Calixarene-based multivalent ligands

L. Baldini, A. Casnati,* F. Sansone and R. Ungaro*

Received 12th October 2006
First published as an Advance Article on the web 6th November 2006
DOI: 10.1039/b603082n

Multivalency is a powerful concept which explains the strong binding observed in biological systems and guides the design and synthesis of ligands for self-assembly and molecular recognition in Chemistry. The phenol–formaldehyde cyclic oligomers, called calixarenes, have been used as scaffolds for the synthesis of multivalent ligands thanks to the fact that they have a variable number of reactive positions for attaching the ligating functions, well defined conformational properties and, in some cases, cavities of molecular dimensions eventually able to encapsulate guest species. This *tutorial review* illustrates the fundamental aspects of multivalency and the properties of calixarene-based multivalent ligands in lectin binding and inhibition, DNA condensation and cell transfection, protein surface recognition, self-assembly, crystal engineering, and nanofabrication.

General aspects of multivalency

In Biology a ligand is usually a low molecular weight species which interacts with a high molecular weight biological entity (receptor, enzyme, nucleic acid, *etc.*), whereas in Supramolecular Chemistry this term is used in a broader sense, indicating any species which self-associates or binds to other species through noncovalent interactions. In order to enhance the binding efficiency and selectivity of the designed ligands, chemists usually try to optimise the complementary matching (in terms of size and nature of binding forces) between the interacting species. However, quite recently they started to take advantage of a new powerful concept, *multivalency*, which is the ability of a particle (or molecule) to bind another particle (or molecule) *via* multiple simultaneous noncovalent interactions (Fig. 1).[1] The valency is therefore the number of ligating functionalities of the same or similar types connected to each of these entities.

Multivalent interactions usually result in high specificity and in thermodynamic and kinetic stability (much higher than those arising from a simple monovalent interaction). They control many important biological processes[1,2] such as the adhesion of cells, viruses/bacteria or antibodies/macrophages to the cell surface. Nature exploits multivalency to convert relatively weak interactions (*e.g.* carbohydrate–protein interactions) into strong and specific recognition events. In the biomimetic approach to Drug Design, multivalency has also been exploited to successfully obtain inhibitors of some of these pathogens, allowing the development of therapeutic agents able to neutralise bacterial toxins or to prevent viral or bacterial infections.

Dip.to di Chimica Organica e Industriale, Università degli Studi, V.le G. P. Usberti 17/A, 43100 Parma and Consorzio INSTM, Via Giusti 9, 50121 Firenze, Italy

Laura Baldini

Laura Baldini (1974) studied Chemistry at the University of Parma and graduated in 1998 under the supervision of Prof. Rocco Ungaro, after developing part of her research thesis at the University of Jyväskylä with Prof. Kari Rissanen. She obtained her PhD with Prof. Rocco Ungaro in 2002, in the field of self-assembly and molecular recognition properties of functionalised calixarenes. Part of her doctoral studies was carried out at the University of Sheffield (UK) with Prof. Chris A. Hunter. In 2002 she joined the group of Prof. Andrew D. Hamilton at the Yale University (USA) as a post-doctoral fellow working in the field of protein surface recognition by synthetic receptors and since 2005 she has been a researcher at the University of Parma.

Alessandro Casnati

Alessandro Casnati (Milano, 1963) obtained his Laurea in Chemistry in 1987 and his PhD in 1991 from the University of Parma under the supervision of Prof. Rocco Ungaro. In 1990 he spent part of his PhD studies at the Twente University (NL) in the laboratory of Prof. David N. Reinhoudt. In 1994 he was appointed researcher and in 1998 associate professor of Organic Chemistry at Parma University. In 1997 he was awarded the ''G. Ciamician'' medal of the Italian Chemical Society. His research interests are in the field of supramolecular chemistry and particularly in the synthesis and study of molecular receptors for ions and neutral molecules.

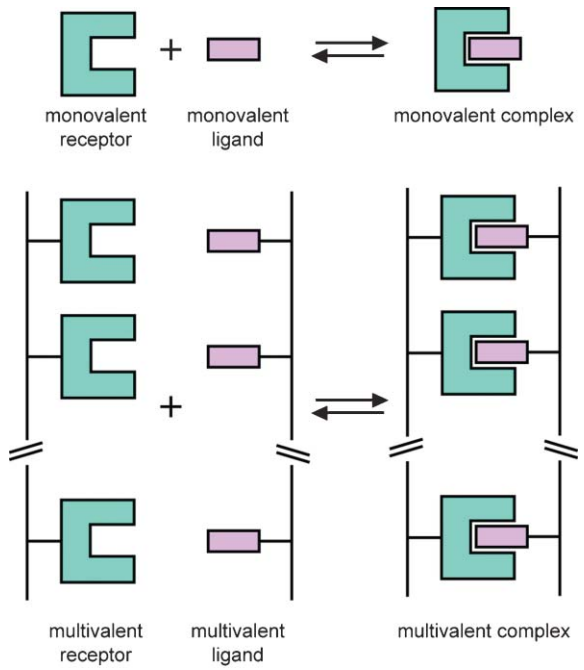


Fig. 1 Example of monovalent *vs.* multivalent complexes.

In the last decade, also workers in the field of Supramolecular Chemistry became more and more interested in applying the multivalency concept to the recognition of biologically important molecules and to nanotechnology. Multivalency, in fact, has different and attractive features common to the supramolecular concept of self-assembly, such as reversibility, self-sorting and self-repairing and the possibility to reach high thermodynamic and kinetic stability.[3] Virtually any of the noncovalent interactions can be used to form multivalent complexes. Moreover, supramolecular systems, sometimes simpler than the natural ones, can help the understanding and quantitative description of the multivalent effect.[4]

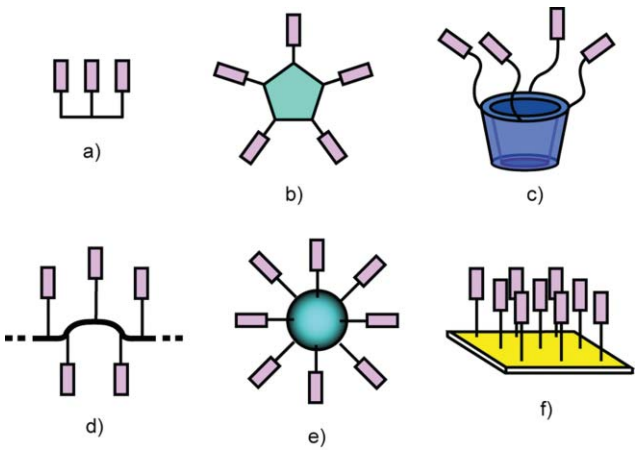


Fig. 2 Different topologies of multivalent ligands. a) 1-D linear arrangement, b) 2-D cyclic/macrocyclic; c) 3-D cavity containing scaffolds (cyclodextrins/calixarenes); d) polymers/peptoids; e) nanoparticles/dendrimers/liposomes; f) 2-D self-assembled monolayers (SAM) on Au/quartz.

A multivalent ligand (Fig. 2) consists of a main core, called the *scaffold*, bearing several covalent connections, *linkers* or *spacers*, to the peripheral *ligating* (binding) units.

Any multivalent scaffold can in principle be used, from those having low valency such as benzene derivatives, monosaccharides, transition metal complexes, azamacrocycles, cyclodextrins or calixarenes (Fig. 2a–c) to high valency ones such as dendrimers, polymers, peptoids, proteins, micelles, liposomes, and self-assembled monolayers (SAMs) on nanoparticles or plane surfaces (Fig. 2d–f).

Often, the concept of cooperativity, generally intended as the influence of the binding of one ligand on the receptor's affinity toward further binding, is associated with multivalency.[1] However, as elegantly pointed out by Ercolani,[5] cooperativity in multivalent systems is extremely scarce. Moreover, although cooperativity can be easily and rigorously

Francesco Sansone

Francesco Sansone was born in 1968 in Parma and graduated in Chemistry in 1993 at Parma University, spending a period of the undergraduate thesis at the Twente University (NL) with Prof. David N. Reinhoudt. In 1995 he developed part of his research on carbohydrates in the group of Prof. Alessandro Dondoni at Ferrara University and in 1998 he received his Ph.D in Organic Chemistry working under the supervision of Prof. Rocco Ungaro. Currently he is a researcher at Parma University, and his scientific interests are in the field of supramolecular chemistry and in particular on the synthesis of calixarenes functionalised with amino acids and carbohydrates.

Rocco Ungaro

Rocco Ungaro was born in Anzi (PZ), in the South of Italy, and studied Chemistry at the University of Parma where he graduated in 1968 in the laboratory of Prof. Giuseppe Casnati. After a short stay in pharmaceutical industry, in 1970 he returned to Parma University as lecturer, becoming an associate professor in Organic Chemistry in 1982 and full professor in 1986. In 1974/75 Prof. Ungaro was post-doctoral fellow at the State University of New York (SUNY) in Syracuse (N.Y.) in the laboratory of Professor Johannes Smid. He has pioneered the chemistry of calixarenes and is currently active in the use of these macrocycles in supramolecular and bioorganic chemistry. In 2002 he received the Research Prize of the Division of Organic Chemistry of the Italian Chemical Society for his studies on organic structure and molecular interactions.

assessed by using Hill or Scatchard plots for the binding of monovalent ligands to a multivalent receptor (the most famous case being the binding of four O_2 molecules to tetrameric haemoglobin), much more difficult is to prove its existence when a multivalent ligand interacts with a multivalent receptor.

In general, to describe multivalent binding, an approach based on the additivity of the free energies[6,7] can be used. The standard binding free energy for multivalent binding $\Delta G^{\circ}_{\text{multi}}$ is

$$\Delta G^{\circ}_{\text{multi}} = n\Delta G^{\circ}_{\text{mono}} + \Delta G^{\circ}_{\text{interaction}}$$

where $\Delta G^{\circ}_{\text{mono}}$ is the standard binding free energy of the corresponding monovalent interaction, n is the valency of the complex and $\Delta G^{\circ}_{\text{interaction}}$ is the balance between favourable and unfavourable effects of tethering.

A rather qualitative but often useful parameter, β, has been introduced by Whitesides *et al.*,[1] as $\beta = K_{\text{multi}}/K_{\text{mono}}$, where K_{multi} and K_{mono} are the association constants for the multivalent and monovalent complexes, respectively. The parameter β, named *enhancement factor*, has been often used in the literature to compare the efficiency of ligands having different topology and/or valency: molecules with high β values are efficient ligands/inhibitors. Sometimes, if the valency n of the complex is known, this enhancement factor can be normalised to n, giving rise to the parameter β/n.

More rigorous approaches have been proposed to quantitatively describe the multivalent binding process. Kitov and Bundle,[7] for instance, developed a model where the standard free energy of multivalent interaction, $\Delta G^{\circ}_{\text{avidity}}$, is a function of three terms $\Delta G^{\circ}_{\text{inter}}$, $\Delta G^{\circ}_{\text{intra}}$ (binding free energies for the first intermolecular and the second intramolecular process, respectively, see Fig. 3) and a statistical term $\Delta S^{\circ}_{\text{avidity}}$, namely *avidity entropy*, calculated on the basis of the topology of the complex. The avidity entropy can grow rapidly with the valency of the complex, always favouring binding, which explains why multivalency can overcome the loss of conformational entropy. By a nonlinear fitting of the measured binding energies for a series of multivalent ligands, $\Delta G^{\circ}_{\text{inter}}$ and $\Delta G^{\circ}_{\text{intra}}$ can be determined, thus allowing to design and maximise the avidity of multivalent ligands.

Very important is also the choice of the spacer which links the scaffold to the ligating units. It should be of the proper length to allow the simultaneous binding of all the ligating groups, without generating enthalpic unfavourable strains (enthalpically diminished binding).

Alternative models proposed for the analysis of multivalency are based on the concepts of *effective concentration* (C_{eff}) and *effective molarity* (*EM*). C_{eff} is defined as the concentration of ligating units locally experienced by one of the neighbouring receptor sites upon intramolecular binding (Fig. 4), whereas C_{sol} is the analytical concentration of ligating units present in solution. If $C_{\text{eff}} > C_{\text{sol}}$ intramolecular binding takes place, whereas if $C_{\text{sol}} > C_{\text{eff}}$ the formation of intermolecular aggregates is favoured. Interestingly, C_{eff} can be theoretically calculated, being a function of some physical parameters of the ligand and receptor units, by using a combination of probability functions and molecular modeling

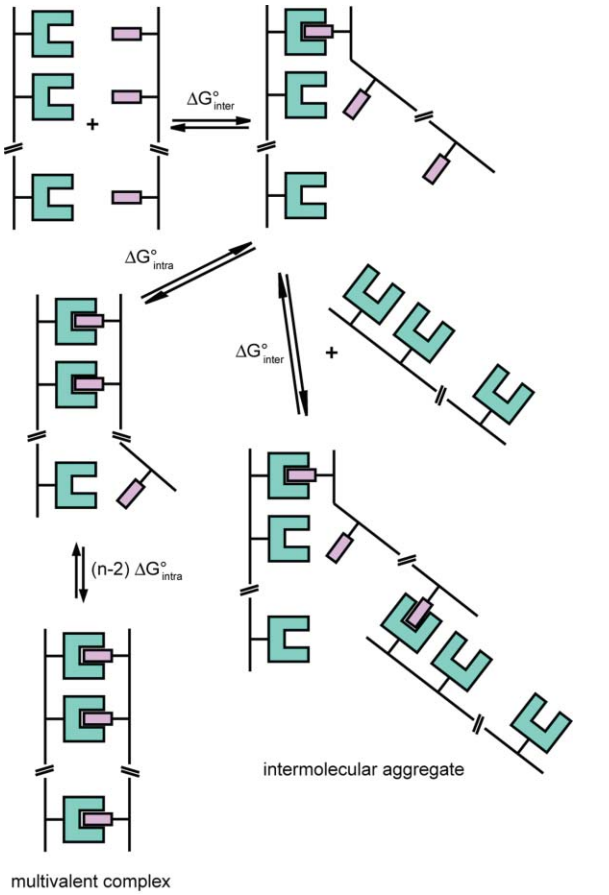


Fig. 3 Intermolecular and intramolecular processes for the formation of a multivalent complex or of an intermolecular aggregate.

calculations.[3] With the assumptions that: (i) binding sites are equivalent, (ii) no cooperativity and (iii) no linker–receptor interactions are present, the binding constant of the multivalent complex K_{multi}, is then given by the equation

$$K_{\text{multi}} = b(\text{K}_{\text{i}})^{n}\,(C_{\text{eff}})^{n-1}$$

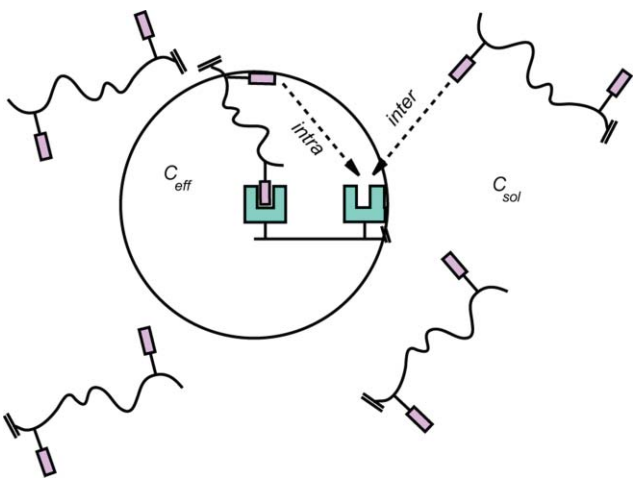


Fig. 4 Intramolecular (multivalent) *vs.* intermolecular binding as a function of C_{eff} and C_{sol}.

where K_i is the intrinsic binding constant and b is a statistical factor representing the number of permutations which allow the conversion of the reagents into products divided by the number of permutation to convert the products back.

Effective Molarity, EM, on the other hand, is the experimental ratio between the intra- and intermolecular association constants ($EM = K_{intra}/K_{inter}$), thus measuring the advantage of the first process taking place over the second one. However, the concepts of C_{eff} and EM are quite similar. The first one can be calculated from the characteristics of the ligand and the receptor, while the latter is usually determined experimentally. In the case of non-cooperative binding, C_{eff} equals EM. Alternatively, EM values much higher than C_{eff} are indicative of a positive cooperativity, while negative cooperativity may occur in the opposite case.

Finally, the kinetics of multivalent interactions is usually characterised by a step-wise dissociation pathway having extremely low dissociation rates. This also explains why the use of very high concentration of monovalent competing species is necessary to avoid rebinding of the partially dissociated multivalent complex.

Calixarenes as multivalent scaffolds

In this tutorial we will exclusively focus our attention on the use of calix[*n*]arenes[8] (Fig. 5) as scaffolds. These macrocycles, obtained by the oligomerisation of phenol and formaldehyde, offer several advantages. For instance, their valency can be easily varied at least from 1 to 8, while the stereochemical orientation of the ligating arms can be properly tuned by shaping, for example, the calix[4]arene macrocycle in one of its four possible conformations (Fig. 6).

On the other hand, the larger and conformationally more mobile calix[6]- and -[8]arenes can adapt their conformations to the stereochemical request of the multivalent receptor entity (induced fit). Moreover, the chemistry of these macrocycles is nowadays well known and efficient procedures are available to easily functionalise both the lower (phenolic OH groups) and the upper (aromatic nuclei) rim (Fig. 5) with most of the functional groups (–CHO, –COOH, $–NH_2$, –NCS, –C≡CH …*etc.*) usually employed for the conjugation of ligating units.[9]

Multivalent glycocalixarenes

The synthesis of multivalent calixarene-based ligands is particularly attractive when the ligating units are carbohydrates. In

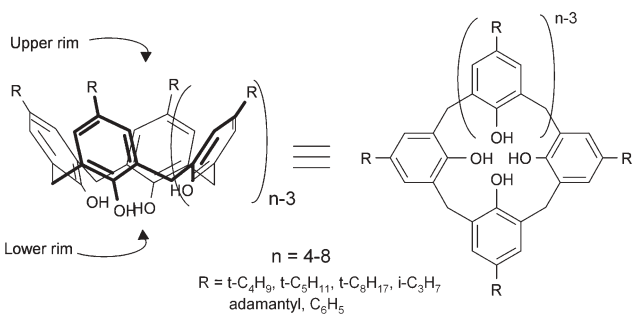


Fig. 5 Calix[*n*]arenes.

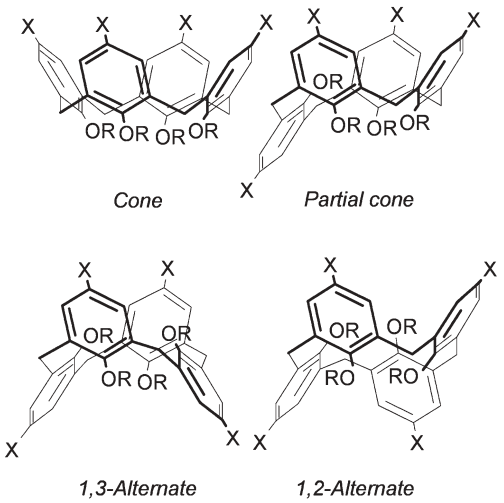
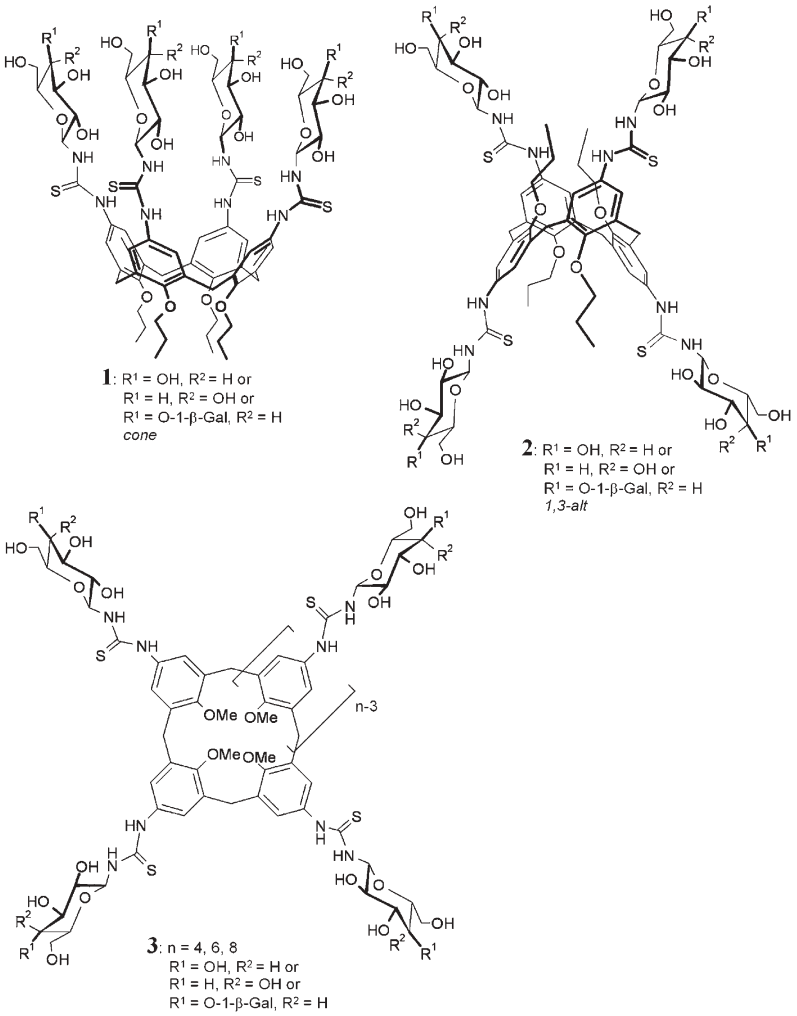


Fig. 6 The four conformations of a calix[4]arene: X and/or R contain the ligating units.

fact, although saccharides have been considered for a long time only as a chemical energy source and as structurally important elements, they are now identified as fundamental substrates for specific receptors in a wide range of biological processes.[2] Intercellular communication, cell trafficking, immune response, infections by bacteria and viruses, growth and metastasis of tumour cells, all occur due to the binding of the sugar residues present on the cell surface by saccharide receptors. In addition, since the affinity of a single carbohydrate unit for its receptor is usually low, the strong binding observed in these recognition events is determined by the simultaneous complexation of several identical glycoside residues, exposed at the substrate surface, by receptors (often proteins) bearing several equivalent binding sites. This particular aspect of multivalency has also been named *glycoside cluster effect*.[10]

Proteins able to recognise carbohydrates of glycoproteins are named lectins. Lectins are present on the surface of cells, bacteria and viruses and are responsible for several physiological and pathological events. For example, the influenza viruses use the trimeric lectin haemagglutinin to target the cell surface, while it has been proposed that the metastasis mechanisms of some tumour cells are based on the selectin (C-type lectin) adhesion system. Lectins are usually present in oligomeric form and thus multivalent glycosylated ligands incorporating their specific glycosylated substrates are extremely attractive as inhibitors of several of these phenomena.

We have first exploited thiourea groups as spacers for the synthesis of glycocalixarenes, since they can be efficiently obtained from easily available glycosyl isothiocyanates and aminocalixarenes with controllable stereochemistry at the anomeric position. Moreover, (thio)urea moieties are known to efficiently bind anions also in polar solvents and, when present at the upper rim of calixarenes, are potentially able to bind anionic drugs and specifically deliver them to a target protein. We have therefore synthesised a series of glycosylthioureido-calix[4]arenes fixed in the cone (*e.g.* **1**)[11] or 1,3-alternate (**2**) structure or having a mobile (**3**, $n = 4$) conformation, and the calix[6]- and -[8]arene (**3**, $n = 6, 8$)



analogues, bearing both mono- (Glc and Gal) and disaccharides (Lac).[12]

These multivalent compounds showed selectivity in lectin binding. Simple and widely used turbidimetric experiments evidenced the specific agglutination ability towards Concanavalin A (ConA) and Peanut Agglutinin by the glucose- and the galactose-bearing macrocycles, respectively. Similar calix[4]- and calix[8]arene glycoclusters, exposing *N*-acetylglucosamine (GlcNAc) units,[13] were successfully tested as ligands of the Wheat Germ Agglutinin, a GlcNAc binding protein, and inhibitors of its agglutination properties towards human erythrocytes. In all cases an amplified lectin affinity with respect to the corresponding simple monosaccharide was observed.

The combination of saccharides with a calixarene scaffold confers a significant amphiphilic character to the resulting glycocalixarenes. Their solubility in water is up to 10^{-3} M, but they tend to aggregate in micelles and/or vesicles, even when very short alkyl chains (*i.e.* methyl) are present at the lower rim and, more significantly, in phosphate buffer.[12] This phenomenon is evidenced by the broadening of the ^{1}H NMR signals in D_2O, by the unambiguous detection by atomic force microscopy (AFM) on mica surfaces of discoid-like aggregates[12] and through dynamic light scattering (DLS) experiments.[14] This behaviour must be carefully taken into account because the valency in action (Fig. 7) could be significantly different and strongly enhanced compared to that expected simply on the basis of the number of saccharide units covalently linked to the monomeric species. Additionally, these glycocalixarenes were shown by NMR spectroscopy and MS to interact with anionic substrates through the thiourea units.[11] This is an important result which establishes the basis for the design and synthesis of site-directed drug delivery systems, exploiting carbohydrate recognition, for high organ or tissue specificity.

Several multivalent glycoclusters have been applied to some toxins and, more in general, to bacterial targets.[4] Many

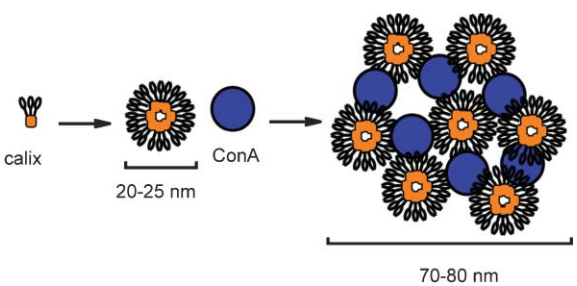
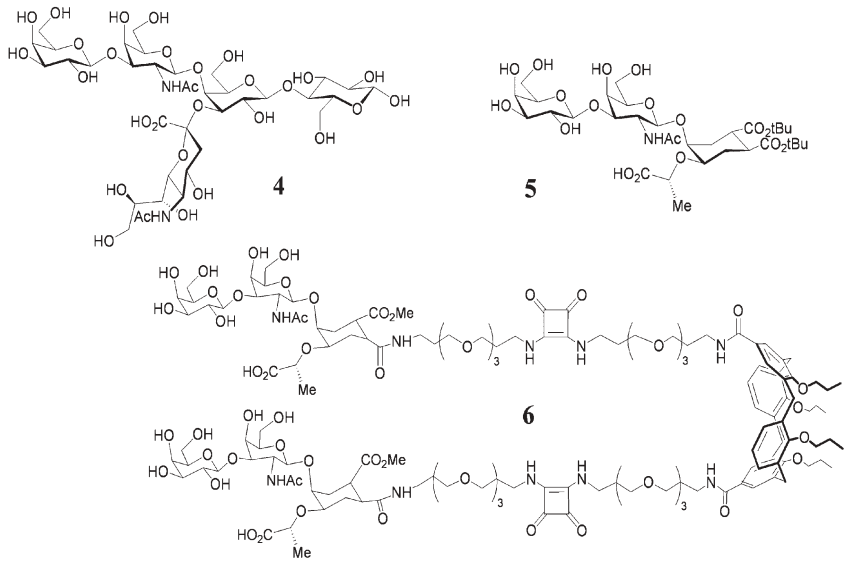


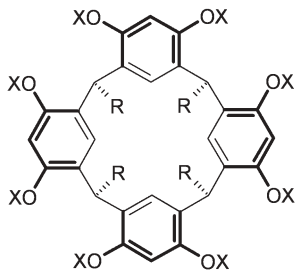
Fig. 7 Self-aggregation of glycocalixarenes and interaction with lectins.



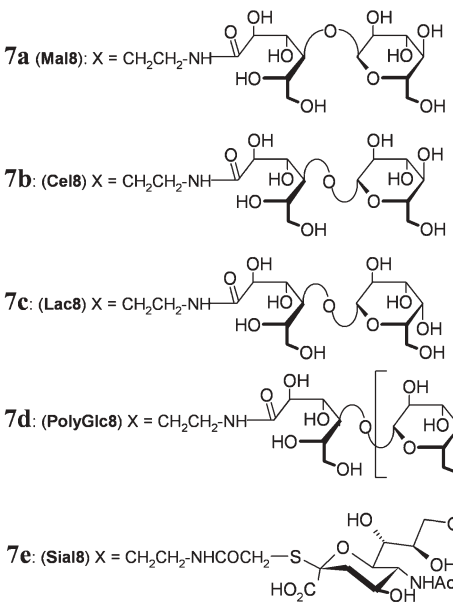
bacterial toxins, such as *Escherichia coli* enterotoxin, pertussis, Shiga and Cholera toxins can indeed be classified as lectins. Cholera toxin (CT) and its homologous heat-labile toxin of *E. coli* are toroid-like AB_5 toxins. The B proteins simultaneously bind to five units of their natural ligand, the GM1 ganglioside, present on the cell membrane. The recognition process primarily involves the galactose and the sialic acid molecules of the GM1 oligosaccharide (o-GM1, **4**), kept in the correct orientation by a lactose scaffold. Recently, Bernardi *et al.*[15] designed and prepared an interesting synthetic mimic (pseudo-GM1, **5**) of the pentasaccharide contained in the natural o-GM1. Pseudo-GM1, in which sialic acid and lactose are replaced by R-lactic acid and cyclohexandiol, respectively, binds to CT with a K_d = 190 μM, which is higher than that of the natural ligand o-GM1 (K_d = 219 nM), but much lower if compared with simple galactose ($K_d = 4 \times 10^4$ μM) and other o-GM1 mimics.[15]

This prompted us to design and synthesise, in collaboration with Bernardi's group, a series of calixarene-based multivalent ligands having pseudo-GM1 units as toxin binding epitopes. So far, results have been obtained with the divalent ligand **6**, which possesses two pseudo-GM1 units linked in diametral position at the upper rim of a calix[4]arene blocked in the cone conformation. The long spacer based on a polyethyleneglycol bis-amine and squaric acid allows the simultaneous interaction of the two pseudo-GM1 units with two non-adjacent binding sites of the CT.[16] This divalent system resulted, by spectrofluorimetric titration experiments and ELISA inhibition tests, a highly efficient ligand for CT, showing a 50% saturation concentration of 48 nM, slightly better than that measured for the natural o-GM1. This value also corresponds to a remarkable affinity enhancement factor (β) close to 4000 (2000 fold per sugar mimic) with respect to the single pseudo-GM1, which is one of the highest ever observed for synthetic CT inhibitors.

Aoyama and co-workers synthesised a series of resorc[4]-arenes (**7a–d**) having four undecyl chains at the lower rim and eight di- (**7a–c**) or oligosaccharide (**7d**) moieties at the upper rim, by reacting the octamine **7f** with the proper oligosaccharide lactone.[17,18] Also these nonionic, amphiphilic glycocalixarenes, having a quite unusual four legged lipophilic tail, are able to self-aggregate in water solution even in the micromolar region, giving rise to small nanoparticles of exceptional stability and unprecedented properties.



R = $(CH_2)_{10}CH_3$



7f: X = CH_2CH_2-NH_2

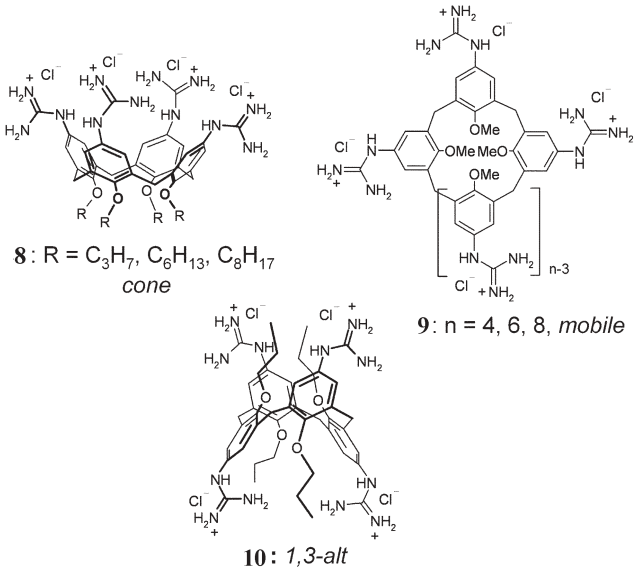
The micellar assemblies, named glycocluster nanoparticles (GNPs), were characterised by NMR, gel permeation

chromatography (GPC), DLS, transmission electron microscopy (TEM) and AFM and showed no notable oligosaccharide chain-length dependent size (2.4–3.5 nm). The molecular weights of the GNPs formed by **7a** and **7d** ($n = 7$) indicate, for example, that they are composed of 4 and 6 molecules, respectively. The unusual stability of GNPs comes from lateral side-by-side inter-saccharide interactions, which immobilise the hydrophobically associated and otherwise labile micelles, pointing out another effect of the multiple presence of saccharides on a single structural core. Moreover, glycocalixarenes of type **7a–d** form complexes in water with a large variety of anionic and cationic dye molecules, having association constants in the range of 10^2–10^6 M^{-1}. Compound **7a** proved to be effective as a molecular delivery system to surfaces such as quartz or gels coated with ConA, octagalactose **7c** is able to deliver fluorescent guests, such as phloxine B or calcein, to hepatocytes, which have receptors for the terminal galactose residues of asialoglycoproteins, while glycocluster **7e**, having eight sialic acid residues successfully tested as a ligand for influenza viruses.[19] The diverse properties of these compounds demonstrate that the selectivity for different biological receptor sites can be finely tuned by changing the saccharide units present on the cluster. These highly specific saccharide-mediated cell recognition events suggest a possible application of glycoclusters in anticancer therapy. If functionalised with the appropriate sugar, they could create a "cancer net" around a solid tumour able to retard its growth and/or inhibit metastasis.

Neutral and charged calixarene multivalent ligands as non-viral gene transfer vectors

Another quite remarkable peculiarity of the resorcarene glycoclusters described in the previous section derives from their ability to interact with phosphates, nucleotides and nucleic acids.[17] While guanosine-5′-monophosphate (GMP) and adenosine mono-, di- or triphosphate (AMP, ADP or ATP) form 1 : 1 complexes with **7a**, HPO_4^{2-} induces formation of large aggregates with an anion/macrocycle ratio higher than 10. The collected evidence clearly indicates that micellar GNPs originating from **7a–d** in pure water are agglutinated or cross-linked in the presence of HPO_4^{2-} to give aggregates whose dimensions (50–140 nm) depend on the oligosaccharide chain length and where the anions (H-acceptors) act as glue for the oligosaccharide (H-donors) units. This property suggested the application of these systems as non-viral gene vectors. Indeed, GNPs are able to compact and coat a plasmid DNA (7040 base pairs) into viral (50 nm) sized particles, named artificial glycoviruses. The aggregation tendency increases in the order Cel8 (**7b**) ≪ Lac8 (**7c**) < Mal8 (**7a**) showing that the alteration in stereochemistry of a single glycoside linkage (Mal *vs.* Cel) or a single OH group (Mal *vs.* Lac), amplified by the presence of a glycosidic cluster, may result in a significant change in size of the glycovirus. These glycoviruses are able to transfect HeLa cells with a remarkable size-regulated activity, the smaller (50 nm) cellobiose glycoviruses being more potent than the larger lactose (200 nm) or maltose (300 nm) ones. Interestingly, for HepG2 cells, which possess receptors for β-galactose, lactose glycoviruses (**7c**) showed a transfection activity 100-fold higher than expected on the basis of size, combining two different effects of their multivalency: a stable aggregation and the specific delivery. Glyconanoparticles represent one of the few examples known of neutral gene transfer synthetic vectors, most of them being cationic[20] since, in this case, the charge–charge interaction with DNA is stronger.

We have recently synthesised a series of guanidinium calixarenes (**8–10**) having different shape and valency and studied their interaction with DNA and transfection properties.[21,22] We were particularly interested in investigating the role played by the macrocycle size, conformation and lipophilicity in DNA binding and condensation. A minimum number of four guanidinium groups is necessary to make these ligands water soluble. Even though differences in the interaction efficiency were evidenced, all the guanidinium calixarenes bind both linear and plasmid DNA, as shown by Electrophoresis Mobility Shift Essay and spectroscopic studies. However, only the calix[4]arene derivatives (**8**, **9** $n = 4$, and **10**) are able to give cell transfection, delivering a 4731 bp plasmid DNA into cells.



AFM investigations allowed the interaction between the different macrocycles and DNA to be directly visualised and the structure of the complexes to be correlated with the transfection properties of the ligands. In fact, images collected with this technique (Fig. 8) show that calix[4]arenes (**8**, **9** $n = 4$, and **10**) condense double stranded DNA in compact blobs which can cross the cell membrane. Conversely, calix[6]- (**9**, $n = 6$) and calix[8]arene (**9**, $n = 8$) originate unmasked and large, multimeric aggregates which cannot be delivered into the cells, thus explaining the absence of transfection for these larger and conformationally mobile ligands. A particularly striking behaviour is shown by the cone calix[4]arenes **8** if compared to the conformationally mobile calix[6]- (**9**, $n = 6$) and calix[8]arene (**9**, $n = 8$) derivatives (Fig. 8). The former, first binds to DNA through guanidinium–phosphate electrostatic interactions and subsequently efficiently condenses a single DNA filament through intramolecular hydrophobic interactions of the liphophilic chains at the lower rim. On the

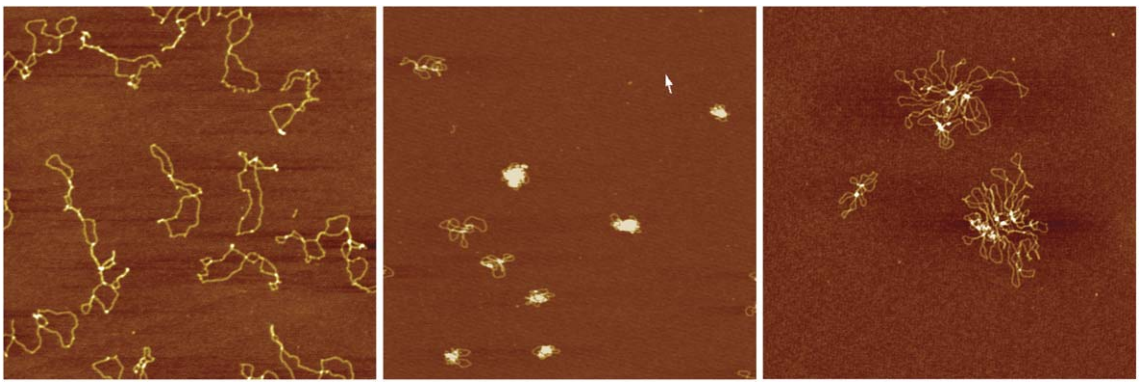


Fig. 8 AFM images of 1 nM supercoiled plasmid DNA (left), condensed blobs of 1 nM supercoiled plasmid DNA in presence of 1 μM calix[4]arene **8** (middle) and large networks of 1 nM supercoiled plasmid DNA in presence of 1 μM calix[8]arene **9**, $n = 8$ (right). Each image represents a 2 × 2 μm scan.

contrary, only electrostatic interactions, which give rise to large interfilament networks, operate in the case of calix[6]- and calix[8]arene derivatives. An intermediate behaviour is shown by the 1,3-alternate derivative **10** and by the conformationally mobile calix[4]arene (**9**, $n = 4$). This is the first time in which DNA condensation, directly visualised through AFM, has been correlated with transfection and conformational properties of multivalent cationic macrocycles.

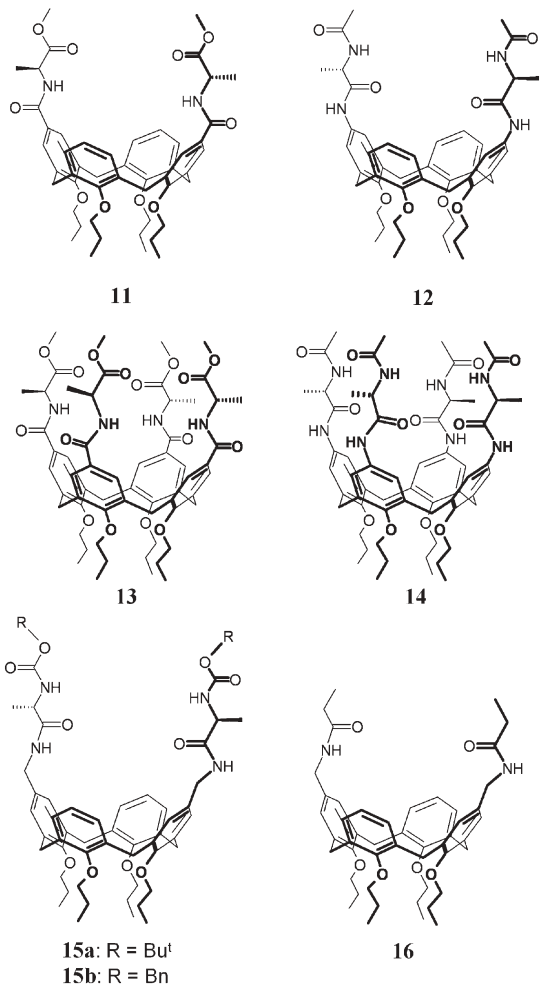
Peptidocalixarenes in self-assembly and protein surface recognition

Multi-armed, calixarene-based ligands, potentially able to give multivalent interactions, can be conveniently obtained also by functionalising the macrocycle with amino acids or small peptides,[9] which are well known for their tendency to form α-helix and β-sheet superstructures and are involved in numberless recognition phenomena at biological level. Amino acids or peptides can be linked to the calixarene scaffold either through the terminal amino or carboxylic acid group, leading to *N*-linked (*e.g.* **11** and **13**) or *C*-linked (*e.g.* **12** and **14**) peptidocalixarenes, respectively.

Self-assembled nanotubes and dimeric capsules

In general, the upper rim, cleft-like peptidocalix[4]arene podands in the cone conformation, both di- (*e.g.* **11**[23] and **12**[24]) and tetrafunctionalised (*e.g.* **13**[23] and **14**[24]), experience in solution intramolecular hydrogen bonding or uncontrolled intermolecular self-association.

Nevertheless, in the case of the difunctionalised *C*-linked peptidocalixarenes **15** and **16**, the presence of two identical ligating groups containing amide and carbamate functionalities proved to be a useful tool to control molecular assembly at the solid state and achieve robust self-assembled architectures.[25] This is an interesting feature in supramolecular chemistry and crystal engineering, since the self-assembly at the solid state of molecular components having a concave structure may offer the possibility of designing functional porous materials. Compounds **15a** and **15b**, where the two amino-protected amino acids are linked through a methylene spacer to opposite aromatic rings of a



tetrapropoxycalix[4]arene, self-assemble in the lattice through a two-dimensional network held together by hydrogen bonds between the amide and carbamate groups of adjacent molecules. In particular, compound **15a**, functionalised with two units of Boc-alanine, is connected to four adjacent molecules through eight hydrogen bonds (Fig. 9): both NH groups of each arm interact with the amide CO group of two adjacent molecules and, contemporarily, the CO amide groups act as acceptor of two hydrogen bonds donated by the two

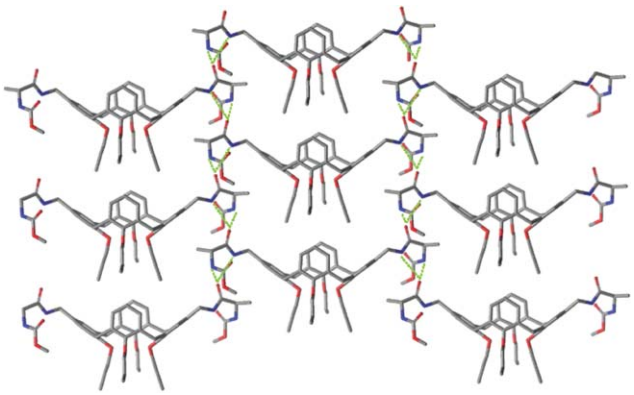


Fig. 9 Self-assembled nanotubes formed in the solid state by cleft-like *C*-linked peptido-calix[4]arenes.

NHs of two other adjacent molecules. This hydrogen bonding motif gives rise to an infinite array of parallel nanotubes formed by the piling of the calixarene cavities and held together by the interdigitation of the amino acid chains.

Despite the presence of a different *N*-protecting group and although it crystallises as two conformers, differing for the orientation of the amino acid chains, the Cbz-alanine compound **15b** originates a substantially similar supramolecular architecture at the solid state. Both conformers, in fact, self-assemble in the lattice through hydrogen bonds forming 2D networks of nanotubes analogous to that described for **15a**. Through the formation of 4 hydrogen bonds involving all the amide groups, also calixarene **16** self-assembles in the lattice with a quite comparable behaviour, with the calixarene cavities disposed, in this case, in a "zig-zag" arrangement. This indicates that the presence of a pair of such amide groups at the upper rim of a cone calix[4]arene is the minimum requested feature for driving this packing motif, the robustness of which is increased when additional hydrogen bonding groups are present in the peptide chains.

Conversely, the introduction of two amino acids or small peptides in opposite orientation (one *C*-linked and the other *N*-linked) at the upper rim of a calix[4]arene amino acid (**17a**) afforded divalent ligands (**18**–**20**) which are self-complementary (the *N*-linked arm is complementary to the *C*-linked arm and *vice versa*) in terms of hydrogen bonding donor and acceptor groups (Fig. 10). As a consequence, in apolar solutions they self-assemble in dimeric structures held together by hydrogen bonds between the CO and NH groups[26] (Fig. 10). The hydrogen bonding array is analogous to an antiparallel β-sheet and brings together the two aromatic cavities in a capsule-like structure. The dimerisation constants could be easily measured by ^{1}H NMR dilution experiments and increase with the number of CO and NH groups (K_{dim} = 74, 105, 776 and 1460 M^{-1} in $CDCl_3$ for compounds **17b**, **18a**, **19** and **20**, respectively). The introduction of bulkier side chains such as those of Phe (**18b**) or Leu (**18c**) causes a decrease in the dimerisation constant (K_{dim} = 63 and 82 M^{-1}, respectively).

Also the trivalent peptidocalixarene **21**, synthesised by de Mendoza, based on a calix[6]arene containing three leucyl amide residues at the lower rim gives rise to a dimeric structure

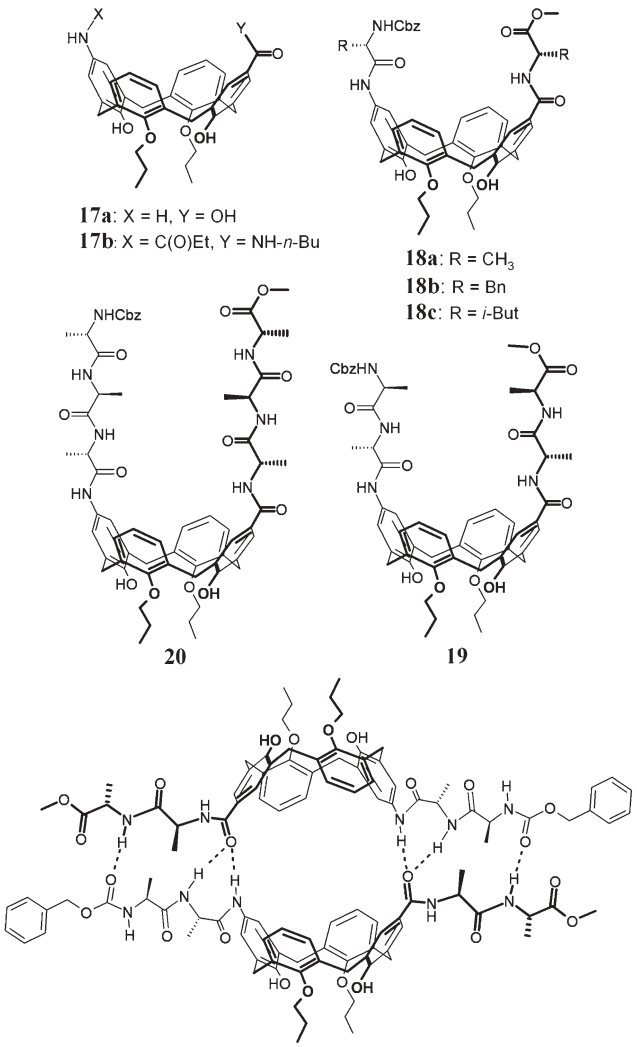
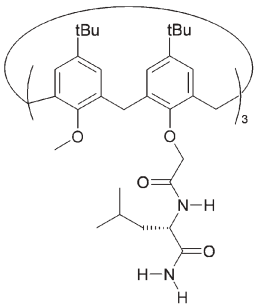


Fig. 10 *C,N*-Linked peptidocalix[4]arenes and the dimer which they form in apolar solvents.

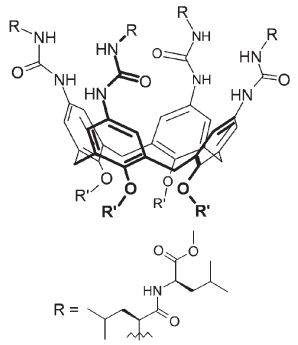
held together exclusively by hydrogen bonding between the amino acids.[27] By molecular mechanics and dynamics calculations, together with NMR studies, it was established that the dimerisation is driven by the formation of up to eighteen hydrogen bonds between the amino acids in an extended conformation. By ^{1}H NMR dilution experiments, the dimerisation constant was measured as K_{dim} = 640 M^{-1} in $CDCl_3$. The inclusion of guests was not observed, probably because the methoxy groups and one of the amino acid flexible chains occupy the free space within the peptide cage.

A secondary, yet interesting role is played by the four units of L-Leu-D-Leu-OMe dipeptide in the tetravalent calix[4]arene ureidopeptide **22**, synthesised with the aim of introducing additional ligating units and enlarging the capsule size[28] with respect to the well known tetraurea calix[4]arenes (R = Ph, Bn in **22**), which form in apolar solvents highly stable, non-covalent dimers held together by a circular array of 16 hydrogen bonds.[29–31] Compared to the latter ones, the capsule structure formed by **22**, elucidated by ROESY NMR, is reinforced by an additional seam of hydrogen bonds provided by the peptide side chains, including the ester carbonyl groups.

The dimerisation constant, measured by ^{1}H NMR dilution experiments, is 20 M^{-1} in CD_2Cl_2 and 5100 M^{-1} in toluene-d_8.



21



22

L-Leu-D-Leu

Protein surface recognition

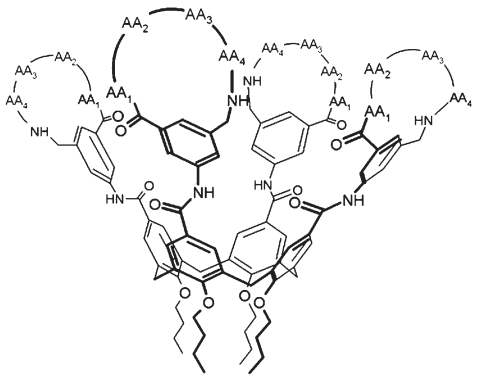
Polyvalent peptidocalixarenes (**23**) have been synthesised by Hamilton to target protein surfaces and inhibit clinically important protein-protein interactions.[32] The design is based on the arrangement of four cyclic pseudo-peptides around the calixarene hydrophobic cavity and is inspired by the observation that a common feature of many natural protein–protein interfaces is, on each partner, a hydrophobic region surrounded by a ring of polar, charged and hydrogen bonding residues. Changing the amino acids AA_1–AA_4 of the cyclic pseudo-peptide results in a modification of the recognition characteristics of the receptors. The calixarene scaffold has the double task of anchoring the ligand groups and providing an additional hydrophobic binding site.

Compound **23a**, with the anionic Gly-Asp-Gly-Asp peptide loops, strongly interacts with the surface of cytochrome c near the heme edge, the region where the protein interacts with its natural partners (such as cytochrome c peroxidase). This region contains a large hydrophobic area, surrounded by positively charged Lys residues, which is complementarily matched by the hydrophobic and anionic features of the calixarene ligand. Moreover, this compound is a potent slow binding inhibitor (with submicromolar activity) of chymotrypsin, an important serine protease, presumably by binding to a patch of several cationic residues which is found near the active site cleft of the enzyme.

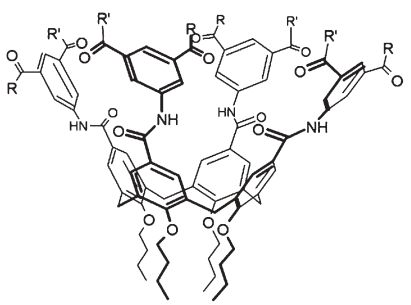
Ligand **23b** (peptide sequence Gly-Asp-Gly-Tyr) is able to bind (K_d is in the nanomolar range) the platelet-derived growth factor (PDGF), whose receptor-binding region is composed of cationic and hydrophobic residues. *In vivo* studies showed that the treatment of nude mice bearing human tumours with **23b** resulted in a significant inhibition of the tumour growth and angiogenesis. The mechanism involved is based on the binding of the synthetic ligand to the growth factor which thus becomes unable to interact with its membrane bound receptor. The lack of growth factor–receptor interaction inhibits the cell signalling pathways which are triggered by this event and would eventually lead to cell growth and angiogenesis. In a similar way, compound **23c**, which has the amino acid sequence Gly-Lys-Gly-Lys, selectively disrupts the binding of the vascular endothelial growth factor (VEGF) to its receptor and inhibits angiogenesis both *in vitro* and *in vivo*, and tumorigenesis and metastasis *in vivo*.[33] More recently, from a second-generation library of multivalent calixarene derivatives, **24** was identified as a potent pharmacological agent which blocks the function of both VEGF and PDGF, inhibiting at the same time the initiation and maintenance of blood vessels around a tumour.[34] Following a similar approach, Cunsolo *et al.* prepared a few calix[8]arene derivatives functionalised with eight basic amino acid residues which are complementary to the surface of tryptase near its active site and are active as tryptase inhibitors[35] and Neri and co-workers synthesised a small library of calix[4]arenes bearing four tetrapeptides at the upper rim which was screened for inhibition activity towards tissue and microbial transglutaminase.[36]

Multivalent calixarenes at surfaces

Guanidinium calix[4]arenes, similar to those used as gene transfer vectors and described above, were very useful also in nanofabrication,[37–39] where clear evidence of the important role played by multivalency was obtained.[37] The group of Huskens and Reinhoudt has introduced and developed the concept of "Molecular printboard",[40] which can be defined as a self-assembled monolayer on a surface that has recognition sites to which molecules can be anchored through specific interactions. Monolayers of β-cyclodextrins self-assembled on gold, silicon wafers and glass (β-CD SAMs) have been particularly useful in this context. We have contributed to

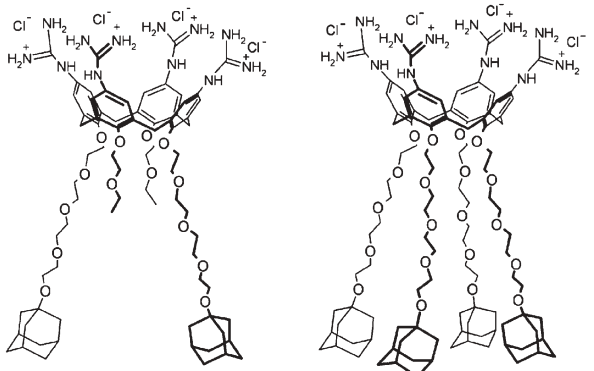


23 **a**: $AA_1 = AA_3$ = Gly, $AA_2 = AA_4$ = Asp
b: $AA_1 = AA_3$ = Gly, AA_2 = Asp, AA_4 = Tyr
c: $AA_1 = AA_3$ = Gly, $AA_2 = AA_4$ = Lys



24 R = OH; R' = OBn

this project by attaching two (**25**) or four (**26**) adamantane groups (Ad) to the lower rim of tetraguanidinium calix[4]-arenes in the cone conformation. These macrocycles are water soluble and their supramolecular chemistry is complementary to that of β-CD SAMs. In fact, it is well known that adamantane derivatives are strongly bound to β-cyclodextrin (β-CD) in water ($K_a \approx 10^5$ M^{-1}) through hydrophobic interactions.



25 **26**

The nature and length of the spacer was chosen in order to ensure water solubility and prevent non specific interactions in the complexation studies. The binding of the divalent ligand **25** (guest) to β-CD (host) was first studied in solution by microcalorimetry, which gave evidence of the formation of a 2 : 1 host : guest complex, implying that **25** is bound by two β-CD cavities with an intrinsic association constant $K_i = 4.6 \times 10^4$ M^{-1}, very similar to that found for the interaction of adamantyl derivatives and β-CD in water. This experiment also proved that the adamantane end groups in ligand **25** bind independently to two β-CDs. The interaction of ligand **25** with a β-CD SAM leads to the formation of a strong divalent complex which was studied by surface plasmon resonance (SPR). An estimated value of the association constant ($K_a \approx 10^{10}$ M^{-1}) was obtained, which is three orders of magnitude higher than that of the 1 : 1 complex of **25** with an EDTA-tethered β-cyclodextrin dimer **27**, used as a reference for divalent binding. This result was rationalised by Huskens' and Reinhoudt's group using a theoretical model which interprets the divalent binding as two consecutive independent monovalent binding events, *i.e.* an intermolecular interaction followed by an intramolecular binding, the latter being associated with an effective concentration term (C_{eff}) accounting for the close proximity of the two interacting species on the surface. These results, together with data obtained with other multivalent systems, were instrumental in developing a general and useful model of multivalency on surfaces[41] which expands the scope of this quite important concept. Moreover, the strong divalent binding experienced by ligand **25** on β-CD SAMs allowed interesting applications in nanofabrication to be developed.[39] By using supramolecular micro-contact printing (μCP) and dip pen nanolithography (DPN) in which **25** was employed as ''ink'', stable patterns were written on molecular printboards represented by β-CD terminated SAMs on gold and silicon oxide. Interestingly, the resolution achieved in writing lines on the molecular printboard was as low as 60 ± 20 nm.

The tetravalent ligand **26** was used to assemble a supramolecular capsule on the molecular printboard, exploiting charge-charge interactions[38] (Fig. 11). When equimolar amounts of the water soluble calix[4]arene tetraguanidinium tetraadamantane **26** and calix[4]arene tetrasulfonate **28**, both in the cone conformation, were mixed together, a precipitate was formed as a consequence of the neutralisation of charges upon capsule formation and of the presence of the four lipophilic adamantyl groups, which further decrease the water solubility of the supramolecular assembly. Upon addition in solution of β-CD, which forms an inclusion complex with the adamantyl units of **26**, the capsule dissolves in water and a

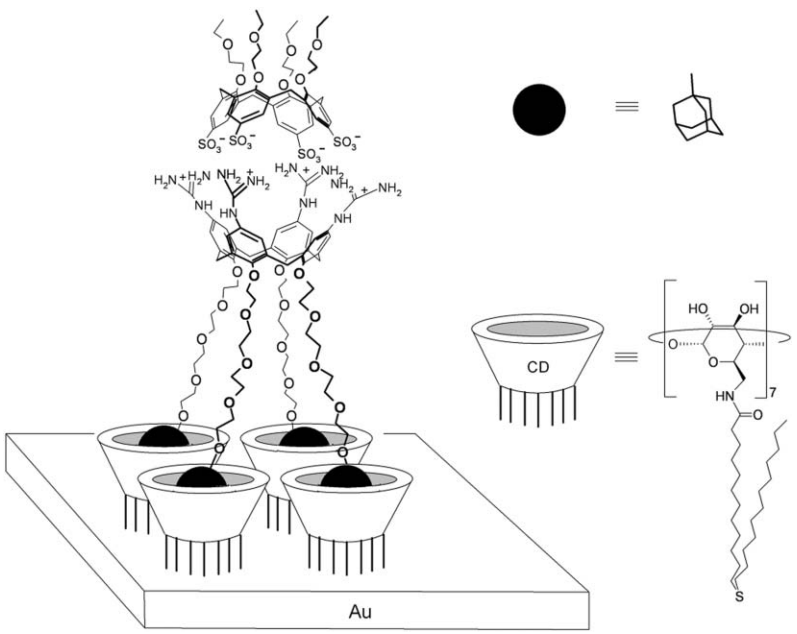
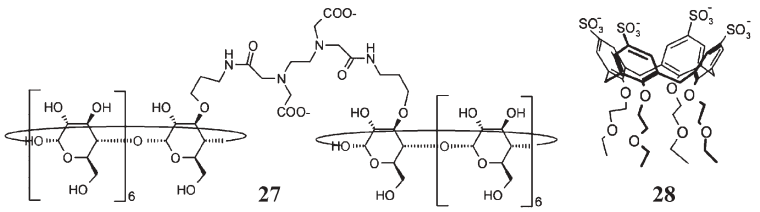


Fig. 11 Schematic representation of the molecular capsule **26**·**28** on the SAM on gold.



clear solution is obtained. Evidence for a capsule formation in MeOH and D_2O containing 10^{-2} M β-CD was obtained by electrospray ionisation (ESI) mass spectrometry and ^1H NMR, respectively. The strength of the capsule was evaluated by isothermal titration calorimetry (ITC) in H_2O containing 10^{-2} M β-CD. The data obtained from the titration were successfully fitted to a 1 : 1 binding model giving an association constant $K_a = (7.5 \pm 1.2) \times 10^5$ M^{-1} for the **26**·**28** complex. The tetravalent interaction of the four adamantyl groups of ligand **26** with the cyclodextrin cavities of β-CD SAM ensures a very strong surface binding. SPR experiments showed that the ligand could not be removed from the surface even upon rinsing with a concentrated β-CD solution. Estimation of the binding constant using existing models of polyvalent binding yielded an association constant in the order of 10^{15} M^{-1}. Once the ligand is bound onto the surface of the β-CD SAM, it exposes the four guanidinium groups upward, where they can eventually interact with the complementary anionic groups. By addition of the tetrasulfonate derivative **28** to the surface covered by **26**, a supramolecular capsule was formed with an association constant $K_a = (3.5 \pm 1.6) \times 10^6$ M^{-1}, very similar to that found in solution by ITC. SPR experiments indicated that the capsule can be disassembled upon rinsing with a concentrated salt solution (1 M KCl) that weakens the guanidinium–sulfonate interactions. In this experiment the cationic part of the capsule remains bound to the β-CD SAM but it can be detached by rinsing the surface with 2-propanol, which reduces the hydrophobic interaction between the adamantyl groups and the β-CD cavity. In this way it was demonstrated that starting from a bare β-CD SAM it is possible to assemble and disassemble the supramolecular capsule in its molecular components in a stepwise, controlled manner, a concept which could be very useful in nanofabrication.

Summary and outlook

For many years calix[*n*]arenes (n = 4–8), like cyclodextrins and crown ethers, have played an important role in Supramolecular Chemistry, mainly as hosts for ions and neutral molecules. More recently, they have been used as platforms for the synthesis of multivalent ligands, thanks to the fact that they possess from 4 to 8 reactive positions at the upper rim (aromatic nuclei) and the same number of reactive OH groups at the lower rim, which can be all functionalised with identical or different binding units.

This circumstance has given a biological perspective to the research in the field of calixarenes and the examples described in this review illustrate the potential of this approach in the rapidly growing field of Bionanotechnology (Fig. 12), whose aim is to develop new tools for biology, new biomaterials, selective sensors and supramolecular

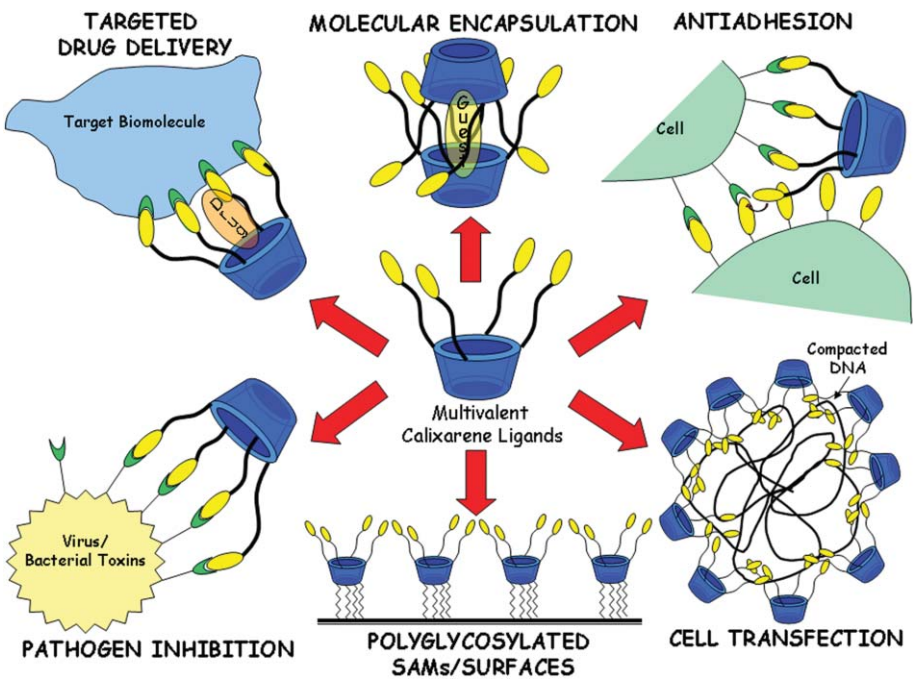


Fig. 12 Multivalent calixarene ligands in Bionanotechnology.

devices for clinical analysis, new therapeutics and smart drug delivery systems.

Acknowledgements

This work was partially supported by the Ministero dell'Istruzione, Università, Ricerca (MIUR), FIRB project RBNE019H9K: Manipolazione Molecolare per Macchine Nanometriche and COST Action D34 (WG D34/0001/05: Multivalent ligands for cholera toxin inhibition and sensing). We thank CIM (Centro Interdipartimentale Misure) of the University of Parma for NMR and AFM facilities.

References

1 M. Mammen, S.-K. Choi and G. M. Whitesides, *Angew. Chem., Int. Ed.*, 1998, **37**, 2755–2794.
2 A. Varki, *Glycobiology*, 1993, **3**, 97–130.
3 A. Mulder, J. Huskens and D. N. Reinhoudt, *Org. Biomol. Chem.*, 2004, **2**, 3409–3424.
4 S.-K. Choi, in *Synthetic Multivalent Molecules*, ed. S.-K. Choi, John Wiley & Sons, Inc., Hoboken, New Jersey, 2004.
5 G. Ercolani, *J. Am. Chem. Soc.*, 2003, **125**, 16097–16103.
6 W. P. Jenks, *Proc. Natl. Acad. Sci. U. S. A.*, 1981, **78**, 4046–4050.
7 P. I. Kitov and D. R. Bundle, *J. Am. Chem. Soc.*, 2003, **125**, 16271–16284.
8 *Calixarenes 2001*, ed. Z. Asfari, V. Böhmer, J. Harrowfield, and J. Vicens, Kluwer Academic Publishers, Dordrecht, 2001.
9 A. Casnati, F. Sansone and R. Ungaro, *Acc. Chem. Res.*, 2003, **36**, 246–254.
10 Y. C. Lee and R. T. Lee, *Acc. Chem. Res.*, 1995, **28**, 321–327.
11 F. Sansone, E. Chierici, A. Casnati and R. Ungaro, *Org. Biomol. Chem.*, 2003, **1**, 1802–1809.
12 F. Sansone, A. Casnati and R. Ungaro, unpublished results.
13 G. M. L. Consoli, F. Cunsolo, C. Geraci and V. Sgarlata, *Org. Lett.*, 2004, **6**, 4163–4166.
14 N. Micali, V. Villari, G. M. L. Consoli, F. Cunsolo and C. Geraci, *Phys. Rev. E*, 2006, **73**, DOI: 10.1103/PhysRevE.73.051904.
15 A. Bernardi, D. Arosio and S. Sonnino, *Neurochem. Res.*, 2002, **27**, 539–545.
16 D. Arosio, M. Fontanella, L. Baldini, L. Mauri, A. Bernardi, A. Casnati, F. Sansone and R. Ungaro, *J. Am. Chem. Soc.*, 2005, **127**, 3660–3661.
17 Y. Aoyama, *Chem.–Eur. J.*, 2004, **10**, 588–593.
18 O. Hayashida, K. Mizuki, K. Akagi, A. Matsuo, T. Kanamori, T. Nakai, S. Sando and Y. Aoyama, *J. Am. Chem. Soc.*, 2003, **125**, 594–601.
19 K. Fujimoto, O. Hayashida, Y. Aoyama, C. T. Guo, K. I. P. J. Hidari and Y. Suzuki, *Chem. Lett.*, 1999, 1259–1260.
20 Monothematic issue on "Cationic Transfection Lipids": *Curr. Med. Chem.*, 2003, **10**(14).
21 M. Dudič, A. Colombo, F. Sansone, A. Casnati, G. Donofrio and R. Ungaro, *Tetrahedron*, 2004, **60**, 11613–11618.
22 F. Sansone, M. Dudič, G. Donofrio, C. Rivetti, L. Baldini, A. Casnati, S. Cellai and R. Ungaro, *J. Am. Chem. Soc.*, 2006, DOI: 10.1021/ja0634425.
23 F. Sansone, S. Barboso, A. Casnati, M. Fabbi, A. Pochini, F. Ugozzoli and R. Ungaro, *Eur. J. Org. Chem.*, 1998, 897–905.
24 M. Lazzarotto, F. Sansone, L. Baldini, A. Casnati, P. Cozzini and R. Ungaro, *Eur. J. Org. Chem.*, 2001, 595–602.
25 L. Baldini, F. Sansone, C. Massera, A. Casnati, F. Ugozzoli and R. Ungaro, *Inorg. Chim. Acta*, 2006, DOI: 10.1016/j.ica.2006.07.023.
26 F. Sansone, L. Baldini, A. Casnati, E. Chierici, G. Faimani, F. Ugozzoli and R. Ungaro, *J. Am. Chem. Soc.*, 2004, **126**, 6204–6205.
27 A. M. Rincón, P. Prados and J. de Mendoza, *Eur. J. Org. Chem.*, 2002, 640–644.
28 A. M. Rincón, P. Prados and J. de Mendoza, *J. Am. Chem. Soc.*, 2001, **123**, 3493–3498.
29 J. Rebek, *Chem. Commun.*, 2000, 637–643 and references therein.
30 A. Bogdan, Y. Rudzevich, M. O. Vysotsky and V. Böhmer, *Chem. Commun.*, 2006, 2941–2952 and references therein.
31 Calix[4]arene tetraurea derivatives form a class of multiarmed calixarene-based receptors which self-assemble in dimeric capsules where a guest can be encapsulated. Their supramolecular properties have been extensively described in recent review articles (*see* ref. 29 and 30) and we preferred not to include them in this tutorial review which mainly focuses on peptido- and glycocalixarene derivatives.
32 H. Yin and A. D. Hamilton, *Angew. Chem., Int. Ed.*, 2005, **44**, 4130–4163 and references therein.
33 J. Z. Sun, M. A. Blaskovich, R. K. Jain, F. Delarue, D. Paris, S. Brem, O. M. Wotoczek, Q. Lin, D. Coppola, K. H. Choi, M. Mullan, A. D. Hamilton and S. M. Sebti, *Cancer Res.*, 2004, **64**, 3586–3592.
34 J. Z. Sun, D. A. Wang, R. K. Jain, A. Carie, S. Paquette, E. Ennis, M. A. Blaskovich, L. Baldini, D. Coppola, A. D. Hamilton and S. M. Sebti, *Oncogene*, 2005, **24**, 4701–4709.
35 T. Mecca, G. M. L. Consoli, C. Geraci and F. Cunsolo, *Bioorg. Med. Chem.*, 2004, **12**, 5057–5062.
36 S. Francese, A. Cozzolino, L. Caputo, C. Esposito, M. Martino, C. Gaeta, F. Troisi and P. Neri, *Tetrahedron Lett.*, 2005, **46**, 1611–1615.
37 A. Mulder, T. Auletta, A. Sartori, S. Del Ciotto, A. Casnati, R. Ungaro, J. Huskens and D. N. Reinhoudt, *J. Am. Chem. Soc.*, 2004, **126**, 6627–6636.
38 F. Corbellini, A. Mulder, A. Sartori, M. J. W. Ludden, A. Casnati, R. Ungaro, J. Huskens, M. Crego-Calama and D. N. Reinhoudt, *J. Am. Chem. Soc.*, 2004, **126**, 17050–17058.
39 T. Auletta, B. Dordi, A. Mulder, A. Sartori, S. Onclin, C. M. Bruinink, M. Peter, C. A. Nijhuis, H. Beijleveld, H. Schonherr, G. J. Vancso, A. Casnati, R. Ungaro, B. J. Ravoo, J. Huskens and D. N. Reinhoudt, *Angew. Chem., Int. Ed.*, 2004, **43**, 369–373.
40 O. Crespo-Biel, B. J. Ravoo, J. Huskens and D. N. Reinhoudt, *Dalton Trans.*, 2006, 2737–2741.
41 J. Huskens, A. Mulder, T. Auletta, C. A. Nijhuis, M. J. W. Ludden and D. N. Reinhoudt, *J. Am. Chem. Soc.*, 2004, **126**, 6784–6797.

