except proline, is detected. Only asparagine and glutamine possess two signals due to the presence of two amide bonds, one for the NH_2 side chain group. When the assignment of each peak is done for the free protein, several information can be obtained about the residues involved in the interactions from the observed perturbations that some of them undergo upon titration. In this way the binding sites can be identified. Additionally, CSPs of the labelled protein in the presence of the ligand can give data on the affinity[7], stoichiometry[8] and also kinetic of the binding.[9] In fact, for the kinetic, when there is fast exchange on the NMR timescale between complex and free protein, the peaks of the protein shift linearly with the

increasing amount of the interacting partner, reaching a plateau when achieved saturation. In this way the observed chemical shifts during titration are due to an weighed combination of free and bound state of the protein.

On the contrary, when the complex dissociation is slow, during the titration the resonances for the free protein gradually disappear and are replaced by resonances for the free protein, with a coexistence of both for a certain range of concentrations. This is referred to the slow exchange on the NMR timetable and the intensity of the two series of peaks represents the concentration of the two species in solution. There is also an intermediate case,[2] where during the addition of the ligand to the protein solution the peak intensities become weaker, less defined and could even disappear from the NMR bidimensional map. In this case the residues belonging to the interacting surface are delineated by very broad signals of low intensity.[10]

For a thermodynamic point of view, dissociation constants K_d of the complexes, and consequently K_{ass}, can be obtained from a titration *via* HSQC experiment. The data allow to calculate the binding curves by monitoring the chemical shift changes of the backbone amides as a function of ligand concentration. The K_d values are achieved fitting chemical shift changes in both ^{1}H and ^{15}N dimension by non-linear regression analysis, using proper softwares Particularly, this is the only technique that straightforward provides both thermodynamic data and knowledge on the binding site from the same experiment. Another feature of using NMR method it is the extension of the range of measurable interactions into the mM range, a region not well covered by traditional biochemical binding assays.[7] NMR is ideally suited for the analysis of protein-ligand interactions with dissociation constants ranging from 2 μM to 10 mM.

This is a very instant technique because if the ligand does not bind protein, there are not CPSs and so not false positive results are obtained. The analysis of the data has a simple concept, because CSPs are measured at every titration point and at the end, following the peak movements and the measurement of the resonances give the information about the recognition phenomenon. It is also a quick method since spectra can be acquired in 10-20 min and with good resolution of resonances. However this technique has some drawbacks such as the very high variability, depending on experimental conditions. This means that it is important to keep every parameters (temperature, pH of the sample, concentrations and so

on) as consistent as possible.[6] Sometimes it is not so easy to determine what is the real cause of CSPs or better the real contact point between protein and ligand by using only NMR data since for example the macromolecule could also significantly change conformation upon binding. In that case some observed shifts could be due to this rearrangement of the folding and involve residue far from the binding site. This can determine mistakes in the interpretation of the data or, in any case, it is necessary to be very careful in the treatment of information that, however, are on the whole very reach and useful.

The low sensitivity is also a parameter that can be considered as disadvantage, since major amount of protein is needed, but nowadays this is overcome by using cryoprobe and TROSY[11,12](Transverse Relaxation Optimized SpectroscopY) experiments. Definitely, the invention of cryogenically cooled probes (probes which have their electronics cooled to 20-25 K) greatly diminish the level of thermal noise. In this way the signal to noise ratio is raised by a factor of three or more and, consequently, the sensitivity of the NMR experiment is improved.[13] This is correlated also to the reduction in the time experiment of approximately 10-fold and also a decrease in the protein concentration needed to achieve the same signal-to-noise with a room temperature probe.[14] On the other hand, thanks to TROSY, larger proteins may be studied, although some problems remain with the high number of cross-peaks that have to be assigned.[15]

4.1.2 Protein-ligand crystallization

Even if NMR spectroscopy can give important information on protein-ligand interactions, the possibility to solve the X-ray structure of the complex can reinforce the NMR data, especially when the situation is ambiguous. Therefore the combination between the HSQC data and the crystal structure of the protein-ligand complex provides a big advantage because the resolved structure can be used to better understand the binding phenomenon. For instance it is possible to know exactly which residues are involved and which type of interactions are taking place.

Protein crystallization is a powerful tool since the function of these macromolecules is determined by their three-dimensional structure. The possibility to visualize it and how proteins and ligands interact is essential to develop therapeutic treatments and new drugs.

A limiting factor for the structural determination of the protein-ligand complexes is their stability and then a certain strength of the interactions that allow the crystallization of the intact complex instead of the single ligand or the single macromolecule. Moreover, crystals containing both counterparts, must have a sufficient size and physical properties and quality to allow diffraction.

In order to permit protein-ligand crystals formation, many parameters and tools must be taken in account and once found the conditions that yield some crystals, it is important to optimize the individual variables to achieve perfect crystals for diffraction analysis.[16]

Therefore, the first step is the identification of the best conditions to obtain some crystalline materials. For the screening, if the protein crystal structure is already characterized, the starting conditions are those found in literature and they are modified, if not completely adequate to attain good complex crystals. Otherwise if the protein structure is not already characterized, it is necessary to identify chemical, biochemical and physical conditions able to yield crystalline objects and then changing them to reach the best crystals of the protein-ligand complex.[17] pH, protein and ligand concentration, precipitant solutions, temperature, etc.. are parameters that can be changed by performing a large number of trials in to reach the desired result.

The screening can be done using different reservoir solutions (crystallization agents), that are composed of salts, PEGs and buffers They can be prepared manually or employing crystallization screening kits, purchased from different companies. These contain several reservoir solutions, already prepared, very different one to another. In this way many tests can be done easily, in the same moment, without loss of time, exploiting even very distinct conditions. There are also various experimental procedures for the screening, employing:[18]

- liquid dispensing robot (authomatized procedures);
- pipetting devices for setting up microbatch screens (manual procedures).

Fig. 1: Crystallization Robot equipment by Douglas Instruments (image taken from Douglas Instruments website)

There are different methods for protein-ligand crystallization,[19] but particularly here in the work described in this chapter, hanging and sitting drop vapour-diffusion (Fig**. 2**) have been used.

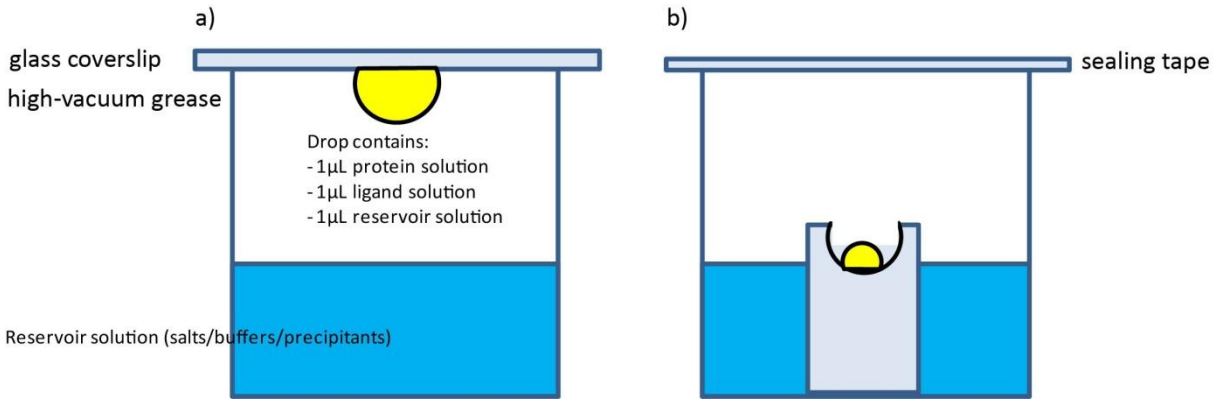


Fig. 2: different vapour diffusion crystallization methods: a) hanging and b) sitting drop.

These conventional set up consist of an aqueous drop where the protein, ligand and the reservoir solution (salts, precipitants etc..) are added separately and mixed. The crystallization mixture (three components) is placed in the vicinity of a reservoir that contains a high concentration of non-volatile precipitating agents or salts. The wells are sealed and an equilibration process starts between the mixture and the reservoir. Slow diffusion of water from the crystallization solution into the concentrated solution of salt is due to the difference in osmotic pressures of them. In this way there is a decrease in the volume of the crystallization mixture, that is necessary to go in the supersaturation of the

crystallization solution.[20] The difference between these methods is the distance between the drop and the reservoir solution[21] but both are easy to perform exploiting several conditions with minimal volumes. The hanging drop technique has as advantages the improvement in the crystal shape and size, thanks to the inverted position of the drop, the possibility to change the drops from one well to another one or reservoir solution, and as disadvantage the use of silicone grease and a siliconized cover slip. On the other hand the sitting drop reduce cost and time, but crystals often adhere to the surface.[20]

In order to understand how to obtain crystals, phase diagram (figure **3**), showing the [protein] against [variable parameter], can help.

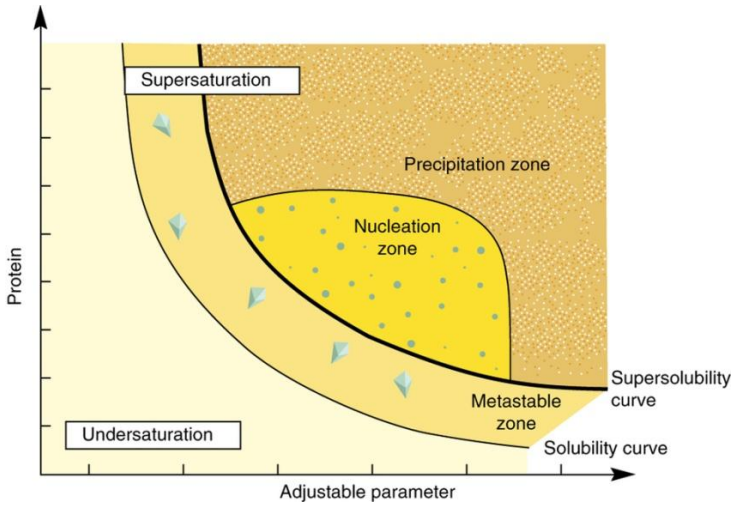


Fig. 3: example of a diagram phase that can help in the crystallization of a protein-ligand complex (figure taken from *Nature Protocols,**9**,1621–1633, (2014))*

The x-axis depends on which conditions are kept constant and which changed (temperature, salts, ligand concentration, etc..). Under the solubility curve, no crystals or precipitates occur. Obviously the precipitation zone does not form crystals because precipitation is faster than crystallization. Crystals will appear when the supersaturated region is reached, so when the protein concentration exceeds solubility. The energetic barrier to overcome, called nucleation, can take time and it is the energy required to create nuclei, microscopic protein clusters, starting point for the crystal growth.22 In the metastable zone the nucleation is too slow, so crystals will need long time to be formed. Indeed in the labile zone (or nucleation zone) the nucleation is spontaneous, but if the supersaturation is too large aggregates and

precipitates are created.

4.1.3 Target proteins

As anticipated in chapter 1, protein-protein and protein-ligand interactions are at the base of many pathological and physiological processes. For this reason these macromolecules are important targets for the development of new bioactive compounds. In some cases, these could be designed to restore functions that a protein lost because of genetic or post-translational alterations. In others, on the contrary, they could be disruptors or inhibitors of protein-protein and protein-ligand interactions if these are at the origin of diseases. In this part of the PhD work, performed during a period spent in the group of Dr Crowley at the National University of Galway (Ireland), we studied the ability of some multivalent calixarenes in the binding to protein surfaces, side chain residues and recognition sites, by using NMR and crystallization methods.

In the next paragraph is reported the description of the proteins selected for the studies, i.e. immunoglobulin-binding domain B1 of streptococcal protein G (GB1), Flavodoxin (Fld), alpha synuclein (α-syn) and a lectin from *Ralstonia solanacearum*. All the proteins, not commercially available, were over-expressed and purified.

4.1.3.1 Immunoglobulin-binding domain B1 of streptococcal protein G (GB1)

Among model systems for protein folding and design one very common is immunoglobulin G (IgG)-binding domain B1 of streptococcal protein G (GB1). GB1[23] is a small single-domain protein (only 56 amino acids), with a molecular weight of approximately 6 kDa, composed by one α-helix, four-stranded β-sheet and two hairpins (figure **4b**). This is an acid protein, since the total number of negative residues Asp/Glu is 10 against 6 for positively charged residues Lys/Arg. This protein structure is well characterized by NMR[24], X-ray diffraction[25] and it is an excellent model protein. This feature is due to many factors such as its high stability and solubility, the occurrence of both α-helical and β-sheet character. Its thermodynamic properties are also characterized,[26] it does not contain cysteines and it has a tryptophan residue that is buried in the folded state. Normally the wild type GB1 is not used but

mutants are created to have higher stability in drastic conditions. Between these derivatives the so called GB1-QDD[27] with mutations T2Q, N8D, and N37D is largely used. The T2Q mutation is introduced to prevent processing of the N-terminal methionine.[28] As reported by Lindman and co-workers,[27] GB1-QDD is stable at several values of pH and has the same structure of the wild type GB1, hence comparisons among the two mutants are valid. Due to introduced negative charges at residues 8 and 37, this protein will have a higher net charge than the wild type at high and intermediate values of pH and affecting the stability. In this work GB1-QDD will be studied and for simplicity the protein will be reported only as GB1.

4.1.3.2 Flavodoxin (Fld)

Flavodoxins[29,30] (Fld) are a family of proteins involved in electron-transfer, that is the main biological function of them. They are implicated in a variety of photosynthetic and non-photosynthetic reactions in anaerobic and aerobic bacteria[31,32], whereas, in eukaryotes, a variation of the flavodoxin gene helps build multidomain proteins. They are characterized by the presence of non covalently (but tightly) bound flavin mononucleotide cofactor (FMN) as a redox active group. In fact the FMN in flavodoxin can exist in three different oxidation states: oxidized, semiquinone (one electron reduced) and hydroquinone (two-electron reduced). Flavodoxins were discovered in the 1960s in cyanobacteria and clostridia growing in low-iron conditions, where they replaced the iron-containing ferredoxin in reactions leading to $NADP^+$ and N_2 reduction, respectively and they are fundamental in some organisms.[33] For that reasons flavodoxins were characterized for their interactions with partners such as ferredoxin(flavodoxin)-$NADP^+$reductase (FNR)[34],for physicochemical properties, including their peculiar redox potentials, the reversible removal of FMN, and the three-dimensional structures. Flavodoxins were discovered to be important as electron donors in the reductive activation of anaerobic ribonucleotide reductase[32], pyruvate formate lyase, biotin synthase[35] and cobalamin-dependent methionine synthase[34].

Furthermore some flavodoxins have found practical applications in different fields, from the determination of iron deficiencies in phytoplankton or the design of mixed protein/metal nanostructures[36,37], while those from pathogenic bacteria are being investigated as potential drug targets.[38]

All the well-known flavodoxins are highly acidic proteins and are constituted by a number of amino acids from 140 to 180. They can be divided in two groups, long chain and short chain flavodoxins, depending on the presence or not of 20 residues of an unknown function. Except to some algae, Fld are present in bacteria, Particularly in this thesis the attention was focalize on Fld from *Escherichia coli,* characterized by the presence of 175 residues with a molecular weight of approximately 19 kDa[39]. It belongs to the class of long-chain flavodoxins which are characterized by a loop inserted in the fifth β-strand (Figure **4a**). Even this protein structure was well characterized by NMR spectroscopy[40] and X-ray diffraction.[41]

The recent interest of scientists in protein folding and protein-ligand recognition has stimulated the use of flavodoxin as a protein model. Effectively since flavodoxin is essential for the survival of some human pathogens [42,43] could make it a drug target on its own.

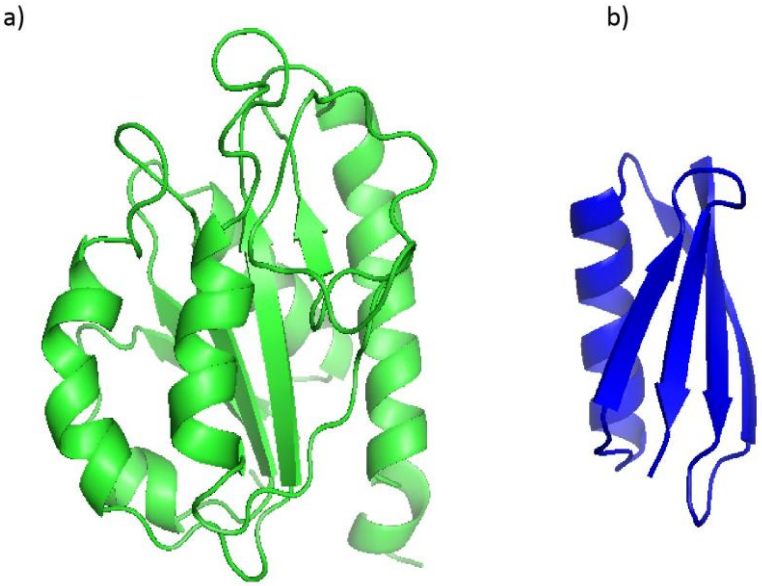


Fig.4: schematic representation of a) Fld and b) GB1 (made by Pymol).

4.1.3.3 α-Synuclein (α-syn)

Another important class of protein are synucleins[44], small and soluble proteins expressed primarily in neural tissue and in certain tumours. The family includes three known proteins: α-synuclein, β-synuclein, and γ-synuclein with a range of amino acids from 127 to 140 in length and they are 55-62% identical in the sequence, with a similar domain organization. The name is due to the discovery of the protein in both *syn*apses and in the *nucl*ear envelope.[45] Further studies have confirmed the presence of synucleins in nerve terminals,

but have failed to confirm a nuclear localization, but the name was kept. Among them α-synuclein (α-syn) is one of the most important. It is predominately expressed in brain, particularly in the neocortex, cerebellum and thalamus. [46]

α-syn is a small protein (14 kDa, 140 aa) expressed at high levels in nervous tissue. It is an acidic protein that has also amphipathic domain and non-amyloid-β component (NAC). The studies of this protein have gained the interest of scientific community because is correlated to neurodegenerative diseases such as Alzheimer's disease (AD), Parkinson's disease[47] (PD), Huntington's disease (HD), and amyotrophic lateral sclerosis (ALS).[46] α-syn is also the primary fibrillar component of Lewy bodies, and α-syn lesions are also observed in cases of dementia with Lewy bodies and multiple system atrophy and Hallervorden-Spatz disease. In fact it is in the class of IDPs (Intrinsically Disordered Proteins), (figure **5**) so it can adopt every conformation ranging from collapsed to fully extended, playing roles in signalling and transcription.

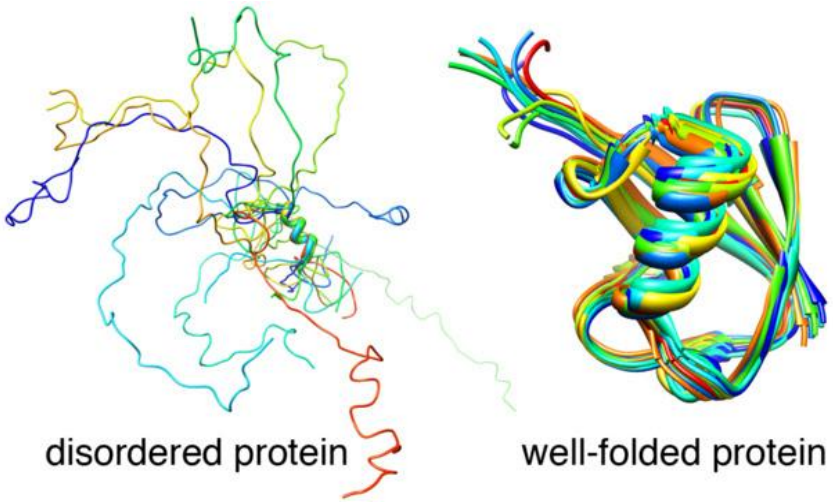


Fig. 5: Structures of an IDP, such as alpha-synuclein, and of a well-folded protein.

For these reasons this protein has a biomedical importance since can be considered as a biomarker for neurodegenerative diseases.[48]

4.1.3.4 Ralstonia solanacearum lectin (RSL): CBP

As well-explained in chapter 2, lectins are CBPs and they are very important biological targets. Among bacterial lectins, *Ralstonia solanacearum* lectins are from this bacteria distributed worldwide and that causes lethal wilt in many agricultural crops. In fact *Ralstonia*

solanacearum is a soil-borne plant pathogen and belongs to the family of β-proteobacteria. The bacteria invades the plants through the roots, which requires usually wounds or natural openings. Then they penetrate the xylem and spread through the plant.[49]

R. solanacearum produces three different soluble lectins, namely RSL (9.9 kDa), RS-IIL (11.6 kDa) and RS20L (20 kDa). However in this chapter the attention is focalized on the first protein: RSL. This is a small protein of 90 amino acids with a tandem repeat in its amino acid sequence and it is a potent L-fucose-binding lectin,[50] having high affinity for fucose residues, especially for α-Fuc1–2Gal and α-Fuc1–6Gal linkages. Especially the sugar affinity order is the following: L -fucose > L -galactose > D -arabinose > D - fructose > D –mannose.[51]

The 3D-structure of this protein is well-known and analysis of RSL crystals shown a symmetrical trimeric complex, forming a so called six-bladed β-propeller, with each monomer built out of two small four-stranded anti-parallel β-sheets (see figure below). The resulting hexavalent trimer has the capacity to invaginate lipid membranes containing fucosylated glycolipids.[52]

Fig. 6: Cartoon diagram of RSL trimer (prepared in Pymol)

Moreover, RSL was co-crystallized with different sugars (α-methylfucose and 2-fucosyllactose). The co-crystallization with α-methylfucose revealed the occurrence of two sugar binding pockets in each monomer: one centred in the monomer and the other at the interface where two monomers can meet, resulting in six binding sites in total for a trimer.[50]

Since it is a CBPs with biological application the possibility to find ligands with higher affinity than L-fucose that can inhibit the native interactions has raised the attention against plant diseases.

4.1.4 Studied ligands

The ability of calixarene-based ligands to bind to proteins was already observed and studied in the group of Prof. Peter B. Crowley.[53,54,55] Therefore the biomacromolecules previously described were studied with some macrocyclic ligands, prepared in the laboratory of Prof. Sansone and Casnati in Parma.

The ligands were selected and designed on the basis of the chemical properties of these proteins. Particularly, derivatives belonging to two classes of calixarenes were studied: the guanidino- and the glycocalixarenes.

The first aim of the work was to explore the binding abilities of a small series of cationic multivalent ligands, the guanidino-calixarenes shown in figure **7**, towards the two negatively charged proteins GB1 and Fld (figure **8**).

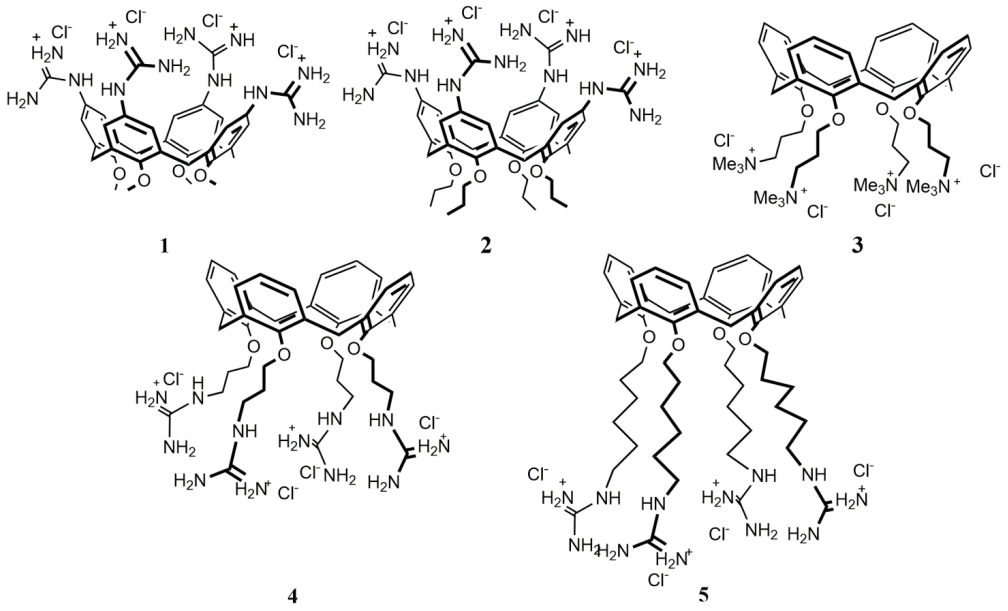


Figure 7: The chloride salts of the five cationic calix[4]arenes used for the binding studies with GB1, Fld and α-synuclein

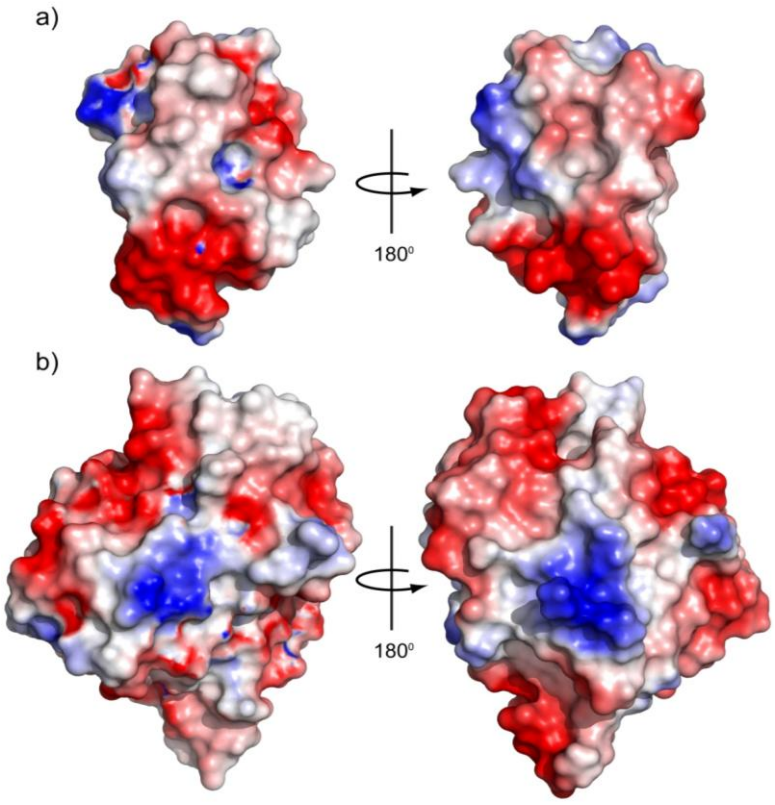


Figure 8: Electrostatic surface potential of a) GB1 and b) Fld with blue and red areas corresponding to regions of positive and negative potential, respectively. Their acidic character suggested their studies with positively charged ligands such as guanidino-calixarenes.

The different ligands **1**, **2, 4**, and **5** were selected on the basis of the well-known binding ability of the guanidinium units towards anionic residues that in proteins can be the Asp and Glu side chains. Moreover, in the four molecules the conserved guanidinium units have a different spatial position and distance respect to the aromatic cavity of the macrocycle which could play an additional role in the binding event. The cationic groups have also a different freedom degree, being more rigid when directly linked to the aromatic units in **1** and **2** and more mobile when linked at the lower rim through aliphatic chains in **4** and **5**. Furthermore, while calixarenes **2**, **4** and **5** are blocked in the cone geometry, the analogue **1** is conformationally mobile so that, while the formers are more preorganized, the latter can adapt itself during the complexation. Eventually, derivative **3,** analogue of **4,** was included in the studies to compare quaternary ammonium and guanidinium in the binding process.
The same cationic calix[4]arenes were used in binding studies with α-syn.

The second aim of the work was the study of the interactions between glycocalixarenes and

the microbial lectin from *Ralstonia solanacearum* (RSL), a plant pathogen.[50] Since this protein is characterized by the presence of six binding sites selective for L-fucose, the glycocalix[6]arene **6** bearing L-fucose was designed for the interaction with this lectin. A galactosylcalixarene (**7**) was prepared as negative control, assuming that this monosaccharide cannot interact with that lectin .

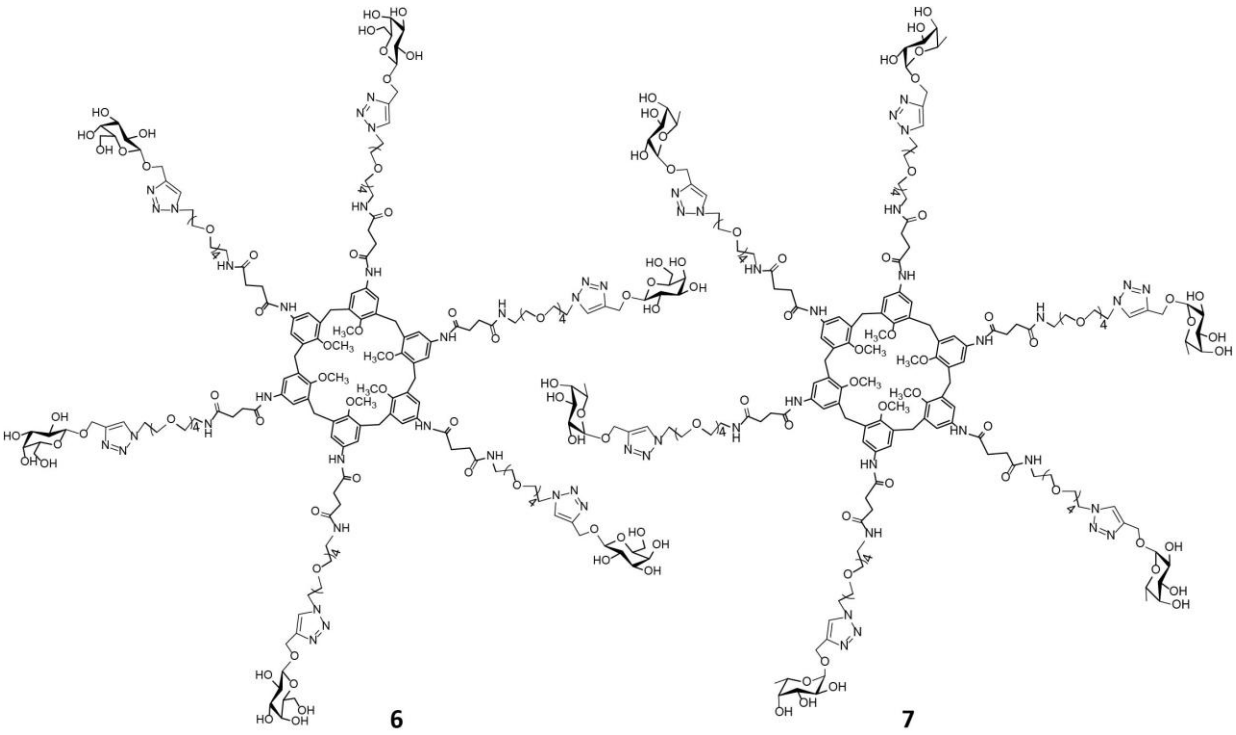


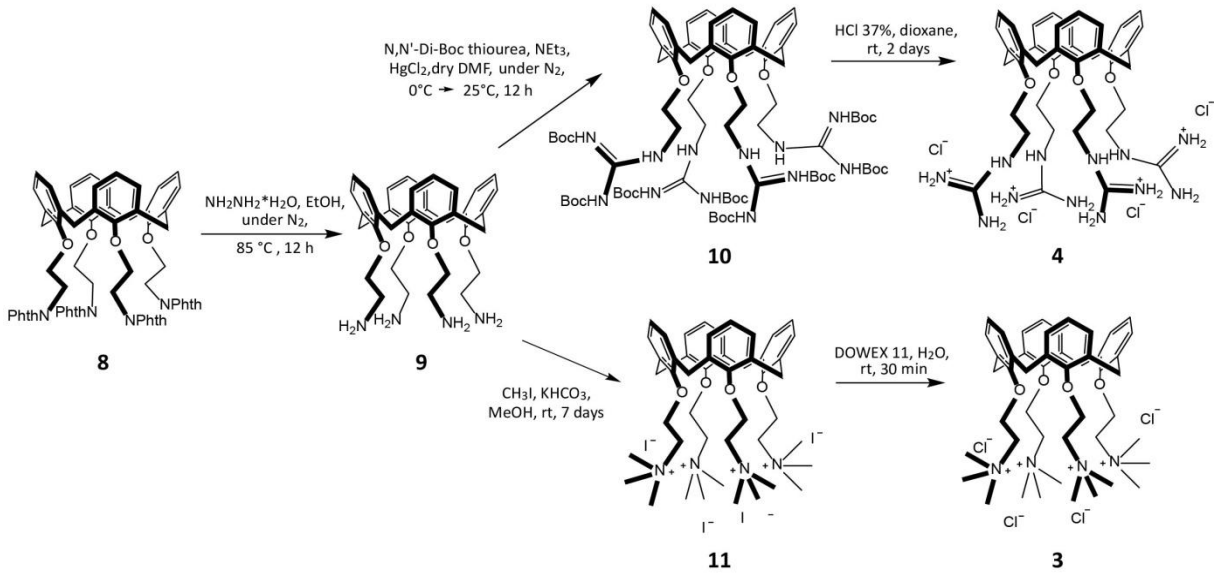
Figure 9: The two glycocalix[6]arenes studied with RSL bearing six units of a) L-fucose (**6**) and b D-galactose (**7**).

4.2 Results and discussion

4.2.1 Synthesis of guanidinocalix[4]arenes

The guanidino derivatives **1** and **2** were already available in the Parma laboratory, while **4** and **5** and their analogue **3** were synthesized as reported in literature.[56,57]

The synthesis of calixarenes **3** and **4**, reported in the scheme **1**, started from the same intermediate **8.**



Scheme 1: synthesis of derivatives **3** and **4** from intermediate **8**

The removal of protecting phthalimide groups was carried out using hydrazine in EtOH, being careful to do not expose too long product **9** to the air. In fact it was observed that lower rim tetraamino calix[4]arenes as **9** are able to rather efficiently absorb two molecules of CO_2 per calixarene, yielding self-assembled dimers, having high stability.[58] As said, compound **9** was obtained and used as intermediate to prepare calixarenes **3** and **4**. The guanidinium calixarene **4** was obtained from amino calixarene **9** with two steps: guanidinination of the lower rim using N,N'-di-Boc-thiourea in dry DMF and removal of the Boc protecting groups with 1,4 dioxane and concentrated HCl (37%). The tetramethylammonium compound **3** was obtained by methylation of amino groups using CH_3I in MeOH affording derivative **14** that was subsequently treated with DOWEX 11 resin (Cl^- form) to exchange the anion and give the final compound **3** in the form of chloride salt.
Below the ^{1}H-NMR spectrum of guanidinocalixarene **4** is reported, where the absence of Boc signals around 2-1.5 ppm confirmed the achievement of the desired product.

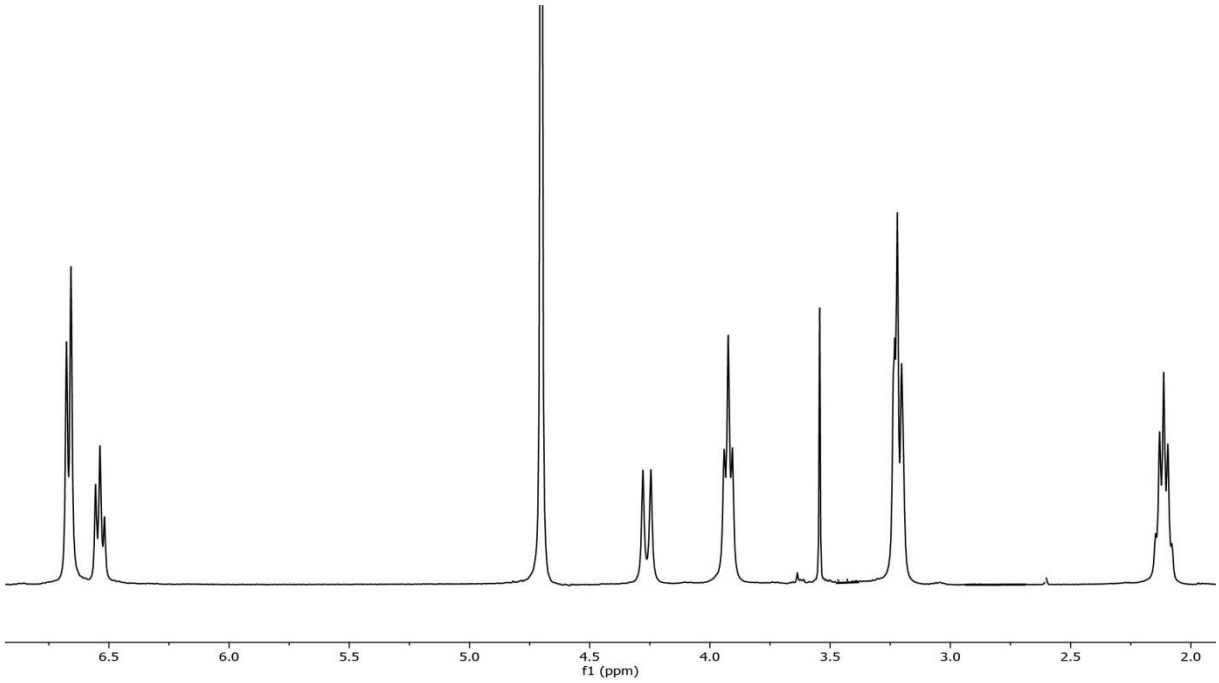


Fig. 10: ^{1}H NMR spectrum of deprotected guanidino calixarene **4** (400 MHz, D_2O)

4.2.2 NMR binding studies

It has been established that calixarenes functionalized with the same groups but in distinct conformations or with the binding units linked at different positions of the macrocyclic platform, can interact with different affinities and high selectivity with the same biomacromolecules, even included in the same family and class.[59,60,61] This the reason why we decided to investigate the properties of the small family of guanidino calixarenes reported above in the recognition of the three negatively charged proteins, GB1, Fld and α-syn. As briefly described before, all these derivatives are fixed in a cone geometry except **1**, conformationally mobile, due to the presence of the small methyl groups at the lower rim.

Protein-ligand titrations were performed by adding µl aliquots of each ligand to the ^{15}N-labelled proteins. For GB1 and Fld monitored by ^{1}H-^{15}N HSQC spectroscopy, the effect of the interaction with the five ligands are summarised in **Table 1**. Calixarenes **2** and **5** caused precipitation of GB1 and Fld, suggesting that interactions were occurring but preventing characterization by NMR spectroscopy. It would appear that the more marked hydrophobicity of these compounds (with lower rim propyloxy or hexyloxy substituents) tend to give rise to aggregation phenomena or to more hydrophobic complexes and then precipitation of the proteins. With Fld protein, also ligand **1** resulted in complete precipitation showing a rather significant difference respect to its behaviour with GB1 for

which the calixarene showed a weak interaction (K_d ≈ 1 mM). The protein-calixarene interactions that were amenable to NMR analysis were observed to be in fast exchange (on the NMR timescale) between the ligand-free and -bound states.

	Ligand				
Protein	**1**	**2**	**3**	**4**	**5**
GB1	Δδ (≈1 mM)	ppt	Δδ (≈800 µM)	Δδ (≈80 µM)	ppt
Fld	ppt	ppt	Δδ (≈600 µM)	Δδ (≈1.2 mM)	ppt

Table 1: ligands studied with Fld and GB1. Δδ indicates that the complex was suited to NMR analysis, while ppt indicates precipitation. In the bracket are reported the K_d of the complexes, obtained from NMR data.

With α-syn only preliminary titrations were carried out, where **2**, **4** and **5** gave precipitation and spectra started to become weaker and going beyond detection. Only **1** was the only interesting ligand for this protein.

4.2.2.1 GB1-calixarene binding

Compounds **1**, **3** and **4** were found to bind GB1 and form soluble complexes. As shown in figure **11**, all three ligands resulted in chemical shift perturbations of the GB1 backbone resonances.

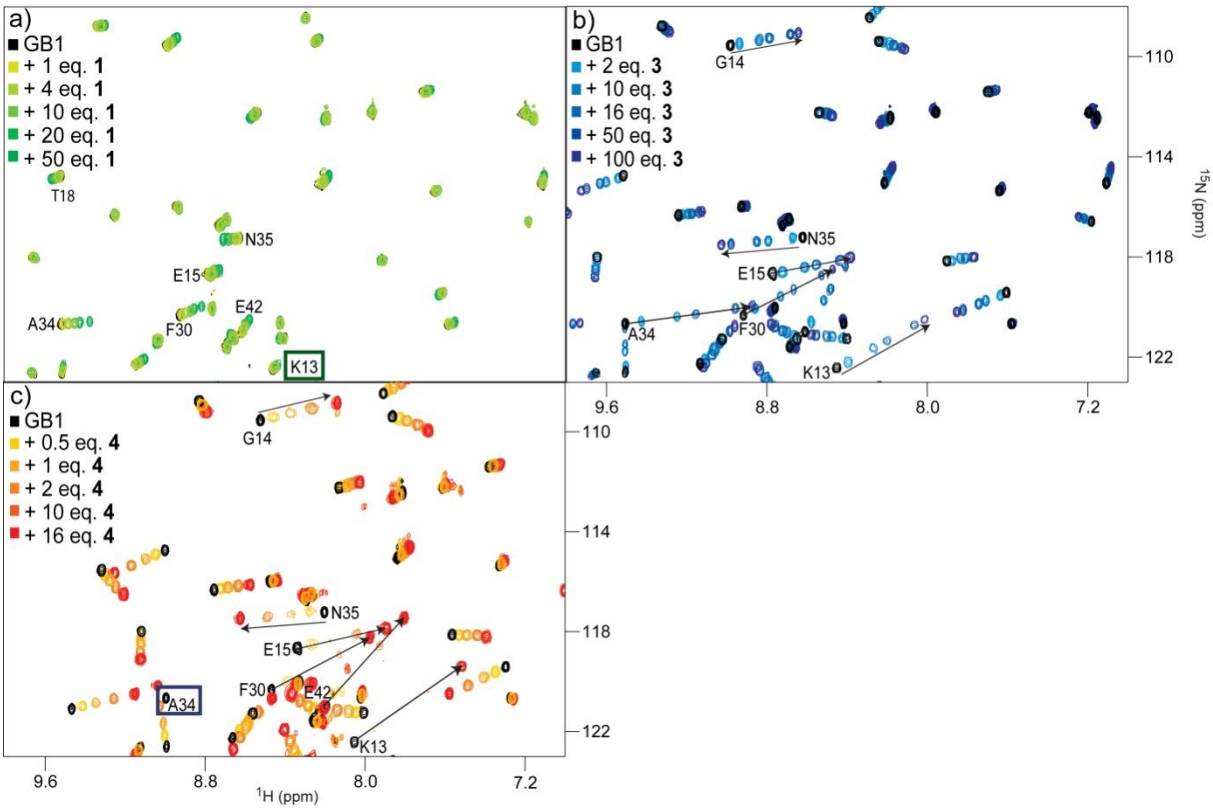


Figure 11: Spectral region from the overlaid ^{1}H-^{15}N HSQC spectra of GB1 in the presence of increasing concentrations of **1**, **3** and **4**. The reference spectrum of 0.05 mM pure GB1 is shown as black contours, while successive titration points are coloured as indicated. K13 for **1** and A34 for **4** are highlighted.

Especially we observed an increase in the Δδ going from **1** to **3** and then to **4** and the most affected residues were G14, E15, F30, A34 and N35 (figure **11**), but the extents of the shifts of each amino acid were quite different between these macrocycles, especially upon the titration with **4**. Some significant differences were observed, for example for K13 that did not undergo remarkable variation with ligand **1**, while it was more affected by **3** and **4**. This could mean that a cation-π interaction was taking place with the available electron rich cavity of the two lower rim substituted macrocycles.[55] In **1** in fact three factors can prevent the inclusion of the ammonium side chain in the apolar cavity: i) the repulsion between positive charges of ammonium and guanidinium units, ii) the hindered access if the ligand, conformationally mobile, is arranged in the cone geometry, iii) the absence of a cavity if on the contrary the ligand is arranged in partial cone or alternate conformation. To this end, it can be useful to say that water soluble conformationally mobile calix[4]arenes with a certain amphiphilic character as **1** tend to assume in aqueous environment the 1,3-alternate geometry. CSPs for E42 were observed only with the guanidinocalixarenes **1** and **4**, but not

with the analogue **3** bearing the quaternary ammonium groups, suggesting that the glutamate has an higher affinity for the guanidinium moieties. Probably this contribute to the lower Kd obtained for **4** (80 μM) respect to **3**.

Following the recognition process by ^{1}H-^{15}N HSQC, the pattern of CSPs, in term of direction, was similar with calixarenes **1**, **3** and **4**, suggesting that the GB1 interaction with these derivatives is based on similar recognition events, apart the interaction involving K13. In addition to CSPs, in the case of calixarene **4** the binding resulted in line broadening effects too. For examples, A34 signals broadened beyond the possibility of detection immediately after the first addition of the ligand (0.025 mM, 0.5 eq) suggesting that this residue is strongly involved in the interactions. E15, E42, as well as N35 signals started to become weaker in the middle of the titration at 0.5-1 eq of ligand but they returned more intense and sharper above these amounts at the last points of the titration. Furthermore, K13, F30 disappeared at certain point of titration (1 eq), to emerge back at higher concentration of the ligand and with a changed chemical shift. This behaviour should be interpreted as a first protein-ligand recognition event, followed by a second binding, that has a larger affinity compared to the first one.[62,6] On the other side for calixarenes **1** and **3** peaks were clearly detectable along the whole titration, not broadening or disappearing, indicating one single binding event.

The previous concepts are also visible in the CSP maps, reported in figure **12** for compounds **1**, **3** and **4** at the same ligand concentration and the same Δδ cut-off. These allowed to better visualize which parts of the protein sequence are more affected by the ligand-protein binding. The profile is common to all three calixarenes and three regions of the primary structure appear more significantly modified by the presence of the ligands including residues 6-18, 30-44 and 52-56, even if, as already said, with strongly different extent for the different macrocycles (Figure **13**).

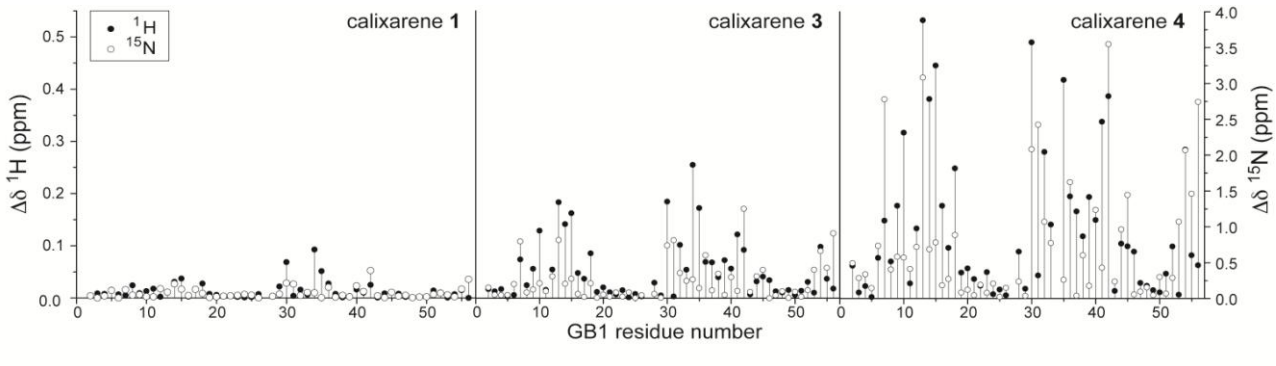


Figure 12: Plot of the chemical shift perturbations measured for GB1 backbone amides at pH 6.0 in the presence of 10 equivalents of compounds **a) 1**, **b) 3**, and **c) 4**. Blanks correspond to: E27 and A34 for calixarene **4**.

This similarity can be interpreted as a preference of the ligands for specific parts of the GB1 backbone. On the other hand, the differences in the $\Delta\delta$ size, being constant the points of contact, are a first element supporting the idea of a different strength for the three complexes. Most of the residues affected by these multivalent ligands are situated in the area of GB1 with a negative charge density (compare Figures **8** and **13**), like E15 and E42 reasonably interacting with the ligands through charge-charge contacts.[54] Also neutral residues (G14, G41) are involved in some way in the recognition process[63] and CH-π and π-π interactions can take place (A34, F30).

Analysing the residues that underwent highest Δδ, it is possible to appreciate better how the calix[4]arenes bound to the protein (Figure **13**).

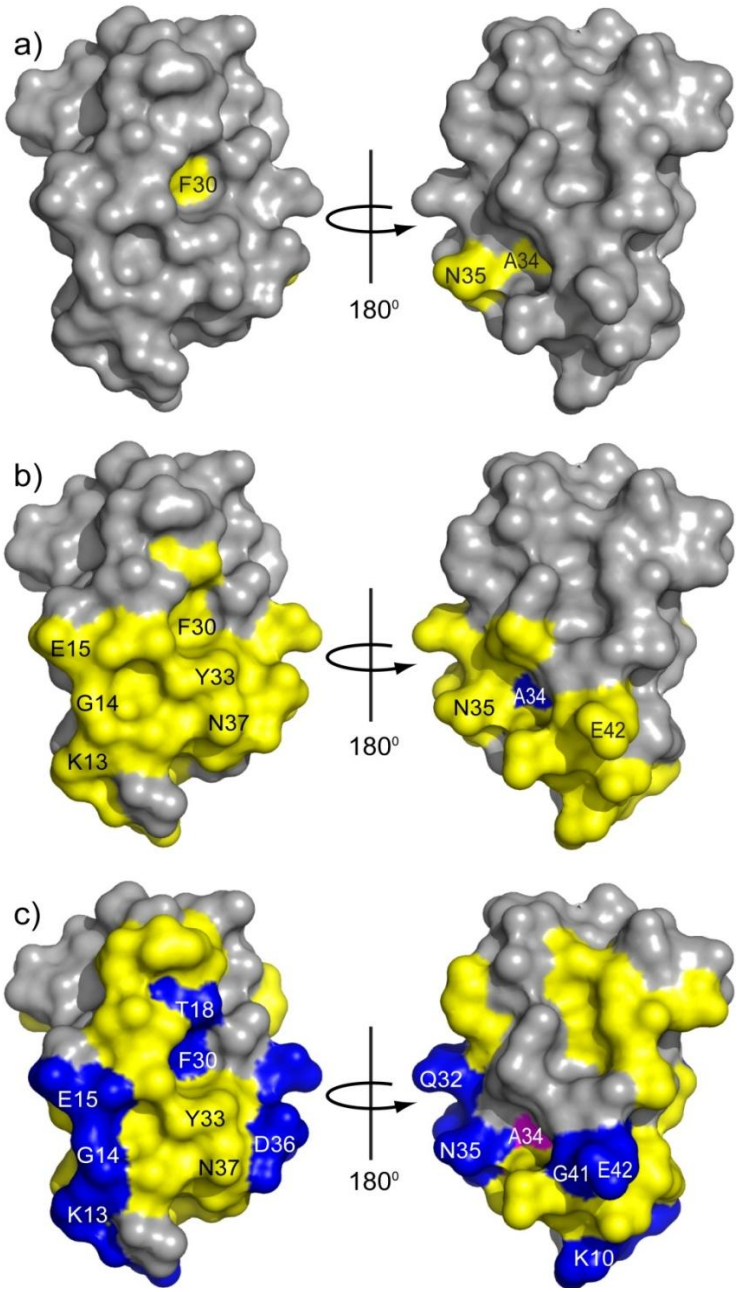


Figure 13: Chemical-shift perturbation map of GB1 in the presence of 10 equivalents of a**) 1**, b**) 3**, and c**) 4**. The molecular surface was generated in PyMOL by using the crystal structure of GB1. Residues for which the amide resonances experienced large ($\Delta\delta_H$ > 0.2 ppm or $\Delta\delta_N$ > 2ppm), or small ($\Delta\delta_H$ > 0.05 ppm or $\Delta\delta_N$ > 0.5 ppm) shifts are colored blue and yellow, respectively. Resonances that broadened beyond detection (A34 with ligand **4**) are magenta.

F30 is located in the core of GB1, but it is strongly affected with all the three ligands. Since in **1** an hydrophobic cavity is indeed not necessarily available because of its conformational mobility, the occurring interaction could be also a cation-π interaction between the aromatic

residue of the amino acid and the positively charged guanidinium and trimethylammonium groups o the macrocycles. In fact it is well known that cone calix[4]arenes have the capability of trapping positively charged groups and side chains of amino acids.[64] This type of interaction could actually take place with reversed roles between K13 and K10, well exposed from the protein surface, and the aromatic nuclei of the ligands, without necessarily considering the inclusion in the cavity of the cation group, but only with an approach to the exterior surface of the macrocycle.[65,66] a

As said before, the cross peak relative to A34 disappeared immediately upon addition of **4**. This was a weird perturbation because the accessible surface area of this residue is particularly small (as marked in figure **13**). With **1** and **3** the peak relative to this residue was present for the whole titration, suggesting for **4** a strong binding of this derivative. In this hypothesis A34 could be affected by the interaction between **4** and the vicinal N35 well exposed to the external environment (see figure **13**). The differences observed in the CSPs of this amino acid with the two ligands **3** and **4** both blocked in the cone geometry and not having an anionic counterpart in that point of the sequence suggest that rather complicated events are occurring.

To this end and more in general, it is also evident that the largest surface of GB1 was affected by the presence of ligand **4** with a decrease going to **3** and then to **1.** The same trend is in this case observed for the extent of the chemical shift variations (figure **12**) and for the complex strength (Table 1). The Kd values in fact that could be calculated from the titration experiments indicate calixarene **4** as the most efficient ligand among the three, even if, on the other hand, it seems to lack selectivity for a region of the protein surface or a class of residues. This could be interpreted as evidence of a non-selective recognition process. The binding isotherms (figure **14**) yielded hyperbolic curves for all three compounds. The fit was good using a one-to-one binding model for all. Other types of fitting were tried, especially for **4**, where the curves seem less appropriated. With **4** the goodness of the fitting is in fact lower, due to some residues (K13, F30) disappearing beyond detection or starting to become weaker at certain ligands concentration. The higher affinity of **4** for GB1 (saturated at approximately 6 equivalents) could be then attributed to the guanidinium units respect to the quaternary ammonium groups of **3** (saturation at 80 equivalents), even if their more effective binding ability is accompanied by a lack of selectivity. The cone geometry of both seems to help respect to **1** in the recognition process. It is rather surprising

that **1**, lacking the cavity and then having substantially only the guanidinium units for binding does not interact with glutamate carboxylates.

This order of affinity highlighted that the different orientation of the ligating unit plays a role in determining the binding behaviour of the calixarene based multivalent ligands to a protein as previously shown with this class of macrocycles.[60,67]

Considering the data, for the control ligand **3** the binding affinity ~10-fold lower than **4**, indicates that the presence of guanidinium moiety had a role in the recognition process since the absence of this functional group decreases the affinity towards GB1. Instead the derivative **1** did not show a good affinity for GB1 (higher K_d) although bearing the same group of **4**. Unfortunately, the impossibility of determining a Kd value for ligands **2** and **5** avoids to deduce if, given the cone geometry, it is better to have guanidinium units at upper or at the lower rim or about the possible effect of the distance between the macrocyclic platform and the cationic units, given the presence of these at the lower rim.

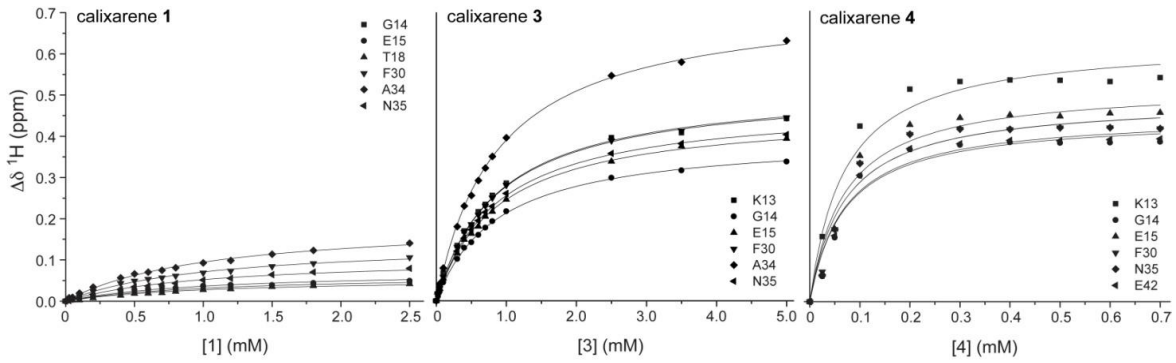


Figure 14: Binding isotherms for GB1 binding to ligands **1** ($K_d \approx$ 1mM), **3** ($K_d \approx$ 800μM), and **4** ($K_d \approx$ 80μM), were fit to a 1:1 binding model. In **4** there is not the data at 0.5 mM for K13 because the resonance disappeared beyond detection.

4.2.2.2 Fld-Calixarene Binding

In the binding studies with Fld protein only ligands **3** and **4** allowed a complete titration without precipitation phenomena and they showed a reversed order of affinity respect to that found with GB1. In fact **4** did not induce significant chemical shift perturbations when added to Fld, suggesting that the interactions were very weak, in agreement with the calculated K_d value (~1.2 mM). The very poor affinity was also at the origin of the scarce

reproducibility of the binding curves in different titration experiments in term of order of affected residues. The ammonium derivative **3** with a Kd ≈600 µM resulted more efficient than **4** and also more efficient in the binding to Fld than to GB1. In the ^{1}H-^{15}N HSQC maps (figure **15**) it is possible to observe CSPs, not for almost all the residues as for GB1, but only for a 10% of resonances, indicating that the recognition process is more selective.

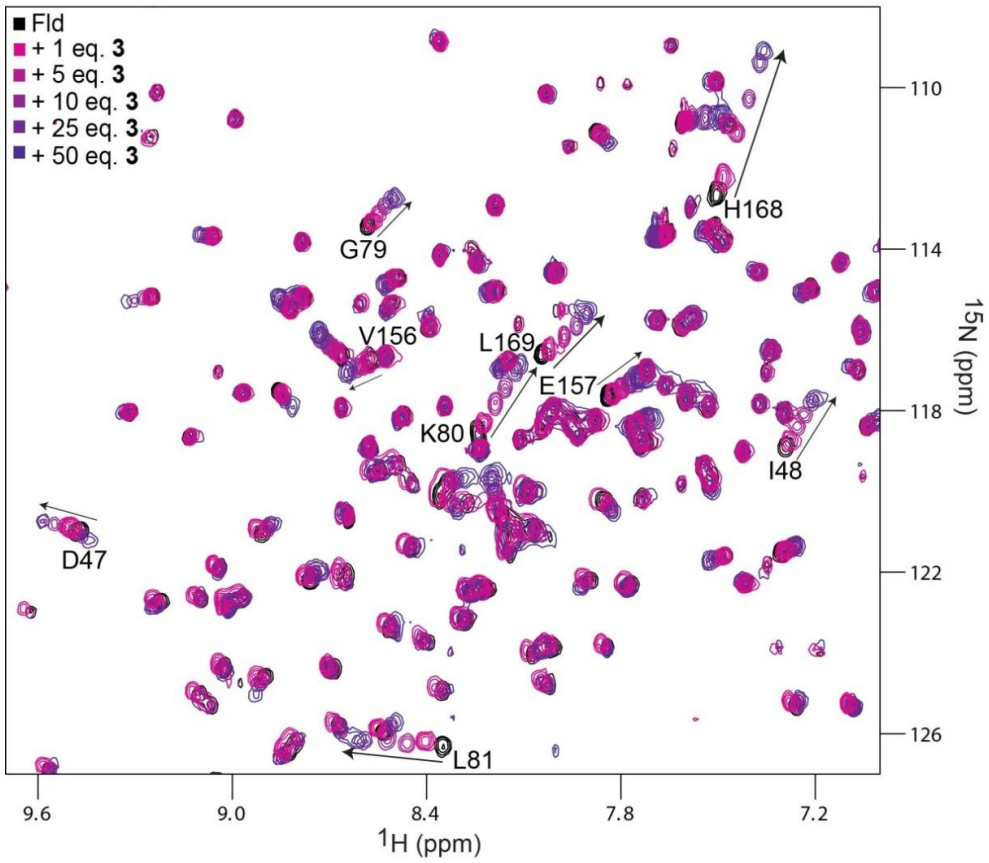


Figure 15: Spectral region from the overlaid ^{1}H-^{15}N HSQC spectra of Fld (0.1 mM) in the presence of increasing concentrations of **3**. The Fld resonance (0.1 mM) is shown as black contours, while successive titration points are magenta to purple. Arrows indicate the direction of the peak shift throughout the titration. The resonances with $\Delta\delta_H > 0.05$ ppm or $\Delta\delta_N > 0.5$ ppm are assigned.

This behaviour is more evident from the plot in Figure **16b.** Only some residues are affected by the presence of the ligand and they identify a specific region of the protein. E47, E157, H168, K80, L81 and L169 were the residues most involved in the binding with Fld. Furthermore, observing the perturbation map (figure **16a**) residues such as L81 E157, H168 are well exposed to the external environment, while K80 and L169 more hidden. It is well

evident that there is mainly one part of the protein able to interact with the calixarene **3**, that is the C terminus. This is a further confirmation that Fld can interact with **3** differently from GB1 and the recognition process is more selective using this more acidic and higher MW protein. The binding is based again on a mix of electrostatic, cation-π, CH-π interaction, but apparently charge-charge interactions would seem less important respect to the complexation process with GB1. In fact, ligand **3** with the ammonium units is more efficient than **4** equipped with guanidinium, known indeed to be a better anion binder. Analysing the potential surface of Fld (figure **16a**) only aspartate/glutamate residues (D31, E42, D47 and E157) recognized the ammonium moieties with charge-charge interactions, while K80 with a cation-π interaction. This residue is hidden to the outside but could be affected by the proximity of L81 as well as L169 with H168. Mostly involved for Fld are hydrophobic amino acids (V32, I48, L81, V156, L169), likely with CH-π or van Der Waals interactions with the scaffold cavity and the linker, respectively. The only binding curves reported (figure **16c**) are for ligand **3** for which, one more time, the fit using a 1:1 binding model was the best choice. The K_d (~600 μM) is lower compared to the value for the same compound with GB1, indicating that the types of interactions in which **3** was were stronger. It is interesting also to note that ligand **1**, determining precipitation of the protein, seems able to interact significantly with Fld while it showed a low affinity for GB1.

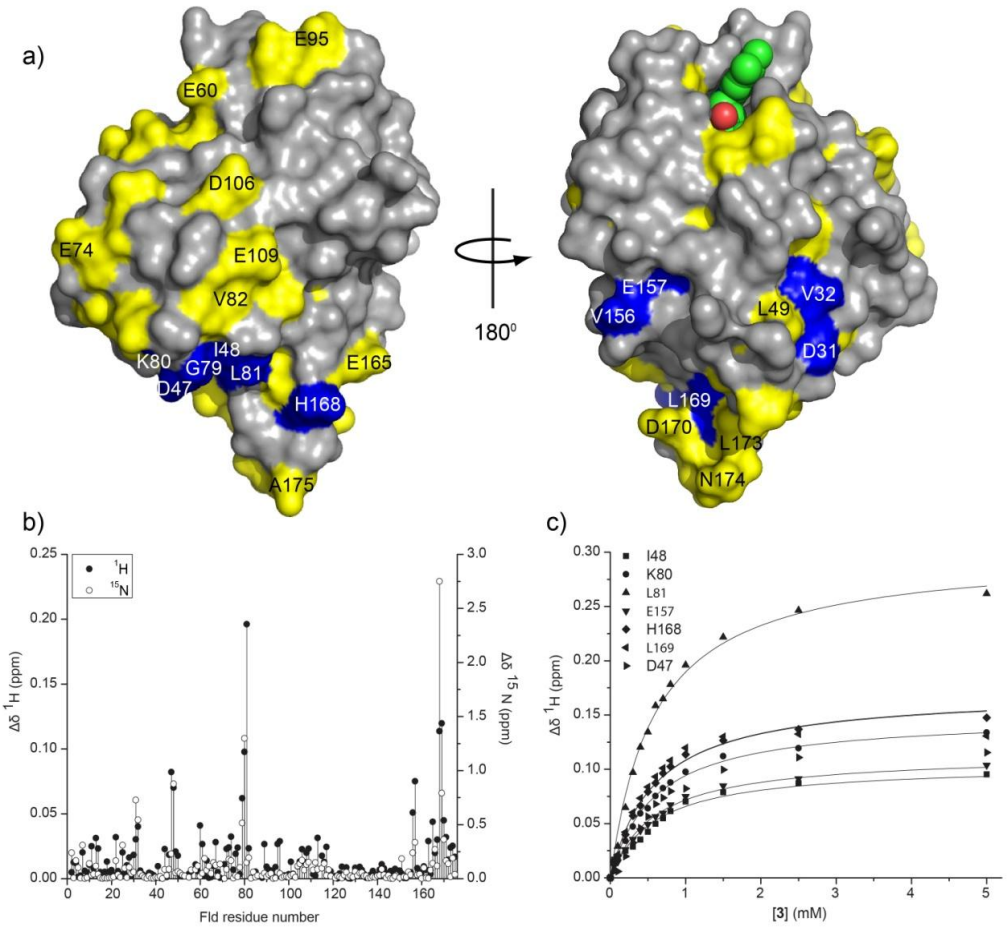


Figure 16: a) Chemical-shift perturbation map of Fld in the presence of 10 equivalents of **3**. The molecular surface was generated in PyMOL by using the crystal structure of Fld.[41] Residues for which the amide resonances experienced large ($\Delta\delta_H$ > 0.05 ppm or $\Delta\delta_N$ > 0.5 ppm), or small ($\Delta\delta_H$ > 0.02 ppm or $\Delta\delta_N$ > 0.2 ppm) shifts are coloured blue and yellow, respectively. **b)** Plot of the chemical shift perturbations measured for the Fld backbone NH amides at pH 6.0 in the presence of 10 equivalents of compound **3**. **c)** Binding isotherms for ligand **3** (K_d ≈ 600μM), with fits to a 1:1 binding model.

4.2.2.3 Binding studies on cationic calixarenes and α-synuclein

The data collected for this protein were preliminary since only two-three points of titration for each ligand were acquired. As it is possible to observe the reference spectrum of alpha synuclein (figure **17** on the left) is completely different from the previously studied proteins. The resonances are well resolved and sharp, but with a limited dispersion of chemical shifts. This behaviour reflects the unfolded nature and the high degree of backbone mobility [68] of

the native protein. It could adopt a β- or α-stranded conformation when ligand binding occurs and so the spectrum become more similar to well-folded proteins such as GB1.

The overlapped spectra (figure **17** on the right) revealed **1** as the only interesting ligand that gave CSPs for few residues. Particularly V37, A69, I112, D119, N122, S129, G132, D135, and E137 were the most affected ones. No broadened or new resonances appeared. On the contrary, the observed changes of the chemical shifts were linear, monotonic functions of the calixarene concentration. This means that fast exchange takes place.

With these first results it is possible to say that for **1** the interaction with the protein involved both aliphatic residues and negatively charged residues. This behaviour can be compared with those observed for the previous proteins. In fact even in this case electrostatic interactions occurred, with D119, D135, and E137, demonstrating the involvement of the guanidinium groups in the recognition process. Furthermore CH-π interactions with V37, A69, I112 could take place. Rather surprisingly the other calixarenes bearing guanidinium groups (**2**, **4** and **5**) did not give significant CPSs,

However these are only very preliminary and partial data that must be deepen with further studies performing whole titrations for each ligand and treating the experimental data to determine K_d values for the complexes.

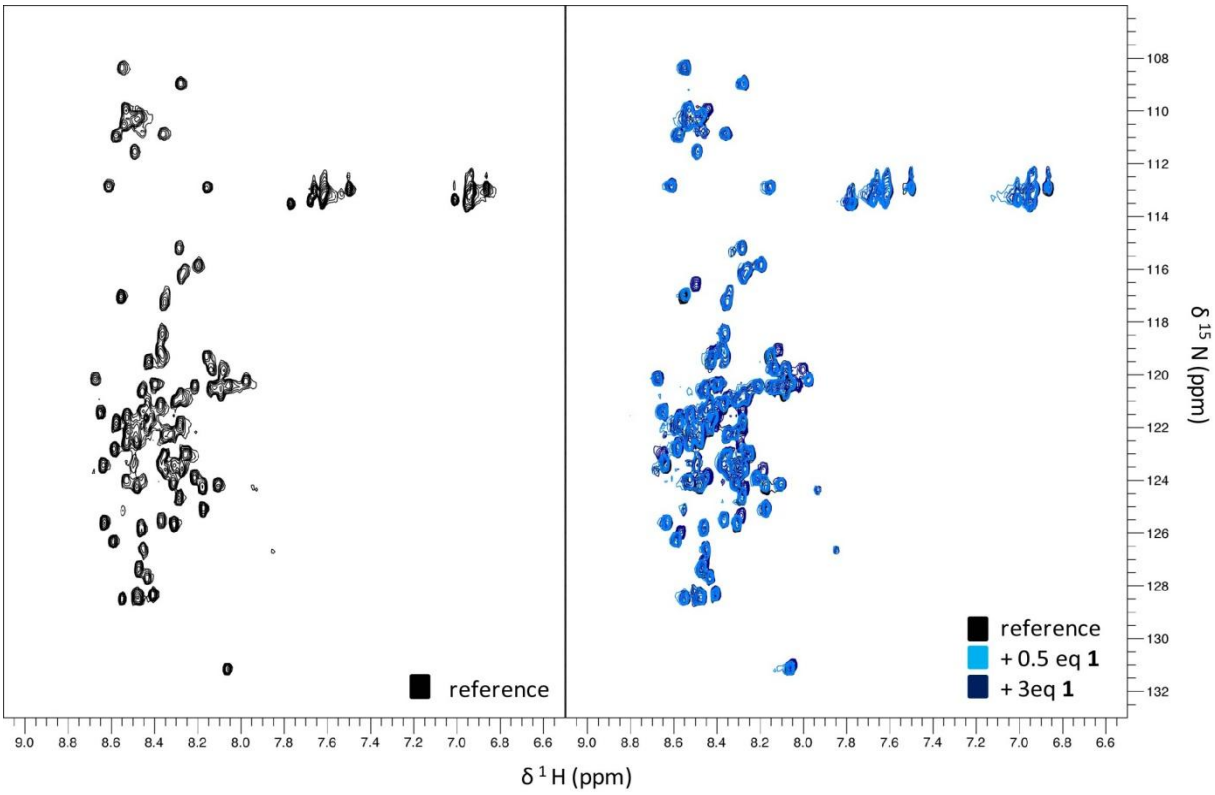


Fig. 17: Spectral region from the overlaid ^{1}H-^{15}N HSQC spectra of α-syn (0.1 mM) in the presence of increasing concentrations of **1**. The Fld reference (0.1 mM) is shown as black contours (left panel), while successive titration points are light blue and blue.

4.2.3 Crystallization trials

4.2.3.1 GB1

Crystallization attempts were done for all the complexes between GB1 and the five ligands. Here only the most important results, obtained with compound **1**, are reported.

As a starting point there were used the conditions found in literature for the crystallization of GB1 alone.[25] Those included the two different ways to afford crystals in diverse crystallization solutions, both consisting of 25 mM NaOAc, pH 4 and 4.5 respectively, and 2-methyl-2,4-pentanediol (MPD). The first one had 50% of MPD and NaCl as salt at 50 mM concentration, while the second one had higher % of MPD (70%) and 300 mM $MgSO_4$. The presence of salt cation can help crystallization forming salt bridges with the protein but for complexation studies is used in lower concentration, in order to facilitate the interaction between the macromolecule and a positively charged ligand and, then, to favour the crystallization of the complex rather than the free protein. Trials were done in hanging drops method by vapour diffusion technique, using 24-well plates and in sitting drops method

using 96-well plates. Crystal conditions were sought in the pH range 3.8-5.0 and MPD in the range of 30-80% was also used. The protein concentration was maintained equal to that reported in literature (1mM). Proper concentrations of the ligand were found to be lower than 5 mM since fast precipitation did not occur. No attractive precipitates were afforded with NaCl so only magnesium sulphate was further explored as salt. Particularly 2 and 3 mM as ligand concentration for **1** were found to give rise to crystalline materials through manual method. The most interesting results were obtained, with crystal growth of 1-2 weeks, by the following conditions, reported for each crystal (images obtained by optical microscopy):

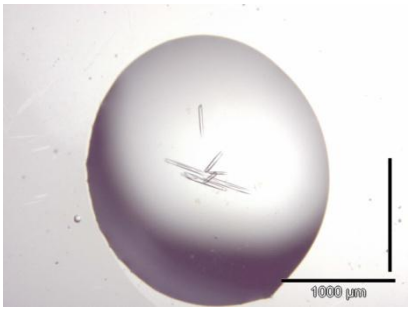


Fig. 18a: 30% MPD, 25 mM Na acetate pH 4.5, 100 mM $MgSO_4$

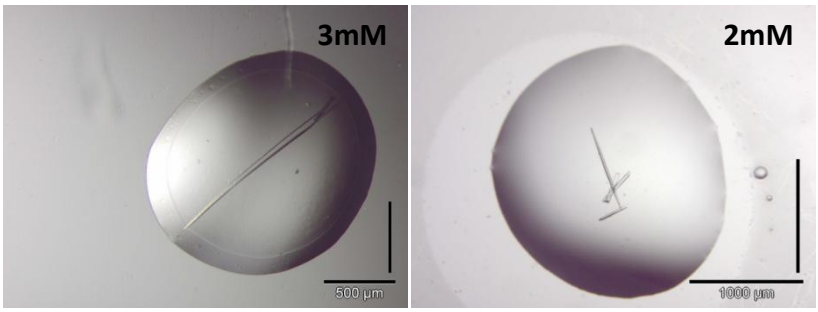


Fig. 18b: Images obtained by optical microscopy of crystals using 40% MPD, 25 mM Na acetate pH 4.5, 100 mM $MgSO_4$

Fig. 18c: Images obtained by optical microscopy of crystals using a) 30% MPD, 100 mM Na acetate pH 4.5, 100 mM $MgSO_4$ b) 30% MPD, 50 mM Na acetate pH 4.5, 80 mM $MgSO_4$ c) 30% MPD, 80 mM Na acetate pH 4.5, 80 mM $MgSO_4$ (all 2 mM)

All the crystals were sent to be analyzed by synchrotron in Paris, but the structure obtained revealed in some cases the presence of only protein and in some cases only salts. Hence the structure of the complex between ligand **1** and GB1 has not be resolved so far and this prevented for the moment a comparison and a possible confirmation of the collected NMR data.

4.2.3.2 Fld

For this second protein the most important results were obtained with compounds **1** and **4**. As starting point for the crystallization of Fld-ligand complexes the conditions found in literature were used.[41] The manually set-up screenings were firstly prepared and since Bis-Tris indicated in the literature Fld crystallization[41] was not available in laboratory, TRIS-HCl was used, changing concentration from 20 to 60 mM. $CaCl_2$ was explored in the range 10-100 mM, while PEG 400 in percentages of 10-50%. Since larger crystals of Fld were obtained using MPD as precipitant[41], this reagent was used to explore more reservoir solutions. The protein concentration was kept 1 mM (as reported in literature). For ligand **1** concentrations lower than 3.5 mM were used to avoid fast precipitation, whereas for calixarene **4** they were lower than 2.5 mM.

However all the trials did not give rise to crystalline materials and so the sitting-drop method using Douglas instrument Robot and screening solutions from JenaBioscience were tested. For ligand **1** two reservoir solutions, employing the sitting-drop method gave rise to the following interesting results reported in figure **19** (images again obtained by optical microscopy).

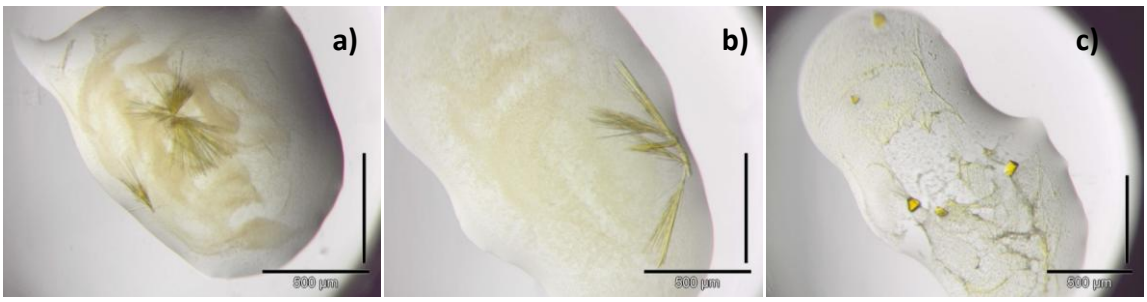


Fig. 19: Images obtained by optical microscopy of crystalline materials using a) 20% PEG 3350, 200 mM ammonium formate pH 6.6, [**1**] = 3 mM and b) with [**4**] = 2 mM; c) 20% PEG 3350, 200 mM magnesium formate pH 5.9, [**4**] = 2 mM.

As it is possible to observe, both reservoir solutions contain formate salt, indicating a role of this anion in the formation of needles/crystals. Probably this is able to form ion bridge with guanidinium moieties on the calixarene scaffold and positively charged residues on Fld. Unfortunately the dimensions of these needles and crystals were too small to be studied via X-ray diffraction, so no proof of this hypothesis could be done.

On the other hand for ligand **4** after some manual plates using different modification of literature conditions, a robot screening was carried out, exploiting 2.5 and 3 mM concentrations for the calixarene. An amorphous crystalline material was obtained with one reservoir solution, different from the literature, as reported in the figure below.

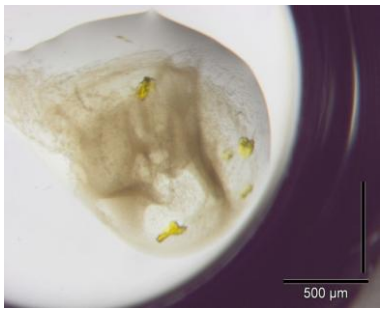


Fig. 20: Image of the crystalline materials obtained using 20% PEG 8000, 100 mM phosphate/citrate pH 4.2, 200 mM NaCl

Trying to improve the quality of the crystal and and to obtain some of them suitable for X-ray studies, other manual methods were prepared using both the previous crystallization solution and some modifications of that. Finally some small yellow crystals were achieved, with both manually-prepared and Jena reservoir solutions (**figure 21**).

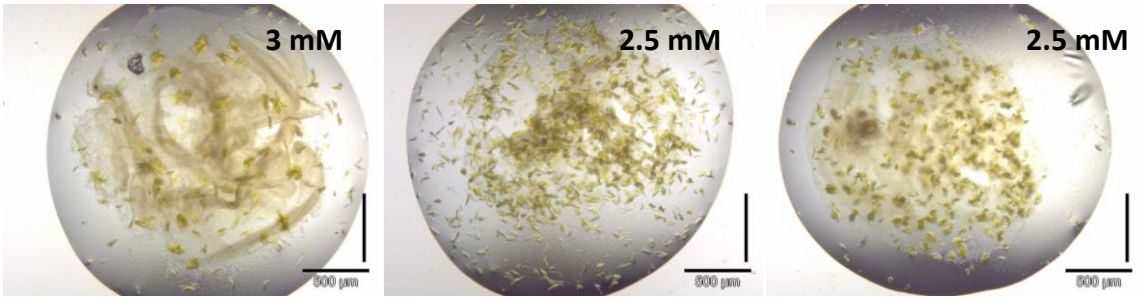


Fig. 21: Crystalline materials obtained with 20% PEG 8000, 100 mM phosphate/citrate pH 4.2, 200 mM NaCl (from JENA solution) with addition of paraffin oil trying to reduce the nucleation rate.

These were very reproducible but even after the addition of paraffin oil, in order to reduce the nucleation rate, the crystals were too small for structural studies with X-ray diffraction. Therefore no suitable crystals of Fld/**1** and Fld/**4** were obtained and so no additional information about these protein-ligand complexes, especially for calixarenes not possible to study by NMR, were achieved.

4.2.4 Synthesis of glycocalix[6]arenes

Two glycocalixarenes were synthesized, with the same valency and different sugars, for studies on their interaction properties towards the L-fucose-binding lectin of the bacterium *Ralstonia solanacearum* (RSL), having six binding site for these carbohydrate (as shown in figure **22**). Particularly the fucosyl calixarene **6** was designed to selectively recognize this lectin, bearing the right sugar, perceived by RSL. The galactosylcalix[6]arene **7** was actually synthesized as a control system, since this monosaccharide has a very low affinity for this lectin and the idea was to exploit it to verify the involvement of the fucose units in the observed complexation between **6** and RSL.

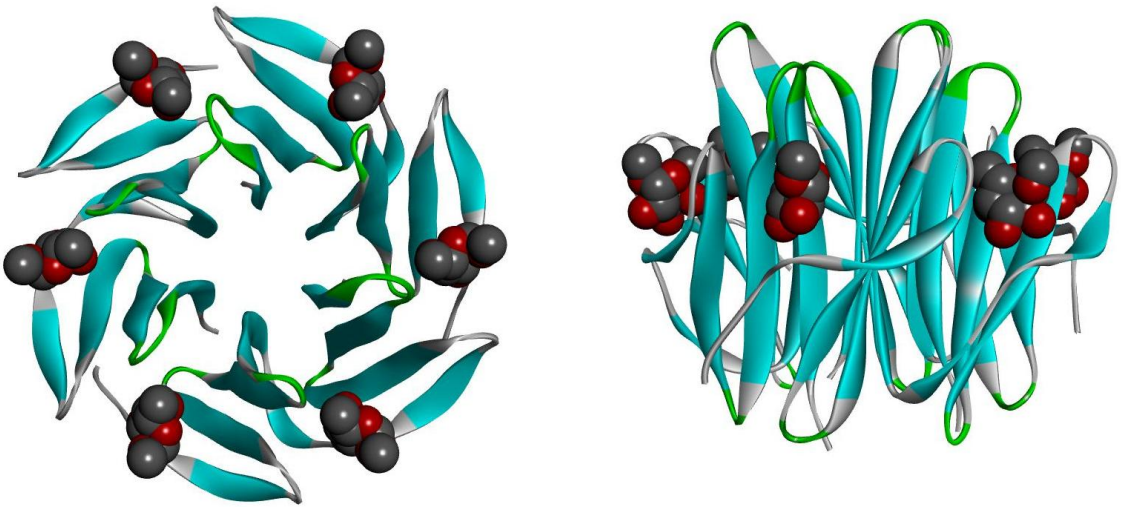
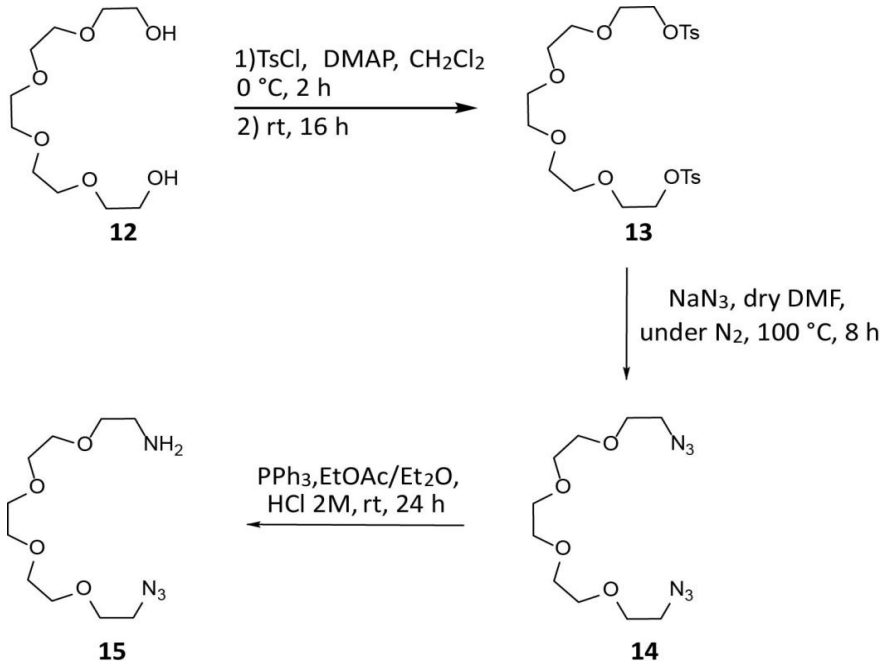


Fig. 22: Frontal and lateral view of RSL with six bound L-fucose units, from Protein Data Bank

4.2.4.1 Synthesis of the linker

The idea of the long PEG containing linker was followed since the aim was to have a ligand potentially able to establish simultaneously a contact with all the six recognition sites of the

lectin with its six fucose units giving rise to a 1:1 complex. Therefore it was necessary at least a linker containing 15 atoms to connect the aromatic core to the planned alkynyl glycosides. To achieve this feature the penta-ethylene glycol derivative **15** was selected, as main component of the spacer, functionalized with an amino group at one end and an azido function at the other one (**Scheme 2**). In this way the amino group could react easily with the carboxylic acid on calixarene **17** creating an amide bond. On the other hand the azido functionality was necessary for the "click" CuAAC with the alkyne on the sugar moieties. Furthermore the presence of polyethylene chains could enhance the solubility of the final compounds in water solution. This feature too was very important since the ligands must be soluble in water, because all the NMR experiments should be performed in an aqueous buffer solution and only a 10% of DMSO is permitted to avoid protein denaturation.



Scheme 2: synthetic strategy to obtain final linker **15**

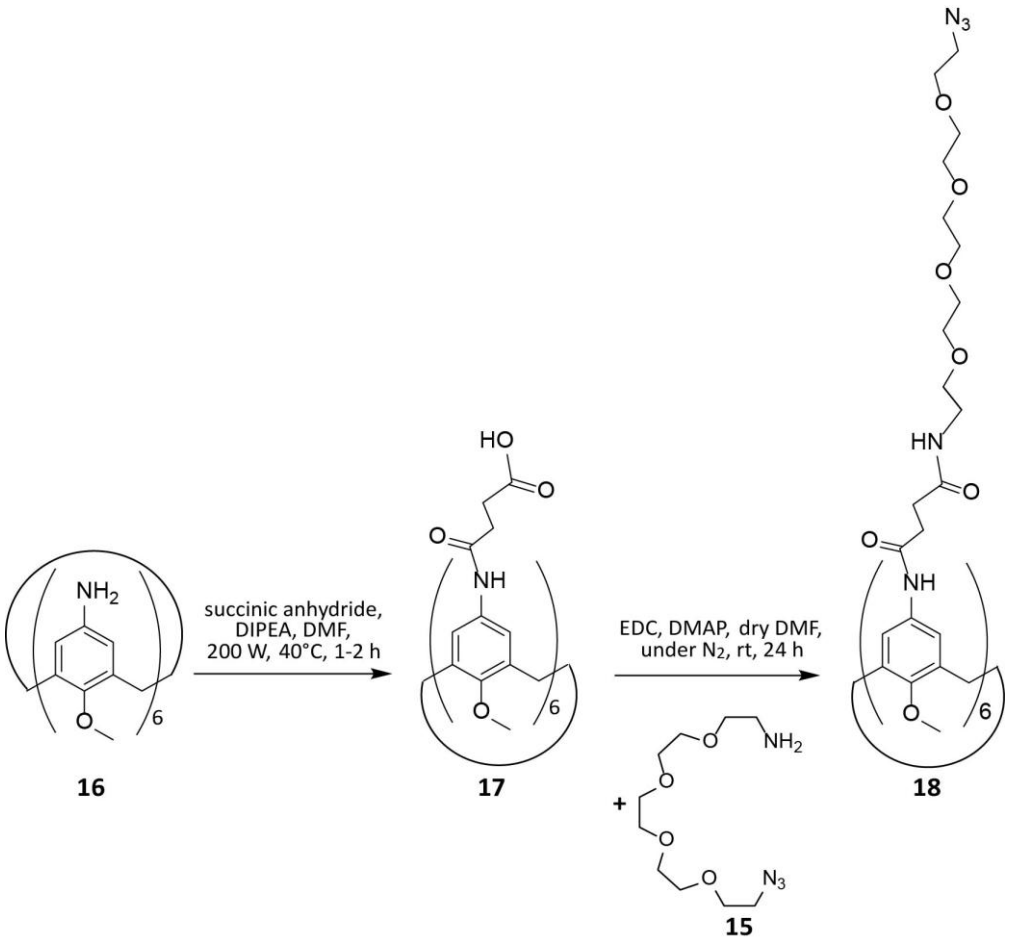
The asymmetric linker **15** was afforded following the strategy reported in **Scheme 2**. The first step was the protection of OH groups in **12** using tosyl chloride, to give rise to ditosylated product **13**, that was in turn reacted with NaN_3 to obtain the diazido derivative **14**. Since a selective reduction of only one azido group to amine was required, a desymmetrization reaction was employed where triphenylphosphine was used as reducing agent in a biphasic

system (EtOAc-Et_2O/aqueous HCl 0.5 M).[69,70] The selective reduction is possible since as soon as the mono amino compound **15** is formed, it becomes protonated and diffuses to the acidic aqueous phase. The high polarity of this derivative does not permit the transfer back to the organic solution. After work up (only extractions) no presence of the diamino compound was found demonstrating how highly selective is the reaction. So no purification steps were required, since all the remaining reagent was in the organic phase.

4.2.4.2 Synthesis of calixarene scaffold and coupling with the linker

The calix[6]arene scaffold was chosen based on the structural characteristics of RSL, since there are two carbohydrate binding sites per monomer, for a total number of six for the three monomers. The exposition of six sugar units at proper distance could maximize the interaction with the protein.

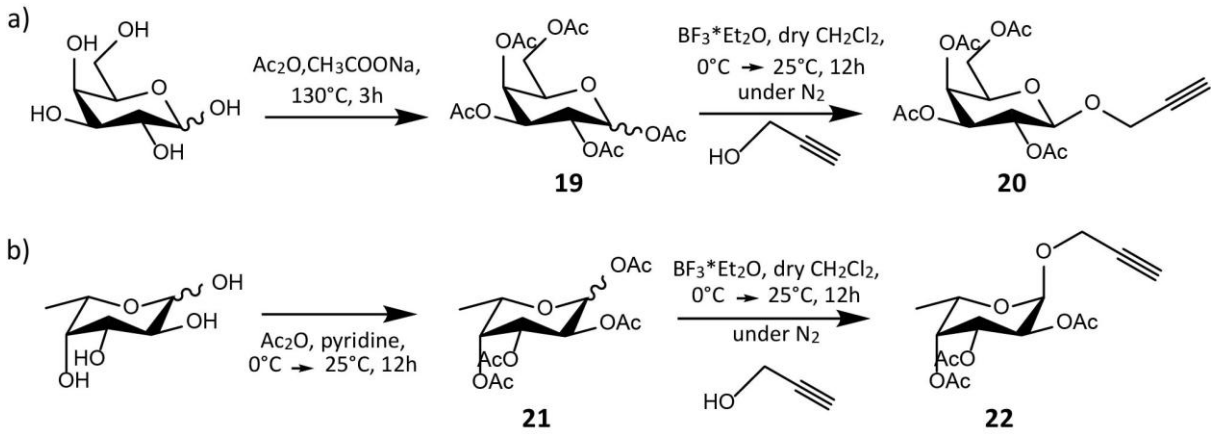
Starting from the amino calix[6]arene **16** (the procedure to obtain this derivative was reported in chapter 3), the first reaction was the formation of the amide bond by reaction between the amino groups on the calixarene and succinic anhydride to form derivative **17**. This compound was fully characterized by NMR and ESI-MS analysis since not already present in literature. Subsequently the PEGylated linker **15**, previously prepared, was introduced to create the longer spacer of the final derivative through the formation of a second amide bond between the amino group on the spacer and the carboxylic acids on the macrocycle. The complete functionalization was achieved as confirmed in particular by ESI-MS analysis. Also **18** was fully characterized. The azide functionality was then used to connect the macrocyclic scaffold to the sugar units, bearing an alkyne on the anomeric carbon (see next paragraph).



Scheme 3: synthetic strategy for the preparation of properly linker-functionalized calix[6]arene **18**

4.2.4.3 Synthesis of the saccharidic unit

The starting points for the formation of the two final sugars **20-22** are the fully unprotected D-galactose and L-fucose respectively. The peracetylation of hydroxyl groups was carried out with acetic anhydride for both and in the presence of sodium acetate for the first one and pyridine for the second one, attaining compounds **19** and **21**. The further step was the functionalization of the anomeric carbon with propargyl alcohol for the insertion of the terminal alkyne, able to react with the azide on the calix[6]arene **18**. The reaction allowed the formation as main product of β-anomer **20** for galactose and α-anomer **22** for fucose.



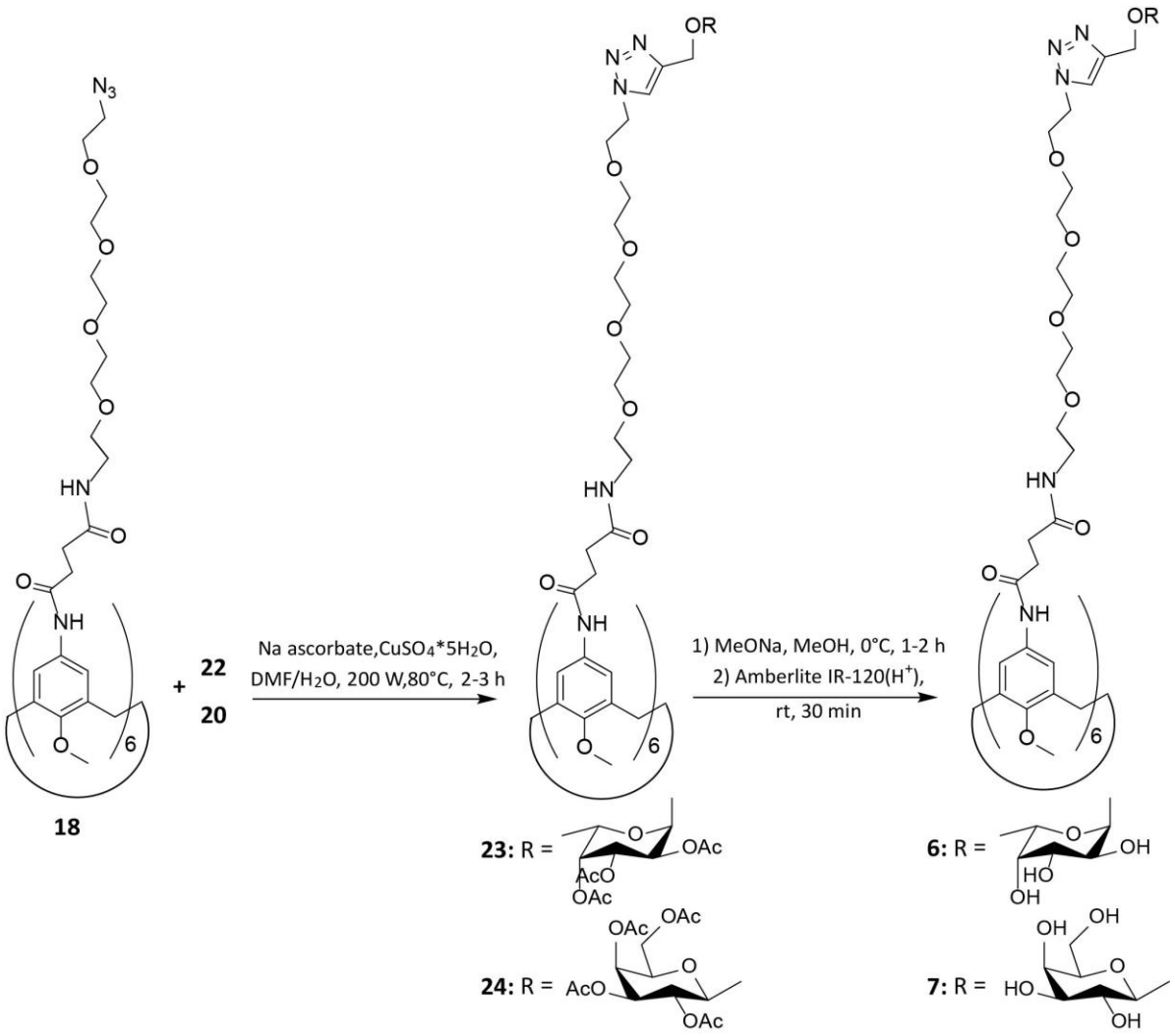
Scheme 4: Synthesis of a) propargyl galactose **20** and b) propargyl fucose analogue **22**

To confirm the success of the reactions ^{1}H-NMR spectroscopy was used. Particularly the signals of H_1 confirmed the formation of β anomer for galactose (J = 8.0 Hz) typical of an axial-axial coupling, with a upfield shift respect to the other anomer. Indeed for fucose the coupling constant was smaller, J = 2.9 Hz, due to an axial-equatorial correlation between H_1 and H_2, with a downfield shift respect to H_1 in a β-linkage. The broad singlet around 2.5 ppm, integrating for 1 proton, confirmed for each derivatives the presence of the alkyne terminal CH.

4.2.4.4 "Click" reaction to form new glycocalix[6]arenes and deprotection

The reaction between the alkyne-terminating glycosides **20** and **22** and the hexa-azide calix[6]arene **18** was carried out using Copper-catalyzed-alkyne cycloaddition (CuAAC),[71] to lead the formation of 1,4 disubstituted triazole ring.

The conditions used were the same for both saccharidic derivatives **20** and **22**, employing DMF and water as solvents, under MW irradiation. The formation of the desired glycocalix[6]arenes **23-24** was reached in less than 2 h and a purification step was necessary only for the removal of the carbohydrate excess, since no evidence of partially functionalized compounds was found by ESI-MS analysis.



Scheme 5: Synthesis of final glycocalix[6]arene **6** and **7** from **18**

Products **23-24** were fully characterized. The presence of the triazole unit was confirmed by the singlet at around 8 ppm for the hydrogen of this ring, corresponding to 6 protons, indication of the complete functionalization of azide groups with glycosides for both intermediates. Particularly the macrocycle **23** showed to be conformationally mobile at room temperature, as the NMR signals of the methylene bridge (ArC*H*$_2$Ar) suggest. In fact in the ^{1}H-NMR spectrum a singlet at 3.9 ppm can be observed, and in the ^{13}C-NMR spectrum the corresponding peak is at 30.0 ppm, both typical values for calixarenes in a mobile conformation.[72]

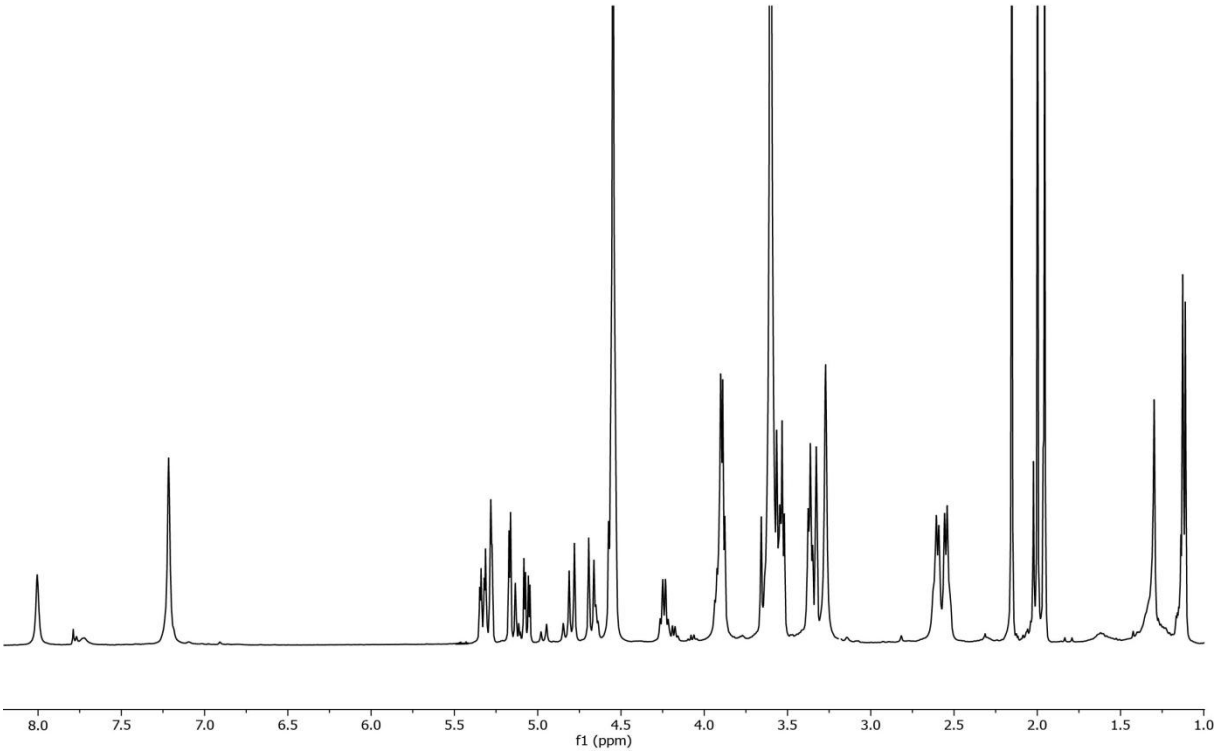


Fig. 23: ^{1}H NMR spectrum (400 MHz, MeOD/$CDCl_3$ 9:1) of fully acetylated fucose calixarene **23** at 50 °C

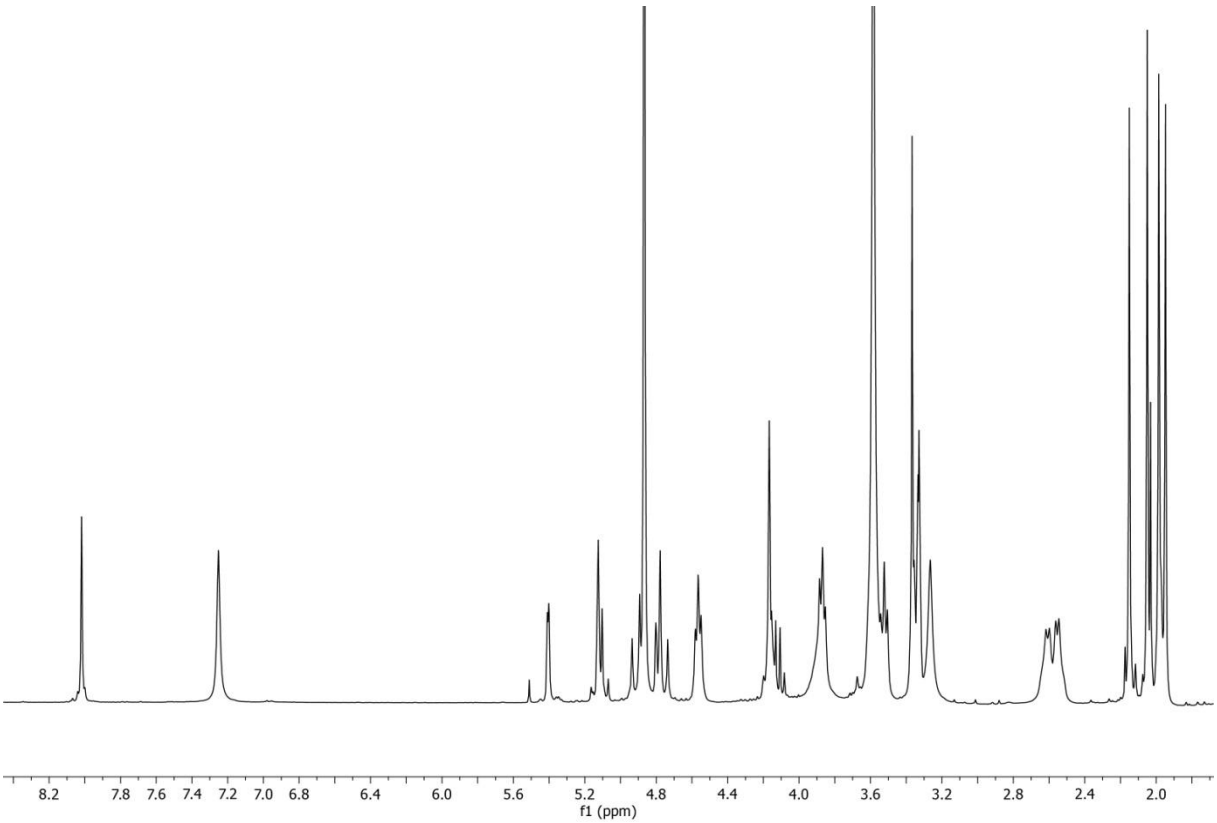


Fig. 24: ^{1}H NMR spectrum (300 MHz, MeOD, 25°C) of fully acetylated galactose calixarene **24**

The deprotection of fully protected fucose- and galactose- calixarenes **23-24** was performed using Zemplèn method, at 0 °C and the total removal of acetyl groups was confirmed by both ^{1}H-NMR and ESI-MS. Complete deacetylation was achieved in less than 3 h. Products **6** and **7** were completely characterized, being new compounds. The complete absence of the acetyl signals between 2.15 and 1.85 ppm in the ^{1}H NMR spectra and no peaks for 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 in ^{13}C NMR spectra) showed that the deacetylation was successful (figure **25**).

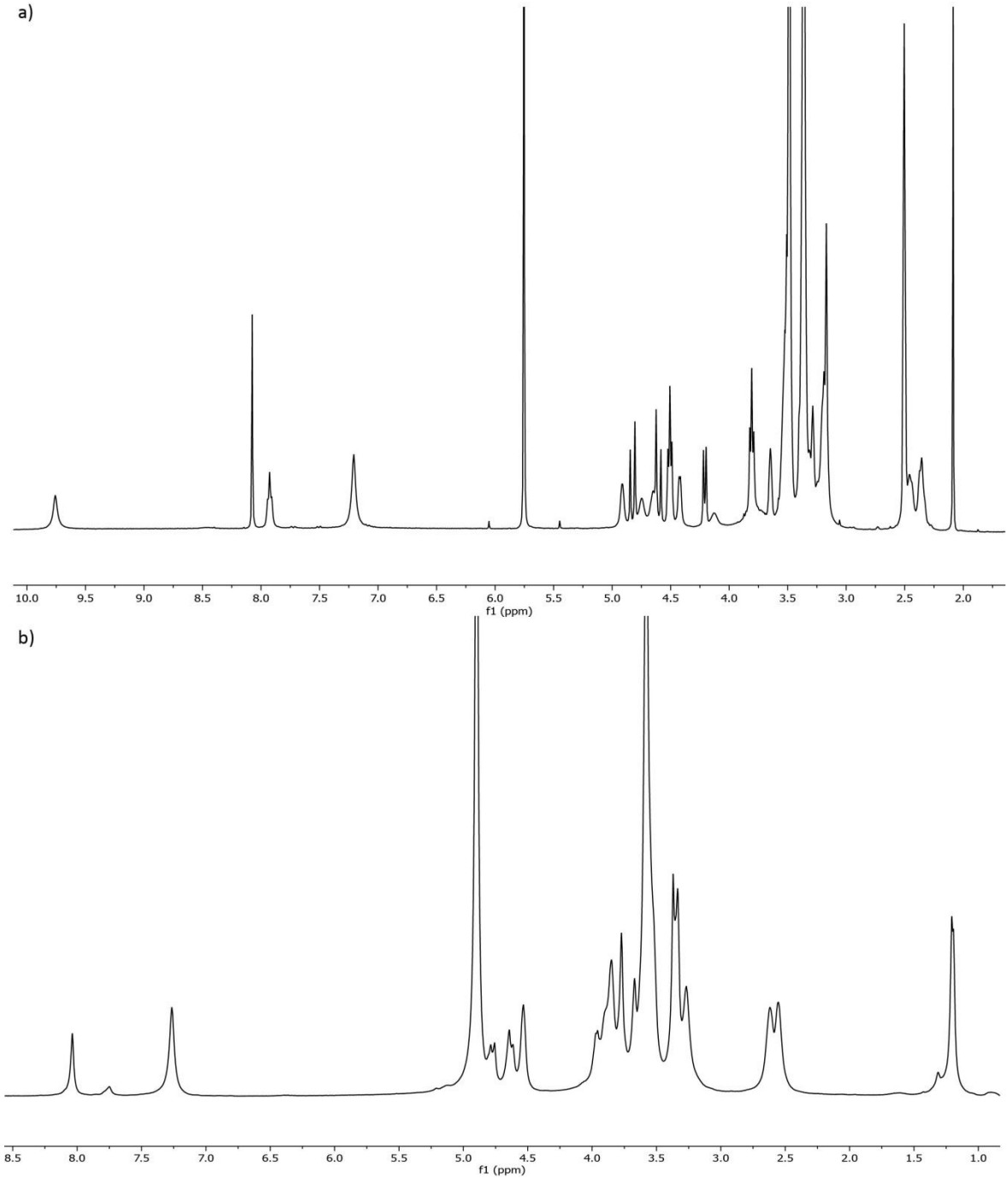


Fig. 25: ^{1}H NMR spectrum of deprotected a) galactose calixarene **6** (300 MHz, DMSO) and b) fucose calixarene **7** (400 MHz, MeOD)

These final derivatives were then studied with ^{15}N-labelled RSL using NMR spectroscopy, in order to disclose and identify the potential interactions that can occur.

4.2.5 RSL-calixarene binding: NMR studies

RSL was firstly over expressed using *E. coli,* and after harvested the cells, these were subjected to lysis and DNase addition. The lysate was purified by using FPLC (Fast Protein Liquid Chromatography) with an affinity column based on a mannose functionalized stationary phase. After that the lectin was exposed to dialysis as the last step of the purification protocol. The obtained pure ^{15}N-labelled protein was finally studied by NMR with compounds **6** and **7**, using the 2D ^{1}H-^{15}N HSQC experiment.

The fucosylated calix[6]arene **6** was used to titrate a sample of 0.3 mM ^{15}N-RSL. Even at the addition of the first aliquot of ligand, we observed significant precipitation in the NMR tube. However the spectrum was acquired. Comparing this spectrum with that of the protein alone broad signals were observed, characterized by the same frequency resonances of the free macromolecule without any chemical shift perturbation. This suggested that the revealed peaks were just leftover of unbound RSL while the precipitate consisted of a mixture of lectin and calixarene interacting one to the other. Another addition of the ligand (100 μM the final concentration) caused the complete disappearance of the signals because of the precipitation of the whole amount of the protein. This indicated the ability of this ligand to bind to the lectin forming either a complex less polar than the free protein or, more probably, a large aggregate due to inter-crosslinking and subsequent agglutination. In this second hypothesis, the 1:1 complex had a very small chance to be present. Interestingly, with the addition of α-methyl fucoside to this sample, the ^{1}H NMR signals were restored strongly indicating that the interaction between **6** and RSL was taking place thanks to the sugar units of the macrocycle. Moreover, only a large amount (6 mM) of the monosaccharide was necessary to reach this result and suggesting that the complex was rather stable and the strength of the lectin-calixarene interactions remarkable. Unfortunately, the precipitation made impossible a quantitative evaluation of the Kd of the complex.

To further confirm the involvement of the sugar units of **6** in the interaction, we proceeded in titration of lectin with the galactosylated derivative **7**.

On the other hand for compound **7** with the galactose units no precipitation occurred and the whole titration was carried out, till 5 eq. of the ligand, without observing any change in

the chemical shifts. This was a clear indication that this glycocalixarene did not interact with the lectin and then a strong proof that, on the contrary, the fucosylated ligand **6** binds to the lectin using its sugar units and does not establish aspecific interactions with other parts of its structure like the aromatic units or the long spacers.

On these basis, derivative **6** could be in the future studied as ligand of this and other lectins selective for α-fucoside in biological assays and used for crystallization trials that in this work we had not enough time to perform.

4.3 Conclusions

NMR spectroscopy, and particular ^{1}H-^{15}N HSQC experiments, was used to study protein-ligand interactions. 5 cationic ligands **1**-**5**, all based on a calix[4[arene scaffold, were employed with three different acidic proteins: GB1, Fld and alpha synuclein.

GB1 and Fld were observed to interact with all five cationic calix[4]arenes but showing different behaviours and affinities. For both proteins the interaction with compounds **2** and **5** resulted in complete precipitation upon addition of the ligand. Among the other three derivatives, **1**, **3** and **4**, for which titration and determination of a Kd value were possible, for GB1 the lower rim guanidinium calixarene **4**, blocked in cone geometry resulted as the most efficient with Kd = ~80 μM while **1**, bearing the same positively charged groups, but at the upper rim and being the macrocycle conformationally mobile was the weakest. The lower efficiency of **3** respect to **4** suggests that guanidinium is more efficient than ammonium for the binding to this protein. From the observed chemical shift perturbations we could conclude that all three ligands have preference for the same regions of the protein surface. On the other hands, the increased affinity of ligand **4** is accompanied by a loss of selectivity since many areas of the proteins result affected by the presence of the macrocycle. From the analysis of the most affected amino acids in the protein sequence, charge-charge, CH-π and cation-π interactions seem to take place.

For Fld the only ligand completely studied by NMR was **3**, because **1** caused precipitation and for **4** the data were not reproducible. The recognition process between the guanidinium derivative **1** and protein appeared more selective than those observed for GB1, with only few amino acid peaks shifted throughout the titration experiments

In the ^{1}H-^{15}N HSQC experiment the ligand-induced changes in f1 and f2 dimensions are

poorly correlated respect to the type of binding that can occur. In fact ^{1}H shifts are influenced both by hydrogen bonding and ring current, while ^{15}N shifts are due to hydrogen bonding and conformational changes. Particularly ^{15}N shifts are strongly correlated to hydrogen bonding with the adjacent carbonyl which has a very low influence on the proton shift. Secondly, electric field effects, as hydrogen bonding, have a proportional effect to the cosine of the angle to the relevant peptide bond, but these angles are different for the two dimensions. For the first one this implies that direct interaction with the ligand can be observed also in the shift in the adjacent amide nitrogen.[6]

For all the ligands with GB1 and Fld the shifts are mainly in the H and so this is a confirmation that the cavity has an important role in the recognition process, involving different type of binding. As a confirmation of this behaviour in **1,** where the cavity is not so available, the shifts were in GB1 very small. However the shifts of negatively charged residues proved the involvement of guanidinium moieties, thanks to electrostatic interactions.

For the cases where complete titration was possible with a consequent treatment of the experimental points the 1:1 complex resulted to satisfy the collected data. There is also for all the proteins only a single binding mode, since the titrations gave rise to straight lines in the spectrum for all the ligands, while multiple binding will in general produce curved plots, since the primary interactions will almost always have different effects on the chemical shifts respect to the secondary binding event.[6] Furthermore it is possible to say that the observed large shift changes are local and therefore due to the calixarene binding, since CSPs from conformational changes on binding are often very small. Hence it is not the case of the binding events here studied between our cationic calixarenes.

Unfortunately no crystals of complexes were obtained, although a rather large screening testing different conditions was tried to obtain single crystal suitable for X-ray analysis useful to clarify the structure of the complexes.

For *Ralstonia solanacearum* lectin L-fucose glycocalix[6]arene **6** was synthesized as potential ligand. This derivative was able to interact with RSL but was not possible to study this binding by NMR spectroscopy, because of precipitation of the complex. Absence of interaction with the galactosylated calixarene **7**, prepared as negative control, and

competition in the binding by α-methyl fucoside both indicated that the complex form thanks to the presence of the fucose units on multivalent ligand **6**and their contacts with the lectin carbohydrate recognition sites. The observed precipitation phenomenon is reasonably due to the formation of a large network because of inter-crosslinking that determines agglutination

4.3.1 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) and then, according to the functional groups present on the molecules, revealed with UV light or using staining reagents: ninhydrin (5% in EtOH), basic solution of $KMnO_4$ (0.75% in H_2O), Pancaldi solution (molybdatophosphorus acid and Ce(IV)sulphate in 4% sulphuric acid). Merck silica gel 60 (70-230 mesh) was used for flash chromatography and for preparative TLC plates. ^{1}H NMR and ^{13}C NMR spectra were recorded on Bruker AV300 and Bruker AV400 spectrometers (observation of ^{1}H 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 ^{1}H NMR spectra recorded in D_2O at values higher that 25°C the correction of chemical shifts was performed using the expression $\delta = 5.060 - 0.0122\times T(°C) + (2.11\times10^{-5})\times T^2(°C)$ to determine the resonance frequency of water protons.[59] Electrospray ionization (ESI) mass analyses were performed with a Waters spectrometer in both positive and negative mode with MeOH/CH_3CN as solvents. Melting points were determined on an Electrothermal apparatus in closed capillaries. Microwave reactions were performed using CEM Discovery System reactor.

Protein production and purification

All the reagents were from Sigma. All the proteins (unlabelled and ^{15}N-labelled) were over-expressed in Escherichia coli BL21 (DE3) and purified according to literature methods. The purity and concentration of the protein was estimated by using 15% SDS polyacrylamide gel electrophoresis (SDS-PAGE) and UV-vis spectroscopy (Perkin Elmer Lambda 35).

Luria-Bertani (LB) medium was prepared with 10 g N-Z amine, 10 g Yeast extract, 5 g NaCl. Add 1 L H_2O and adjust pH to 7.0. LB Agar plates were prepared using 10 g N-Z amine, 10 g Yeast extract, 5 g NaCl, 15 g Agar,2 mM $MgSO_4$. Add 1 L H_2O and adjust pH to 7.0.

Minimal medium (MM) ^{15}N labelled minimal medium was produced with 50 mM Na_2HPO_4, 50 mM KH_2PO_4, 2g/L D-glucose, 1 g/L NH_4SO_4 ^{15}N-labelled, 20 mM citrate/20 mM succinate pH 7.0, 20 ml/L 50x5052, 30 mg/L Thiamine), 1 µL/mL Carbenicillin (75 mg/mL). Note: 50x5052 contains: 25 g glycerol, 7 ml water, 2.5 g glucose, 10 g lactose monohydrate.

Protein over-expression

E. coli BL21, a strain which is optimized for IPTG-inducible recombinant protein over-expression, was transformed with the plasmid encoding for the chosen protein (see paragraphs below). 1 µl DNA was added to 50 µl BL21 competent cells, incubated for 30 min on ice followed by heat shock for exactly 45 s at 42 °C. To allow the bacteria to recover, 250 µl LB was added and the sample was incubated for 45 min at 37 °C. 100 and 150 µl were then plated on a LB Carbenicillin plate, which was incubated overnight at 37 °C. The next day a pre-culture was set-up by picking a single colony and inoculating 5 ml LB medium containing 2 mM $MgSO_4$ and 1 mM Carbenicillin for selection. The pre-culture was incubated overnight at 30 °C. The next day LB medium (250 ml-1000 ml) containing again $MgSO_4$ and Carbenicillin was inoculated with the pre-culture. The cells were grown at 37 °C until an OD600 ~ 0.6-0.8 was reached in 1h 30 min-2h. Protein over-expression was induced by addition of IPTG(Isopropyl β-D-1-thiogalactopyranoside) to a final concentration of 1 mM. After 4 h of over-expression at 37 °C the cells were harvested by centrifugation for 20 min at 4,500 rpm. The pellet was then re-suspended in 5 mM of lysis buffer (10 mM TRIS-HCl, 1 mM EDTA, pH 7.5) and frozen at -20 °C, for at least one night to be ready for the further purification steps.

For preparation of ^{15}N-proteins when the OD at 600 nm ~ 0.6-0.8 the cells were harvested by centrifugation for 20 min at 4.500 rpm and the pellet was re-suspended in minimal medium.

After 30 min of shaking (300 rpm) at 37 °C the culture was split in aliquots of 250 ml volumes in 2 L Erlenemeyer flasks. Protein overexpression was then induced similarly by adding IPTG to a final concentration of 1 mM.

SDS-PAGE

A total of 20 µl cell/protein extract was mixed with 4x SDS buffer and boiled at 95°C for 5min. The purity of denatured, reduced samples were tested on 15% SDS polyacrylamide gel electrophoresis (80 minutes, 140 V). As a protein size marker 5 µl of EZ-Run™ Pre-Stained Rec Protein Ladder (Fisher BioReagents) were loaded. Coomassie Blue staining was performed as described elsewhere.[73]

GB1 purification

Protein over-expression was carried out according to the protocol by F.W Studier.[73]

The GB1K10H plasmid was transformed onto *E. coli* BL21 (DE3) and the protein purification followed the method explained above. After that the cells were harvested by centrifugation (Hettich 4723 Rotor) for 45 minutes, 10°C, 4500 rpm. After removal of the medium, the cells were left in lysis buffer, freezing the sample overnight.

Protein isolation and purification was performed according to the literature.[74] The re-suspended cells from the culture were heated to 80°C and then directly cooled on ice. They were centrifuge at 10°C, 11000 rpm, 10 min and after poured into 15 ml of buffer A (20 mM TRIS/HCl pH 7.5). The protein was purified using a procedure similar to Fld. The column was equilibrated in buffer A (20 mM TRIS HCl pH 7.5). When all the positively charged proteins came off, a gradient from 0 to 50% of buffer B (20 mM TRIS HCl, 1 M NaCl pH 7.5). Fractions containing GB1-QDD were pooled,concentrated and loaded into a second DEAE column, performed with a manual gradient of buffer B. After pooled together and concentrated all the fractions containing GB1, the sample was loaded onto an SEC column (20 mM KP_i, 50 mM NaCl pH 6.0). The presence of GB1 and its purity were checked with SDS-PAGE and NMR (for ^{15}N-labelled samples). The concentration of GB1 was determine via UV spectroscopy using an extinction coefficient ε_{280} = 99.7 $mM^{-1}cm^{-1}$.[24]

Fld purification

Protein over-expression was carried out according to the protocol present in literature.[34] The pDH01 plasmid was used and transformed onto *E. coli* BL21 (DE3). After over-expression the re-suspended cells from the daily culture were thawed and sonicated at 10 microns for 20 min (5 sec on and 2 sec stop), checking the absence of grey precipitate. The suspension was centrifuge at 10°C, 11000 rpm, 30 min. After the addition of DNase (1 mg/mL) and allowing to react, the liquid was suggested to $(NH_4)_2SO_4$ precipitation, adding to the solution 30% of the volume of a saturated solution, at 0°C. Then dialysis and a first DEAE (diethylaminoethyl) anion exchange column were required. Buffer A (100 mM NaCl, 20 mM Na_2HPO_4 pH 7.4) was used to introduce the sample into the column to elute positively charged proteins. Flavodoxin was eluted using a gradient from 0% to 100% buffer B (1M NaCl, 20 mM Na_2HPO_4 pH 7.4). The fractions containing Fld were pooled together.

A 2nd DEAE column was performed starting with 10 % buffer B and when all the impurities were come off a manual gradient (10% stepwise) was started. The fractions containing Fld were identified using SDS-PAGE and UV-visible spectrometer to determine the A ratio (A_{274} /A_{467})[40] and concentrated by ultracentrifugation before the purification with size exclusion chromatograph (SEC). This last step of purification was carried out using a G57 sepharose XK 16/75 column (1.6 cm diameter 65 cm bed height) with buffer 20 mM KP_i, 200 mM NaCl pH 7.4. The final concentration was determined by UV spectroscopy using an extinction coefficient ε_{466} = 8256 $M^{-1}cm^{-1}$.[34]

α-synuclein purification

Protein over-expression was carried out according to the protocol present in literature.[75] The pT7-7 plasmid was used and transformed onto *E. coli* BL21 (DE3). This was a kind gift from Gary Pielak (University of North Carolina, Chapel Hill, NC). After expressing α-synuclein for 4 hours the daily culture was harvested by centrifugation at 4500 rpm for 30 min at 4°C. The pellet was re-suspended in lysis buffer (~ 8 ml) before freezing. The sample was later thawed and boiled for 20 minutes with regular agitation before cooling on ice for 5 min. A few crystals of DNase and RNase were added to the slurry along with 1X protease inhibitor cocktail. The mixture was incubated on ice for a further 5 minutes. The entire slurry was divided into 15 x 500 ul aliquots and centrifuged at 14000 rpm for 10 min at rt. The

supernatant was collected, mixed and diluted to 10 ml before dividing into 10 X 1 mL aliquots. The aliquots were frozen. Directly before use a single aliquot was thawed and centrifuged to remove debris. The supernatant was diluted to 5 mL before loading onto a DEAE column equilibrated in 100 mM NaCl, 10 mM Na_2HPO_4 pH 7. After 20 fractions, a 100 % buffer B (1 M NaCl, 10 mM Na_2HPO_4 pH 7) gradient was introduced over 70 min. The fraction corresponding to α-syn were concentrated and loaded onto an SEC column equilibrated in 100 mM NaCl, 10 mM Na_2HPO_4 pH 7. The purity of final sample was determined by NMR spectroscopy and SDS-PAGE.

RSL purification

RSL samples were produced in *E. Coli* BL21 transformed with the plasmid pET25rsl.[50] After over-expression, RSL was purified by affinity chromatography (D-mannose-agarose resin, SIGMA) on an ÄKTA FPLC. During an initial purification the entire cell extract was loaded onto the column, however due to the limited binding capacity only a fraction of the RSL could be purified. Therefore, it was performed a re-purification applying multiple sample injections to maximize the protein yield. Firstly buffer A (100 mM NaCl, 200 mM TRIS-HCl pH 7.4) was used to eliminate other impurities. Then buffer B (100 mM NaCl, 200 mM TRIS-HCl, 0.1 mM D-mannose pH 7.4) was introduced to elute RSL and all fractions were then collected, concentrated by ultrafiltration. Extensive dialysis was required as last purification step to remove D-mannose.[50] The protein concentration was determined by UV spectroscopy using an extinction coefficient ε_{280} = 44.6 $mM^{-1}cm^{-1}$ and the purity was checked by SDS-PAGE and NMR spectroscopy.

Ligand stock solutions

All the stock solutions of the ligands were prepared in H_2O, with a concentration of 10 mM (except to **3** that had concentration of 80 mM) and pH adjusted to 7.0. Even fucose calixarene was prepared with a different concentration of 5 mM.

NMR spectroscopy

Samples were prepared using 0.05 mM GB1/RSL/α-syn or 0.1 mM Fld in 20 mM KP_i, 50 mM NaCl, pH 6.0, 10 % D_2O with a final volume of 550 µL. For α-syn the buffer was 10 mM Na_2HPO_4 , 100 mM NaCl, pH 7.4.[68] The sample pH was verified to be constant before and

after data collection (maximum variation ± 0.05 pH units). The ligand stock solutions were prepared in H_2O and pH adjusted to 7.0. ^{1}H-^{15}N HSQC spectra were acquired at 30° C on a 600 MHz Varian NMR spectrometer equipped with a HCN cold probe, with spectral widths of 16 ppm (^{1}H) and 40 ppm (^{15}N). Titrations were performed by the addition of µL volumes of the ligands. Spectra were acquired with 8 (GB1/RSL/α-syn) or 16 (Fld) scans and 64 increments. Data processing was in Biopack (with linear prediction in the ^{15}N dimension) and NMRpipe.[76] Titrations were repeated three times to ensure reproducibility (except to compounds **2**, **5** and all the glycocalix[6]arenes). Spectral changes due to ligand binding were assessed in CCPN[77].

Binding Isotherms

Binding isotherms were obtained by plotting the magnitude of the chemical shift perturbations (Δδ) as a function of the ligand concentration. The data were fit (non-linear least squares) to a one-site binding model, with Δδ and [ligand] as the dependent and independent variables, respectively, and the maximum chemical-shift change ($\Delta\delta_{max}$) and the dissociation constant (K_d) as the fit parameters. A global data analysis was performed in which the curves were fit simultaneously to a single K_d value, while $\Delta\delta_{max}$ was varied for each resonance.

Crystallization

The hanging- and sitting drop vapour diffusion method were used for crystallization at 20°C. Crystals were grown starting from similar conditions to those reported for GB1 and Fld. The drops were prepared by combining 1 µL of protein, calixarene ligands and reservoir solution. JBScreen JCSG solutions were used for the crystallization robot. Diffraction data for single crystals were collected at the synchrotone in Paris. Data processing and scaling were performed in CCP4-programs COMBAT or POINTLESS and SCALA.[78] The structures were solved by molecular replacement in PHASER.

Refinement and manual rebuilding were performed in REFMAC5 as implemented in CCP4[79] and COOT.[80] The presence of protein-ligand complexes was analysed in COOT.

Synthesis of the ligands

Compounds **1, 2** and **5** were already present in the laboratory of Crowley and synthesized previously in Parma as reported in literature.[57]

25,26,27,28-Tetrakis[3-aminopropoxy]calix[4]arene (9)

In a two neck round bottom flask, under nitrogen atmosphere, calixarene **8** (1.5 g, 1.28 mmol) was dissolved in EtOH abs (30 ml). Hydrazine hydrate (12.5 ml, 255.7 mmol) was added and the reaction was kept under reflux for 24 h. When it was finished (TLC eluent: hexane/EtOAc 1:1), the solvent was removed under reduced pressure and then the mixture was extracted with DCM (3X50 ml). The organic layers were then washed with brine and was concentrated under vacuum to afford compound **9** as a white solid (quantitative yield).

^{1}H-NMR ($CDCl_3$, 300 MHz): δ (ppm) 6.61-6.55 (m, 12 H, ArH meta and ArH para); 4.40 (d, 4H, *J* = 13.9 Hz, $ArCH_2Ar$ H ax.); 3.96 (t, 8H, *J* = 7.3 Hz, OCH_2CH_2); 3.17 (d, *J* = 13.9 Hz, 4H, $ArCH_2Ar$ H eq.); 2.09-2.00 (t, *J* = 6.6 Hz, 8H, OCH_2CH_2); 1.60 (br s, 8H, NH_2); 1.23 (quin, 8H, *J* = 7.3, *J* = 6.6 Hz, $OCH_2CH_2CH_2$).

The product shown the same chemical-physical data reported in literature.[81]

25,26,27,28-Tetrakis[3-(bis-Boc-guanidine)propoxy]calix[4]arene (10)

To a solution of calixarene **9** (0.3 g, 0.46 mmol) in dry DMF (15 ml), N,N'-di-Boc-thiourea (1.25 eq. per NH_2 group) and Et_3N (0.96 ml, 6.89 mmol) were added at rt. Then the mixture was cooled down to 0°C and $HgCl_2$ (2 eq for each NH_2 group) was added and the reaction was stirred at rt for 24 h. When it was finished (checked via TLC, hexane/EtOAc 9:1) the mixture was filtered off to eliminate HgS and the solvent was concentrated under reduced pressure. The crude was purified by flash chromatography column (eluent: hexane/EtOAc 9:1) and then with preparative TLC on silica gel (eluent: cyclohexane/EtOAc 8:2) to afford compound **10** as a white foam (yield = 30%). **^{1}H-NMR** (300 MHz, $CDCl_3$) δ (ppm): 11.51 (s, 4H, BocN*H*); 8.40 (t, 4H, J = 5.1 Hz, CH_2NH); 6.59 (br s, 12H, Ar*H*); 4.39 (d, 4H, J= 13.4 Hz, $ArCH_2Ar$ ax.); 3.98 (t, 8H, J = 7.0 Hz, OCH_2); 3.59 (m, 8H, CH_2CH_2NH); 3.19 (d, 4H, J= 13.4 Hz, $ArCH_2Ar$ eq.); 2.18 (quint, 8H, J = 7.1 Hz, OCH_2CH_2); 1.48 (s, 36H, t-Bu); 1.47 (s, 36H, t-Bu). The product shows the same physical-chemical data reported in literature.[56,82]

25,26,27,28-Tetrakis(3-guanidiniumpropoxy)calix[4]-arene, Tetrachloride (4)

In a round bottom flask compound **10** was dissolved in 1,4-dioxane (0.1 mmol/10 ml) and HCl 37% was added dropwise (10 equiv for each Boc group) The reaction mixture was stirred for 48 h, and the solvent was removed under reduced pressure to obtain the pure product as a white solid (quantitative yield). **^{1}H-NMR** (400 MHz, D_2O) δ (ppm): 6.67 (d, 8H, J = 8 Hz, Ar*H* meta); 6.54 (t, 4H, J = 8 Hz, Ar*H* ipso); 4.26 (d, 4H, J = 12 Hz, Ar*CH*$_2$Ar ax.); 3.92 (t, 8H, J = 8 Hz, OC*H*$_2$); 3.18 (m, 8H, CH_2CH_2NH); 2.08 (quint, 8H, J = 8 Hz, OCH_2CH_2). The product shows the same physical-chemical data reported in literature.[56,82]

25,26,27,28-Tetrakis[3-(trimethylammonium)-propoxy]calix[4]arene, Tetraiodide (11)

The tetraminocalix[4]arene **9** (54 mg, 0.07 mmol) was dissolved in MeOH (5 ml), then $KHCO_3$ (34 mg, 0.34 mmol) and MeI (0.06 g, 0.44 mmol) were added and the mixture was stirred at room temperature. The reaction was followed by ESI-MS and stopped after 7 days. A mixture MeOH/CH_2Cl_2 9/1 v/v (10 ml) was added and the insoluble inorganic salts were filtered off, removing the solvent under reduced pressure. The mixture was purified by trituration in Et_2O for 1h, attaining compound **11** as a white solid (quantitative yield).

^{1}H-NMR (400 MHz, D_2O) δ (ppm): 6.75 (d, 8H, J = 8 Hz, Ar*H* meta); 6.67 (t, 4H, J = 8 Hz, Ar*H* ipso); 4.28 (d, 4H, J = 12 Hz, Ar*CH*$_2$Ar ax.); 4.16 (t, 8H, J = 8 Hz, OC*H*$_2$); 3.38 (m, 12H, Ar*CH*$_2$Ar eq., CH_2CH_2NH); 3.07 (s, 36H, CH_3); 2.26 (br s, 8H, OCH_2CH_2). The product shows the same physical-chemical data reported in literature.[82]

25,26,27,28-Tetrakis[3-(trimethylammonium)-propoxy]calix[4]arene, Tetrachloride (3)

The iodide salt **11** (0.53 g, 0.4 mmol) was dissolved in H_2O (5 ml), and 5 ml of DOWEX-11 resin were added and the mixture stirred for 30 min. After removal of the resin, pure compound **3** was obtained by evaporation of the solvent under reduced pressure as a white solid in 98% yield. **^{1}H-NMR** (300 MHz, D_2O) δ (ppm): 6.57 (d, 8H, J = 7.4 Hz, Ar*H* meta); 6.42 (t, 4H, J = 7.2 Hz, Ar*H* ipso); 4.36 (d, 4H, J = 13.5 Hz, Ar*CH*$_2$Ar ax.); 4.14 (t, 8H, J = 8 Hz, OC*H*$_2$); 3.42 (m, 12H, Ar*CH*$_2$Ar eq., CH_2CH_2NH); 3.12 (s, 36H, CH_3); 2.30 (br s, 8H, OCH_2CH_2). The product shows the same physical-chemical data reported in literature.[82]

Pentaethyleneglycol-di-*p*-toluenesulfonate (13)

To a solution of *p*-toluenesulfonyl (5.76g, 30.21 mmol) and pentaethylene glycol (3g, 12.59 mmol) in CH_2Cl_2 (30 ml), cooled to 0°C with an ice-water bath, a solution of Et_3N (1.83 ml, 13.09 mmol) was added dropwise. A catalytic amount of DMAP was added and the solution was stirred at 0°C for 2 h and then a room temperature for other additional 16h. The reaction was checked via TLC (eluent: EtOAc) and the mixture was washed with HCl 0.5 M (2x50 ml) and brine (50 ml), then dried over Na_2SO_4 and filtered. The solvent was removed under reduced pressure to give a mixture that was purified by flash column chromatography (eluent: hexane/EtOAc 7:3 to 100% EtOAc) to afford compound **13** as a yellow oil in 86% yield.

^{1}H-NMR (400 MHz, $CDCl_3$): δ (ppm) 7.78 (d, 4H J = 7.6 Hz, Ar*H* ortho); 7.32 (d, 4H J = 7.6 Hz, Ar*H* meta); 4.14(t, 4H J = 4.8 Hz, 2xCH_2OTs); 3.67 (t, 4H J = 4.8 Hz, 2xCH_2CH_2OTs); 3.59-3.57 (m, 16H, OCH_2); 2.43 (s, 6H, ArCH_3). The product shows the same physical-chemical data reported in literature.

1,14-diazido,3,6,9,12-Tetraoxatetradecane (14)

To a solution of **13** (5.9g, 10.81 mmol) in dry DMF (40 ml), under nitrogen, NaN_3 (7.03g, 108.1 mmol) was added. The mixture was stirred at 100°C for 12h and checked via TLC (eluent: EtOAc). The reaction was quenched with cold water (50 ml) and the product extracted with Et_2O (5x40 ml). The combined organic phases were concentrated under vacuum and extracted again with brine (3x40ml) and evaporated. The product was obtained without purification as a yellow oil (yield = 88%).

^{1}H-NMR (300 MHz, $CDCl_3$) δ (ppm) 3.63-3.62 (m, 16H OCH_2); 3.37-3.35 (m, 4H, CH_2N_3). The spectroscopic data are in agreement with those found in literature.[83]

14-azido,3,6,9,12-Tetraoxatetradecane-1-amine (15)

To a solution of **14** in EtOAc/Et_2O 1:1 (20 ml totally) were added HCl 2M (ratio organic phase/HCl 5:1) and PPh_3. The reaction was stirred for 24 h at rt (TLC eluent: DCM/MeOH 9:1). The aqueous phase was washed with DCM (3X50 ml), basified with NaOH aqueous solution and extracted again with DCM (3X50 ml). The organic phase was dried over Na_2SO_4, filtered and then the solvent was removed under reduced pressure to afford compound **15** as a yellowish oil, in 56% yield.

^{1}H-NMR (300 MHz, $CDCl_3$) δ (ppm) 3.68 (m,14H), 3.54 (t, 2H, *J* = 5.0 Hz, C*H*$_2$CH$_2$N$_3$), 3.42 (t, 2H, *J* = 5 Hz, C*H*$_2$N$_3$), 2.90 (t, 2H, *J* = 5.0 Hz, C*H*$_2$NH$_2$), 1.90 (br s, 2H, N*H*$_2$). The spectroscopic data are in agreement with those found in literature.[70]

5,11,17,23,29,35-Hexa[diacetamido]-37,38,39,40,41,42-hexamethoxycalix[6]arene (17)

The amino calix[6]arene **16** (0.30g, 3.70 mmol) and succinic anhydride (0.29g, 3.96 mmol) were dissolved in DMF (5 ml) and then DIPEA (258 µl, 1.48 mmol) was added. The mixture was irradiated under MW (200 W) for 1-2 h at 40°C. The reaction was monitored via TLC (eluent: DCM/MeOH 7:3) and then evaporated under reduced pressure to afford **17** as a brownish solid (yield = 77%).

^{1}H-NMR (300 MHz, MeOD) δ (ppm): 7.22 (s, 12H, Ar*H* meta); 3.92 (s, 12H, ArC*H*$_2$Ar); 3.27 (s, 18H, OC*H*$_3$); 2.61 (m, 24H, NHC*H*$_2$C*H*$_2$NH). **^{13}C-NMR** (300 MHz, MeOD) δ (ppm): 174.9 (*C*ONH); 171.1 (*C*ONH); 152.4 (C-ipso); 134.7 (C-orto); 133.9 (C-para); 120.7 (C-meta); 59.7 (O*C*H$_3$); 30.9 (NH*C*H$_2$*C*H$_2$NH); 28.6 (Ar*C*H$_2$Ar). **ESI-MS:** m/z 1537 [100%(M+6Me+K)$^+$].

M.p. > 185°C dec

5,11,17,23,29,35-Hexa[(14-azide-3,6,9,12-tetraoxatetradecane-1-amino)diacetamido]-37,38,39,40,41,42-hexamethoxycalix[6]arene (18)

A solution of calixarene **17** (0.20g, 0.14 mmol), linker **15** (0.22g, 0.85 mmol), DMAP (0.14g, 1.14 mmol) and EDC(0.18g, 0.97 mmol) in dry DMF (10 ml), under N_2 atmosphere, was stirred at rt for 24h. The reaction was monitored by TLC (eluent: DCM/MeOH 85:15). The mixture was diluted with DCM and washed with brine (3x20 ml) and water (20 ml). The combined organic phases were washed with an aqueous solution of HCl 1M (20 ml), saturated $NaHCO_3$ (20 ml) and then concentrated under vacuum. The crude was purified by flash chromatography (eluent: DCM/MeOH 9:1) to give compound **18** as a yellowish oil (40%).

^{1}H-NMR (400 MHz, MeOD) δ (ppm): 9.52 (s, 6H, CON*H*); 7.92 (s, 6H, CON*H*); 7.20 (s, 12H, Ar*H* meta); 3.83 (br s, 12H, ArC*H*$_2$Ar); 3.59-3.53 (m, 16H, OC*H*$_2$); 3.46 (t, 2H, J = 5 Hz, ArNHC*H*$_2$); 3.30 (t, 2H, J = 5 Hz, C*H*$_2$N$_3$); 3.21 (s, 18H, OC*H*$_3$); 2.57 (s, 12H, NHC*H*$_2$*CH*$_2$NH); 2.50 (s, 12H, NH*CH*$_2$C*H*$_2$NH). **^{13}C-NMR** (400 MHz, MeOD) δ (ppm): 173.2 (*C*ONH); 171.1 (*C*ONH); 152.3 (C-ipso), 134.8 (C-orto), 134.2 (C-para), 120.4 (C-meta), 70.2-69.2 (*C*H$_2$O); 59.9 (O*C*H$_3$); 48.1

(*C*H_2NH); 39.1 (*C*H_2N_3); 31.7(NH*C*H_2*C*H_2NH); 30.0 (Ar*C*H_2Ar). **ESI-MS:** m/z 1460.8 [100%$(M+2Na)^{2+}$].

1,2,3,4,6-penta-O-acetyl-β-D-galactopyranoside (19)

To a solution of sodium acetate (19.12 g, 0.233 mol) and acetic anhydride (80.0 ml; 0.824 mol) at 130°C, D-galactose (10 g, 55.51 mmol) was added and the mixture was allowed to reacto for 3h. The reaction was followed via TLC (eluent: hexane/EtOAc 4:6). When it was finished a H_2O/ice bath (5-fold volume respect to the organic phase) was added and the mixture was stirred for 24h. The precipitate was then filtered on Buchner funnel and washed with hexane. The solid was triturated with EtOH for 1h 30 min and then filtered again to afford the tetra acetylated galactose as a white solid (yield = 35%).

^{1}H-NMR (400 MHz, $CDCl_3$) δ (ppm) 5.70 (d, 1H, J_{1-2} = 8.0 Hz, H_1); 5.42 (d, 1H, *J* = 3.4 Hz, H_4); 5.36-5.31 (m, 1H, H_2); 5.07 (dd, 1H, J_{2-3} = 10.4 Hz, J_{3-4} = 3.4 Hz, H_3); 4.18-4.10 (m, 2H, H_{6a},H_{6b}); 4.05 (t, 1H, *J* = 6.5 Hz, H_5); 2.15 (s, 3H, OAc); 2.11(s, 3H, OAc); 2.03 (s, 3H, OAc); 1.98 (s, 3H, OAc).

The chemical-physical data are in agreement with those found in literature.[84]

1-propargyl-2,3,4,6-tetra-acetyl-β-galactopyranoside (20)[85]

To a solution of peracetylated galactose **19** (1.0 g, 2.56 mmol) in dry DCM (25 ml), propargyl alcohol (194 µl, 3.33 mmol) and BF_3*Et_2O (485 µl, 4.09 mmol) at 0°C, under N_2 atmosphere. The mixture was stirred overnight at rt, for 2h and checked via TLC (eluent: hexane/EtOAc 1:1). The reaction was quenched adding K_2CO_3 (350 mg) and stirred for 30 min at rt. The solid was filtered and washed with DCM (3x40 ml), dried over Na_2SO_4 and the organic layers were concentrated under reduced pressure.[86] The β-product was obtained without purification step as a brownish oil in 97% yield.

^{1}H-NMR (400 MHz, $CDCl_3$): δ (ppm) 5.76 (d, 1H, *J* = 3.0 Hz, H_4); 5.18 (dd, 1H, *J* = 7.9 Hz and 10.4 Hz, H_2); 5.03 (dd, 1H, J_{3-2} = 10.5 Hz, J_{3-4} = 2.4 Hz, H_3); 4.72 (d, 1H, J_{1-2} = 8.0 Hz, H_1); 4.39 (d, 2H, *J* = 2.4 Hz, -C*H*$_2$CCH); 4.51-4.46 (m, 2H, $H_{6a,6b}$); 4.02-3.85 (m, 1H, H_5); 2.45 (s, 1H, CC*H*); 2.17, 2.09, 2.07, 2.01 (s, 12 H, 4xCOC*H*$_3$).

The chemical-physical data are in agreement with those found in literature.[87,88]

1,2,3,4-Tetra-O-acetyl-L-fucopyranoside (21)

To a solution of pyridine (5.72 ml, 70.66 mmol) and acetic anhydride (3.7 ml, 39.0 mmol), L-fucose (0.80g, 4.9 mmol) was added at 0°C. The mixture was kept at this temperature until the complete dissolution of the sugar and then was allowed to go at rt.[89] The reaction was stirred for 24 h and checked via TLC (eluent: hexane/EtOAc 4:6). It was quenched [90] with H_2O/ice for 2h. The aqueous layer was extracted with DCM (2x20 ml)and the combined organic phases were washed with a saturated solution of $NaHCO_3$ (2x20 ml) and water(1x20 ml). After dried over Na_2SO_4, the crude was evaporated under vacuum. Since pyridine was still present the crude was dissolved in DCM and extracted with saturated solution of NaCl. The product was obtained without purification in 99% yield, as a white oil, as a mixture of α/β anomers (1:0.3). NMR data shown are for the α anomer.

^{1}H-NMR (300 MHz, $CDCl_3$): δ (ppm) 6.32 (d, 1H, J_{1-2} = 3.0 Hz, H_1); 5.28–5.25 (m, 3H, H_2, H_3, H_4); 4.29-4.21 (q, 1H, *J* = 6.5 Hz, H_5); 2.12, 2.09, 1.98, 1.95 (s, 12H, COCH_3); 1.14 (d, 3H, *J* = 6.5 Hz, CH_3 fucose). The spectroscopic data are in agreement with those found in literature.[90,91]

1-propargyl-2,3,4,6-tetra-acetyl-α-fucopyranoside (22)

To a solution of peracetylated compound **21** (1.59 g, 4.81 mmol) in dry DCM (40 ml), propargyl alcohol (1,2 ml, 19,2 mmol) and BF_3*Et_2O (2.44 ml, 19.2 mmol) were added at 0°C, under nitrogen atmosphere. The mixture was stirred overnight at rt, for 24h and checked via TLC (eluent: hexane/EtOAc 6:4). The reaction was quenched washing with a saturated solution of $NaHCO_3$ (45 ml) and water (45 ml), then extracted with DCM (3x45 ml), dried over Na_2SO_4 and concentrated under reduced pressure.[86] The crude was purified by flash column chromatography (eluent: hexane/EtOAc 7:3) and the α-anomer was afforded as a white foam in 39% yield.

^{1}H-NMR (400 MHz, $CDCl_3$) δ (ppm) 5.18-5-13 (m,2H, H_3, H_4); 5.06 (d, 1H, J_{1-2} = 2.9 Hz, H_1); 4.96 (dd, 1H, J_{2-3} = 10.7 Hz, , J_{2-1} = 3.0 Hz, H_2); 4.10-4.04 (m, 4H, CH_2CCH); 3.94 (m, 1H, H_5); 2.40 (s, 1H, CH_2CC*H*); 2.00, 1.91, 1.86 (s, 9H, 3xOAc);0.99 (d, *J* = 6.3 Hz, CH_3 fucose).

The spectroscopic data are the same of those found in literature.[84,92]

General procedure for the "click"conjugation reaction (to obtain compounds 23 and 24)

Calixarene **18** (1 eq) and sugar **20/22** (1.5 eq. for each amino group) were dissolved in DMF, in a microwave tube. $CuSO_4$*$5H_2O$ (0.3 eq), Na ascorbate (0.6 eq) and H_2O (500 µl) were

added. The mixture was heated at 80 °C for 2h (200 W). When the reaction was finished (checked via ESI-MS and TLC, eluent DCM/MeOH 9:1), it was quenched by adding water. The mixture was extracted with EtOAc (2x10 ml). Then the combined organic layers were washed with brine (10 ml) and water (10 ml). The organic solvent was removed under reduced pressure and the crude was purified by flash chromatography to afford pure protected glycocalix[6]arenes.

5,11,17,23,29,35-Hexa[(14-azide-3,6,9,12-tetraoxatetradecane-1-amino)diacetamido]-37,38,39,40,41,42-hexamethoxycalix[6]arene-2,3,4,6-Tetra-*O*-acetyl-β-D-fucopyranosyl) (24)

The crude was purified via flash chromatography (eluent: CH_2Cl_2/MeOH 94:6) to obtain the pure product as a yellow oil (30%). **^{1}H-NMR** (400 MHz, MeOD/$CDCl_3$ 9:1, 50°C) δ (ppm): 8.00 (bs, 6H, C*H* triazole); 7.22 (s, 12H, Ar*H*); 5.33 (dd, 6H, J_{1-2}= 3.2 Hz, J_{2-3} = 10.8 Hz, H_2); 5.28 (br s, 6H, H_4); 5.16 (d, 6H, J_{1-2}= 3.6 Hz, H_1); 5.06 (dd, 6H, J_{3-4}= 3.6 Hz, J_{2-3} = 10.8 Hz, H_3); 4.80 (d, 6H, *J* = 12.4 Hz, OCH*H*-triazole); 4.68 (d, 6H, *J* = 12.4 Hz, OC*H*H-triazole); 4.55 (m, 12H, NHCH_2); 4.27-4.22 (m, 6H, H_5); 3.92-3.88 (m, 24H, ArH_2Ar and NHCH_2CH_2); 3.60-3.52 (m. 96H, OCH_2); 3.38-3.33 (m, 12H, CH_2CH_2triazole); 3.19 (s, 18H, OCH_3); 2.60 (dd, 24H, *J* = 6 and 20 Hz, NHCH_2CH_2NH); 2.15, 1.99, 1.96 (s, 54H, COCH_3); 1.12 (d, 18H, *J* = 6.4 Hz, CH_3 fucose). **^{13}C-NMR** (100 MHz, MeOD/$CDCl_3$ 9:1, 50°C) δ (ppm): 173.2 (*C*ONH); 171.1 (*C*ONH); 170.8, 170.6, 170.3 (3x*C*OCH_3); 152.6 (C-ipso); 143.2 (C_q triazole); 134.7 (C-ortho); 134.0 (C-para); 124.8 (*C*H triazole); 120.8 (C-meta); 95.6 (C_1); 71.3 (C_4); 71.2-68.1 (O*C*H_2); 68.0 (C_2); 67.9 (C_3); 64.7 (C_5); 60.6 (O*C*H_2-triazole); 59.9 (O*C*H_3); 50.1 (CONH*C*H_2); 39.2 (*C*H_2CH_2-triazole); 31.2 (*C*H_2NHCO); 30.9 (*C*H_2NHCO); 29.3 (Ar*C*H_2Ar); 19.5, 19.4, 19.3 (3xCO*C*H_3); 14.9 (*C*H_3 fucose). **ESI-MS (+):** m/z 1639.3 [100%; $(M+3Na)^{3+}$]; m/z 1234.3 [40%; $(M+Ca)^{2+}$].

5,11,17,23,29,35-Hexa[(14-azide-3,6,9,12-tetraoxatetradecane-1-amino)diacetamido]-37,38,39,40,41,42-hexamethoxycalix[6]arene-2,3,4,6-Tetra-*O*-acetyl-β-D-galactopyranosyl) (24)

The crude was purified by flash chromatography (eluent: DCM/MeOH 9:1) to give glycocalixarene **23** as a brownish oil (50%).**^{1}H-NMR** (300 MHz, MeOD) δ (ppm): 8.01 (bs, 6H, C*H* triazole); 7.25 (s, 12H, Ar*H*); 5.40 (d, 6H, *J* = 2.4 Hz, H_4); 5.11 (br s, 12H, H_2, H_3); 4.93-4.76 (m, 18H, H_1, OCH_2-triazole); 4.56 (t, 12H, *J* = 5 Hz, NHCH_2); 4.16 (m, 18H, H_5, $H_{6a,b}$); 3.87 (br s,

12H, $NHCH_2CH_2$); 3.57-3.52 (m, 72H, OCH_2); 3.52 (t, 12H, *J* = 5 Hz, CH_2CH_2triazole); 3.36-3.32 (m, 24H, CH_2CH_2triazole, $ArCH_2Ar$); 3.27 (s, 18H, OCH_3); 2.60 (d, 12H, *J* = 6, $NHCH_2CH_2NH$); 2.55 (d, 12H, *J* = 6, $NHCH_2CH_2NH$); 2.15, 2.05, 1.99, 1.95 (s, 72H, 4x$COCH_3$). **^{13}C-NMR** (100 MHz, MeOD) δ (ppm): 173.2 (*C*ONH); 171.1 (*C*ONH); 170.8, 170.6, 170.3 (3x$COCH_3$); 152.6 (C-ipso); 143.2 (C_q triazole); 134.7 (C-ortho); 134.0 (C-para); 124.8 (*C*H triazole); 120.8 (C-meta); 95.6 (C_1); 71.3 (C_4); 71.2-68.1 (OCH_2); 68.0 (C_2); 67.9 (C_3); 64.7 (C_5); 60.6 (OCH_2-triazole); 59.9 (OCH_3); 50.1 ($CONHCH_2$); 39.2 (CH_2CH_2-triazole); 31.2 (CH_2NHCO); 30.9 (CH_2NHCO); 29.3 ($ArCH_2Ar$); 19.5, 19.4, 19.3 (3x$COCH_3$); 14.9 (CH_3 fucose). **ESI-MS (+):** m/z 1753.3 [100%; $(M+3Na)^{3+}$].

General procedure for the deacetylation of compounds 23-24

The protected glycocalixarene was dissolved in MeOH and then cooled down to 0°C. Thus solid CH_3ONa was added and the solution was stirred at 0°C for 1-3 h until complete deactylation (ESI-MS). Amberlite IR-120 was added till neutral pH (30 min). The resin was filtered off and the solvent was removed under reduced pressure.

Deprotected Fucose calix[6]arene (6)

The crude was triturated in DCM to afford compound **7** as a yellow oil (81%). **^{1}H-NMR** (400 MHz, MeOD) δ (ppm): 8.03 (s, 6H, C*H* triazole); 7.26 (s, 12H, Ar*H*); 4.90 (s, 6H, H_1); 4.77 (d, 6H, *J* = 17.0 Hz, OCH*H*-triazole); 4.62 (d, 6H, *J* =12.4 Hz, OC*H*H-triazole); 4.53 (m, 12H, $NHCH_2$); 4.00-3.57 (m, 120H, H_2, H_3, H_4, H_5, OCH_2, $ArCH_2Ar$); 3.33 (br s, 12H, CH_2CH_2-triazole); 3.30 (s, 18H, OCH_3); 2.68 (m, 24H, $NHCH_2CH_2NH$); 1.19 (d, 18H, *J* = 8.5 Hz, CH_3 fucose). **^{13}C-NMR** (100 MHz, MeOD) δ (ppm): 173.4 (*C*ONH); 171.1 (*C*ONH); 152.3 (C-ipso); 144.1 (C_q triazole); 134.8 (C-ortho); 134.2 (C-para); 124.6 (*C*H triazole); 120.4 (C-meta); 98.8 (C_1); 72.2-68.3.4 (C_2, C_3, C_4, C_5, CH_2 chain); 60.3 (OCH_2-triazole); 59.9 (OCH_3); 50.0 ($CONHCH_2$); 39.1 (CH_2CH_2-triazole); 31.7 (CH_2CO); 30.6 (CH_2CO); 29.9 ($ArCH_2Ar$); 15.4 (CH_3 fucose). **ESI-MS (+):** m/z 1639.3 [100%; $(M+3Na)^{3+}$]; m/z 1234.3 [40%; $(M+Ca)^{2+}$].

Deprotected Galactose calix[6]arene (7)

The crude was triturated in DCM to give compound **6** as a yellow-brown oil (95%). **^{1}H-NMR** (300 MHz, DMSO-d_6) δ (ppm): 9.75 (br s, 6H, ArCON*H*); 7.91 (t, 6H, *J* = 5.6 Hz, $CONHCH_2$);

7.21 (s, 12H, Ar*H*); 4.91 (br s, 6H, O*H*); 4.82 (d, 6H, *J* = 12 Hz, OCH*H*-triazole); 4.62 (br s, 6H, O*H*); 4.56 (d, 6H, *J* = 12 Hz, OC*H*H-triazole); 4.46 (t, 12H, *J* = 5 Hz, NHCH_2); 4.41 (s, 6H, O*H*); 4.20 (d, 6H, *J* = 6.8 Hz, H_1); 3.81 (t, 12H, *J* = 5 Hz, $NHCH_2CH_2$); 3.64 (s, 6H, O*H*); 3.54-3.15 (m, 162H, H_2, H_3, H_4, H_5, OCH_2, ArCH_2Ar, OCH_3); 2.50 (br s, 12H, NHCH_2CH_2NH); 2.36 (br s, 12H, NHCH_2CH_2NH). **^{13}C-NMR** (75 MHz, DMSO-d_6) δ (ppm): 172.0 (*C*ONH); 170.7 (*C*ONH); 151.7 (C-ipso); 144.5 (C_q triazole); 134.9 (C-ortho); 134.5 (C-para); 125.0 (*C*H triazole); 120.4 (C-meta); 103.2 (C_1); 75.8 (C_4); 74.0 (C_3); 71.0-69.1 (C_2, C_5, OCH_2); 61.8 (OCH_2-triazole); 60.9 (CH_2NH-triazole); 59.9 (OCH_3); 40.7 (CONHCH_2); 39.6 (ArCH_2Ar); 31.9 (CH_2CO); 30.6 (CH_2CO).
ESI-MS (+): m/z 1418.6 [100%; $(M+3Na)^{3+}$]; m/z 1069.5 [60%; $(M+Ca)^{2+}$].

4.4 References

(1) Ferentz, A. E.; Wagner, G. *Q. Rev. Biophys.* **2000**, *33* (1), 29–65.

(2) Zuiderweg, E. R. *Biochemistry* **2002**, *41* (1), 1–7.

(3) Takeuchi, K.; Wagner, G. *Curr. Opin. Struct. Biol.* **2006**, *16* (1), 109–117.

(4) Clarkson, J.; Campbell, I. D. *Biochem. Soc. Trans.* **2003**, *31*, 1006–1009.

(5) Meyer, B.; Peters, T. *Angew. Chemie - Int. Ed.* **2003**, *42* (8), 864–890.

(6) Williamson, M. P. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 1–16.

(7) Fielding, L. *Prog. Nucl. Magn. Reson. Spectrosc.* **2007**, *51* (4), 219–242.

(8) Bodner, C. R.; Dobson, C. M.; Bax, A. *J. Mol. Biol.* **2009**, *390* (4), 775–790.

(9) Lian, L.-Y. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *71*, 59–72.

(10) Zuiderweg, Erik R. P., Hamers, Laurentius F. , Rollema, Harry S., de Bruin, Simon H. and Hibers, C. W. *Eur. J. Biochem.* **1981**, *118*, 95–104.

(11) Hajduk, P. J.; Augeri, D. J.; Mack, J.; Mendoza, R.; Yang, J.; Betz, S. F.; Fesik, S. W. *J. Am. Chem. Soc.* **2000**, *122* (33), 7898–7904.

(12) Pervushin, K.; Riek, R.; Wider, G.; Wüthrich, K. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94* (23), 12366–12371.

(13) Keifer, P. A. *Prog. Drug Res.* **2000**, *55*, 137–211.

(14) Wyss DF, McCoy MA, S. M. *Curr. Opin. Drug Discov. Devel.* **2002**, *5* (4), 630–647.

(15) Skinner, A. L.; Laurence, J. S. *J. Pharm. Sci.* **2008**, *97* (11), 4670–4695.

(16) MCPHERSON, A. *Methods* **2004**, *34* (3), 254–265.

(17) Luft, J. R.; Collins, R. J.; Fehrman, N. A.; Lauricella, A. M.; Veatch, C. K.; Detitta, G. T. *J. Struct. Biol.* **2003**, *142*, 170–179.

(18) D'Arcy, A.; Sweeney, A. Mac; Haber, A. *Methods* **2004**, *34* (3), 323–328.

(19) Hassell, A. M.; An, G.; Bledsoe, R. K.; Bynum, J. M.; Carter, H. L.; Deng, S. J. J.; Gampe, R. T.; Grisard, T. E.; Madauss, K. P.; Nolte, R. T.; Rocque, W. J.; Wang, L.; Weaver, K. L.; Williams, S. P.; Wisely, G. B.; Xu, R.; Shewchuk, L. M. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2006**, *63* (1), 72–79.

(20) Bolanos-Garcia, V. M.; Chayen, N. E. *Prog. Biophys. Mol. Biol.* **2009**, *101* (1-3), 3–12.

(21) Cudney, R., Patel, S., Weisgraber, K., Newhouse, Y. and McPherson, A. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1994**, No. Che 276, 414–423.

(22) Asherie, N. *Methods* **2004**, *34* (3), 266–272.

(23) Gronenborn, a M.; Filpula, D. R.; Essig, N. Z.; Achari, a; Whitlow, M.; Wingfield, P. T.; Clore, G. M. *Science* **1991**, *253* (5020), 657–661.

(24) Byeon, I. L.; Louis, J. M.; Gronenborn, A. M. *J. Mol. Biol.* **2003**, *2836* (3), 141–152.

(25) Gallagher, T.; Alexander, P.; Bryan, P.; Gilliland, G. L. *Biochemistry* **1994**, *33*, 4721–4729.

(26) Alexander, P.; Fahnestock, S.; Lee, T.; Orban, J.; Bryan, P. *Biochemistry* **1992**, *31* (14), 3597–3603.

(27) Lindman, S.; Xue, W.-F.; Szczepankiewicz, O.; Bauer, M. C.; Nilsson, H.; Linse, S. *Biophys. J.* **2006**, *90* (April), 2911–2921.

(28) Smith, C. K.; Withka, J. M.; Regan, L. *Biochemistry* **1994**, *33* (18), 5510–5517.

(29) Sancho, J. *Cell. Mol. Life Sci.* **2006**, *63*, 855–864.

(30) Vervoort, J.; Heering, D.; Peelen, S.; Van Berkel, W. *Methods Enzymol.* **1985**, *243* (2), 188–203.

(31) Smillie R.M. *Biochem. Biophys. Res. Commun.* **1965**, *20* (5), 621–629.

(32) Bianchi, V.; Eliasson, R.; Fontecave, M.; Mulliez, E.; Hoover, D. M.; Matthews, R. G.; Reichard, P. *Biochemical and biophysical research communications*. 1993, pp 792–797.

(33) Goñi, G.; Herguedas, B.; Hervás, M.; Peregrina, J. R.; De la Rosa, M. a.; Gómez-Moreno, C.; Navarro, J. a.; Hermoso, J. a.; Martínez-Júlvez, M.; Medina, M. *Biochim. Biophys. Acta - Bioenerg.* **2009**, *1787* (3), 144–154.

(34) Hall, D. a; Vander Kooi, C. W.; Stasik, C. N.; Stevens, S. Y.; Zuiderweg, E. R.; Matthews, R. G. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (17), 9521–9526.

(35) Fuhrmann, M. *Journal of Biological Chemistry*. 1995, pp 19158–19165.

(36) Inda, L. A.; Luisa Peleato, M. *Phytochemistry* **2003**, *63* (3), 303–308.

(37) Cremades, N.; Bueno, M.; Toja, M.; Sancho, J. *Biophys. Chem.* **2005**, *115* (2-3), 267–276.

(38) Astuti, Y.; Topoglidis, E.; Briscoe, P. B.; Fantuzzi, A.; Gilardi, G.; Durrant, J. R. *J. Am. Chem. Soc.* **2004**, *126* (25), 8001–8009.

(39) Osborne, C.; Chen, L. M.; Matthews, R. G. *J. Bacteriol.* **1991**, *173* (5), 1729–1737.

(40) Ponstingl, H.; Otting, G. *Eur. J. Biochem.* **1997**, *244*, 384–399.

(41) Hoover, D. M.; Ludwig, M. L. *Protein Sci.* **1997**, *6*, 2525–2537.

(42) Cremades, N.; Velázquez-Campoy, A.; Martínez-Júlvez, M.; Neira, J. L.; Pérez-Dorado, I.; Hermoso, J.; Jiménez, P.; Lanas, A.; Hoffman, P. S.; Sancho, J. *ACS Chem. Biol.* **2009**, *4* (11), 928–938.

(43) Sancho, J. *J. Med. Chem.* **2013**, *56*, 6248–6258.

(44) George, J. M. *Genome Biol.* **2002**, *3* (1).

(45) Maroteaux, L.; T. Campanelli, J.; H.Scheller, R. *J. Neurosci.* **1988**, *8*, 2804–2815.

(46) Goedert, M. *Nat. Rev. Neurosci.* **2001**, *2* (7), 492–501.

(47) Lücking, C. B.; Brice, A. *C. Cell. Mol. Life Sci* **2000**, *57*, 1894–1908.

(48) El-Agnaf, O. M. A. *FASEB J.* **2006**, *20* (3), 419–425.

(49) Rodrìguez-Navarro, D.N., Dardanelli, M.S., Rìz-Saìnt, J. E. *FEMS Microbiol .Lett.* **2007**, *272* (2), 127–136.

(50) Kostlánová, N.; Mitchell, E. P.; Lortat-Jacob, H.; Oscarson, S.; Lahmann, M.; Gilboa-Garber, N.; Chambat, G.; Wimmerová, M.; Imberty, A. *J. Biol. Chem.* **2005**, *280* (30), 27839–27849.

(51) Sudakevitz, D.; Kostlánová, N.; Blatman-jan, G.; Mitchell, E. P.; Lerrer, B.; Wimmerová, M.; Katcoff, D. J.; Imberty, A.; Gilboa-garber, N. *Mol. Microbiol.* **2004**, *52* (3), 691–700.

(52) Arnaud, J.; Tröndle, K.; Claudinon, J.; Audfray, A.; Varrot, A.; Römer, W.; Imberty, A. *Angew. Chemie - Int. Ed.* **2014**, 9267–9270.

(53) McGovern, R. E.; Fernandes, H.; Khan, A. R.; Power, N. P.; Crowley, P. B. *Nat. Chem.* **2012**, *4* (7), 527–533.

(54) McGovern, R. E.; McCarthy, A. a.; Crowley, P. B. *Chem. Commun.* **2014**, *50* (72), 10412.

(55) McGovern, R. E.; Snarr, B. D.; Lyons, J. a.; McFarlane, J.; Whiting, A. L.; Paci, I.; Hof, F.; Crowley, P. B. *Chem. Sci.* **2015**, *6* (1), 442–449.

(56) Bagnacani, V.; Sansone, F.; Donofrio, G.; Baldini, L.; Casnati, A.; Ungaro, R. *Org. Lett.* **2008**, *10* (18), 3953–3956.

(57) Dudic, M.; Colombo, A.; Sansone, F.; Casnati, A. *Tetrahedron* **2004**, *60*, 11613–11618.

(58) Baldini, L.; Melegari, M.; Bagnacani, V.; Casnati, A.; Dalcanale, E.; Sansone, F.; Ungaro, R. *J. Org. Chem.* **2011**, *76*, 3720–3732.

(59) André, S.; Sansone, F.; Kaltner, H.; Casnati, A.; Kopitz, J.; Gabius, H. J.; Ungaro, R. *ChemBioChem* **2008**, *9* (10), 1649–1661.

(60) Sansone, F.; Dudic, M.; Donofrio, G.; Rivetti, C.; Baldini, L.; Casnati, A.; Cellai, S.; Ungaro, R. *J. Am. Chem. Soc.* **2006**, *128* (8), 14528–14536.

(61) Arosio, D.; Fontanella, M.; Baldini, L.; Mauri, L.; Bernardi, A.; Casnati, A.; Sansone, F.; Ungaro, R. *J. Am. Chem. Soc.* **2005**, *127* (11), 3660–3661.

(62) Gordo, S.; Martos, V.; Santos, E.; Menéndez, M.; Bo, C.; Giralt, E.; de Mendoza, J. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 16426–16431.

(63) Casnati, A.; Sansone, F.; Ungaro, R. *Acc. Chem. Res.* **2003**, *36* (4), 246–254.

(64) Arena, G.; Casnati, A.; Contino, A.; Lombardo, G. G.; Sciotto, D.; Ungaro, R. *Chem. - A Eur. J.* **1999**, *5*, 738–744.

(65) Han, C.; Yu, G.; Zheng, B.; Huang, F. *Org. Lett.* **2012**, *14* (7), 1712–1715.

(66) Arnecke, R.; Böhmer, V.; Cacciapaglia, R.; Dalla Cort, A.; Mandolini, L. *Tetrahedron* **1997**, *53* (13), 4901–4908.

(67) Sansone, F.; Chierici, E.; Casnati, A.; Ungaro, R. *Org. Biomol. Chem.* **2003**, *11*, 1802–1809.

(68) Eliezer, D.; Kutluay, E.; Bussell, R.; Browne, G. *J. Mol. Biol.* **2001**, *307* (4), 1061–1073.

(69) Iyer, S. S.; Anderson, A. S.; Reed, S.; Swanson, B.; Schmidt, J. G. *Tetrahedron Lett.* **2004**, *45*, 4285–4288.

(70) Schwabacher, A. W.; Lane, J. W.; Schiesher, M. W.; Leigh, K. M.; Johnson, C. W. *J. Org. Chem.* **1998**, *63* (5), 1727–1729.

(71) Hein, J. E.; Fokin, V. V. *Chem. Soc. Rev.* **2010**, *39* (4), 1302.

(72) Jaime, C.; De Mendoza, J.; Prados, P.; Nieto, P. M.; Sanchez, C. *J. Org. Chem.* **1991**, *56* (10), 3372–3376.

(73) Studier, F. W. *Protein Expr. Purif.* **2005**, *41*, 207–234.

(74) Lindman, S.; Linse, S.; Mulder, F. a a; André, I. *Biochemistry* **2006**, *45* (47), 13993–14002.

(75) Croke, R. L.; Sallum, C. O.; Watson, E.; Watt, E. D.; Alexandrescu, A. T. *Protein Sci.* **2008**, *17*, 1434–1445.

(76) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6* (3), 277–293.

(77) Vranken, W. F.; Boucher, W.; Stevens, T. J.; Fogh, R. H.; Pajon, A.; Llinas, M.; Ulrich, E. L.; Markley, J. L.; Ionides, J.; Laue, E. D. *Proteins Struct. Funct. Genet.* **2005**, *59* (4), 687–696.

(78) Evans, P. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2006**, *62* (1), 72–82.

(79) Collaborative Computational Project, N. 4. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1994**, *50*, 760–763.

(80) Emsley, P.; Cowtan, K. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *D60*, 2126–2132.

(81) Casnati, A.; Ca, D.; Fontanella, M.; Sansone, F.; Ugozzoli, F.; Ungaro, R.; Liger, K.; Dozol, J. *European J. Org. Chem.* **2005**, *11*, 2338–2348.

(82) Bagnacani, V.; Franceschi, V.; Fantuzzi, L.; Casnati, A.; Donofrio, G.; Sansone, F.;

Ungaro, R.; Organica, C.; Parma, U.; Area, P. *Bioconjug. Chem.* **2012**, *23*, 993–1002.

(83) Brauch, S.; Henze, M.; Osswald, B.; Naumann, K.; Wessjohann, L. a.; van Berkel, S. S.; Westermann, B. *Org. Biomol. Chem.* **2012**, *10*, 958.

(84) Wahler, D.; Boujard, O.; Lefèvre, F.; Reymond, J. L. *Tetrahedron* **2004**, *60* (3), 703–710.

(85) Percec, V.; Leowanawat, P.; Sun, H. J.; Kulikov, O.; Nusbaum, C. D.; Tran, T. M.; Bertin, A.; Wilson, D. a.; Peterca, M.; Zhang, S.; Kamat, N. P.; Vargo, K.; Moock, D.; Johnston, E. D.; Hammer, D. a.; Pochan, D. J.; Chen, Y.; Chabre, Y. M.; Shiao, T. C.; Bergeron-Brlek, M.; André, S.; Roy, R.; Gabius, H. J.; Heiney, P. a. *J. Am. Chem. Soc.* **2013**, *135* (24), 9055–9077.

(86) Nierengarten, J.; Iehl, J.; Oerthel, V.; Holler, M. *Chem. Commun.* **2010**, *46*, 3860–3862.

(87) Schmid, S.; Mishra, A.; Bäuerle, P. *Chem. Commun. (Camb).* **2011**, *47* (4), 1324–1326.

(88) Mereyala, H. B.; Gurrala, S. R. *Carbohydr. Res.* **1998**, *307* (3-4), 351–354.

(89) Wolfrom, M. L.; Orsino, J. A. *J. Am. Chem. Soc.* **1934**, *56*, 985–987.

(90) Babiuch, K.; Dag, A.; Zhao, J.; Lu, H.; Stenzel, M. H. *Biomacromolecules* **2015**, *16* (7), 1948–1957.

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

(92) Béha, S. Giguère, D. Patnam, R. Roy, R. *Synlett* **2006**, *11*, 1739–1743.