

Review

Selective Recognition of Amino Acids and Peptides by Small Supramolecular Receptors

Joana N. Martins , João Carlos Lima  and Nuno Basílio * 

LAQV—REQUIMTE, Departamento de Química, Faculdade de Ciências e Tecnologia, Universidade NOVA de Lisboa, 2829-516 Caparica, Portugal; jmn.martins@campus.fct.unl.pt (J.N.M.); lima@fct.unl.pt (J.C.L.)

* Correspondence: nuno.basilio@fct.unl.pt; Tel.: +351-212-948-355

Abstract: To this day, the recognition and high affinity binding of biomolecules in water by synthetic receptors remains challenging, while the necessity for systems for their sensing, transport and modulation persists. This problematic is prevalent for the recognition of peptides, which not only have key roles in many biochemical pathways, as well as having pharmacological and biotechnological applications, but also frequently serve as models for the study of proteins. Taking inspiration in nature and on the interactions that occur between several receptors and peptide sequences, many researchers have developed and applied a variety of different synthetic receptors, as is the case of macrocyclic compounds, molecular imprinted polymers, organometallic cages, among others, to bind amino acids, small peptides and proteins. In this critical review, we present and discuss selected examples of synthetic receptors for amino acids and peptides, with a greater focus on supramolecular receptors, which show great promise for the selective recognition of these biomolecules in physiological conditions. We decided to focus preferentially on small synthetic receptors (leaving out of this review high molecular weight polymeric systems) for which more detailed and accurate molecular level information regarding the main structural and thermodynamic features of the receptor biomolecule assemblies is available.



Citation: Martins, J.N.; Lima, J.C.; Basílio, N. Selective Recognition of Amino Acids and Peptides by Small Supramolecular Receptors. *Molecules* **2021**, *26*, 106. <https://doi.org/10.3390/molecules26010106>

Academic Editor: Roland Pieters
Received: 17 November 2020
Accepted: 22 December 2020
Published: 28 December 2020

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: amino acids; peptides; host-guest systems; molecular recognition; supramolecular receptors

1. Introduction

Peptides are fundamental molecules with several biological functions, acting as neurotransmitters, neuromodulators and hormones in numerous biochemical processes, such as quorum sensing, immune response, pain and metabolism, to name a few [1,2]. Given their utmost importance in different biochemical contexts, the discovery of high affinity and selective synthetic receptors for peptide recognition holds a strong potential for enabling new therapeutic agents, such as inhibitors/activity modulators or vehicles for drug delivery, and important molecular components for advanced biological and diagnostic tools [3–7]. Furthermore, the discovery of the fundamental binding interactions controlling the recognition of small peptide sequences is expected to pave the way towards the rational design of synthetic receptors for surface protein recognition. [3,6,8–10].

However, selective and high-affinity receptors for peptides are recognized to be very difficult to design and synthesize in a rational, bottom-up manner. Being biological molecules, these targets present a difficulty just by the fact that their recognition should be made in water, imposing practical difficulties related to the solubility of synthetic organic receptors, which may require further functionalization to improve their solubility in aqueous solution. Moreover, the solvation of these biomolecules by water and the need to remove the solvent molecules and destabilize the solvation sphere also imposes an energetic penalty that must be overcome upon binding. Directional and strong interactions based on hydrogen bonding, which are frequently and successfully employed in organic solvents,

are completely or partially neutralized in aqueous solutions. Therefore, it is necessary to explore other mechanisms, such as the hydrophobic effect and multiple interactions acting cooperatively to achieve the high affinity required to bind the peptide targets at relevant μM concentration or below [11]. In addition, the flexible, rather ill-defined structure of peptide sequences also increases the difficulties associated with the rational design of small receptors [6,12–14]. While the challenges associated to the factors enumerated above may be common to other biological targets, peptide recognition, in its completeness, requires receptors that are sequence-selective, discriminating not only the type of amino acids residues composing the peptide but also the order by which they are arranged.

Although general methods for high affinity peptide recognition by synthetic receptors are still missing, several research groups have reported encouraging results that show that it is possible to bind specific peptide sequences with simple receptors, such as macrocyclic cavitands, molecular clips, and tweezers or self-assembled coordination cages. Other strategies based on templating methods, such as molecular imprinted polymers and micelles, also showed very promising results. In this review we will cover selected examples of cavitand-based receptors that have been shown to recognize amino acids and peptides in aqueous solution and use these examples to emphasize the noncovalent interactions and geometric factors that control binding stability and selectivity of the different complexes. For this reason, we focus on small molecule receptors for which this type of information can be extrapolated in a more straightforward way and potentially used for the rational optimization of current receptors and design of new ones. Then we will also provide some selected examples of biological and technological applications based on the binding properties of the receptors discussed here to illustrate the potential of these molecules and approaches.

2. The Biological Targets: Amino Acids

Amino acids are the basic building blocks of small peptides and proteins. Even though they are important targets by themselves, the development of supramolecular receptors that can bind specific amino acids with high affinity and selectivity constitutes a first step towards the development of multivalent, sequence specific receptors for peptides. However, the selective binding of free amino acids already presents a difficult task since some side chains of amino acids are very similar to each other. In this section we have divided the 20 common amino acids into four categories: basic, aromatic, acid, and neutral non-aromatic amino acids and discussed their properties. In the next section, their recognition by selected supramolecular receptors will be further discussed.

2.1. Basic Amino Acids

L-Lysine (Lys) and L-Arginine (Arg) are amino acids that play very important roles in biological systems—these are sites of recognition by the enzyme trypsin [15] and also for methylation in several proteins, as is the case of histones, the proteins involved in the structure of chromatin in the nucleus [16]; both take part in protein–protein binding [17], among other functions [18,19]. Arginine has a very important role in the membrane penetrating properties of antimicrobial peptides, as well [20].

The positive charges on these two amino acids make them easier targets for the development of receptors in water, many of them relying on the formation of several ionic interactions to obtain the receptor-analyte complexes. This is the case of *p*-sulfonatocalix[n] arene [21–23], some examples of molecular tweezers [24,25], carboxylatopillar [5]arene [26], all of which have negative charges, to form ionic bonds with the charged amines of Lys and Arg.

L-Histidine (His), despite the presence of two nitrogen atoms, is near equilibrium between its deprotonated and protonated form, at physiological pH, due to its pK_a of 6.5 [27]. Due to this, it can act as a proton donor or acceptor, depending on the environment near the residue [27]. Despite this property, histidine is usually paired with the other two

basic amino acids. Regardless, it presents a lower charge density than the other two and thus His is difficult to target selectively by ionic interactions, in favor of Lys and Arg [26,28].

2.2. Aromatic Amino Acids

Aromatic amino acids—L-tryptophan (Trp), L-tyrosine (Tyr) and L-phenylalanine (Phe)—are essential for protein function having important roles in protein–protein interaction [29] and electron transport in peptides [30] and proteins [31]. Moreover, these amino acids are precursors of several neurotransmitters—dopamine, serotonin, epinephrine, etc. [32]. Besides these functions, tryptophan also has a role in the ability of antimicrobial peptides to form pores on membranes, through the formation of cation–π bonds with arginine [33].

In terms of protein detection and analytical studies, aromatic amino acids are important because they give proteins the characteristic 280 nm absorption band and are the only amino acids that present fluorescence, specially tryptophan, with a higher quantum yield than the other aromatic amino acids [34]. Many strategies for the recognition of these amino acids are based on hydrophobic, π–π and ion dipole interactions, making use not only of the aromatic group's characteristic properties but also of the bigger size of these amino acids' side chain [35–45].

2.3. Acidic Amino Acids

The two anionic amino acids L-aspartate (Asp) and L-glutamate (Glu) exert several roles beyond their function as protein monomers. Glutamate is well known for its role as excitatory neurotransmitter in vertebrates, as well as precursor for gamma-Aminobutyric acid, GABA, an inhibitory neurotransmitter [46]. This amino acid is also the precursor to several other biomolecules and takes part in the process of elimination of ammonium, in the urea cycle [47]. Like glutamate, aspartate is also an important precursor in several biosynthetic pathways and takes part in the urea cycle [48]. However, this anionic amino acid is also essential in energetical processes in eukaryotic cells, being part of the malate–aspartate shuttle [49] and also having a role in the gluconeogenesis pathway [50].

Unlike cationic and aromatic amino acids, none of the most used macrocycles or supramolecular systems show particular affinity towards these two amino acids, without the need for further functionalization. Nevertheless, the functionalization of macrocyclic compounds with cationic groups has been demonstrated to yield potential receptors for these amino acids [51–53].

2.4. Neutral Non-Aromatic Amino Acids

Neutral non-aromatic amino acids are difficult to target selectively in aqueous solution due to their high solvation, similar structures, reduced hydrophobic surface and the lack overall positive or negative charge [3,54]. However, these amino acids still have key roles in proteins and peptides, providing sites for glycosylation [55,56] and other post translational modifications [7]. They are also fundamental in ammonia regulation [57], anabolic and catabolic processes in eukaryotic cells (acting as precursors to other macronutrients or sources of energy) [50] and redox homeostasis [58].

Despite the difficulty of targeting these molecules, some receptors were developed for several of these amino acids. Many of these, however, only reach mM affinities and many were not tested in physiological conditions due to low solubilities in water [54]. Furthermore, many receptors that bind to these biomolecules with higher affinity present very low selectivity and are frequently able to bind to other amino acids with similar or even higher affinities [35,41,43,59].

Thiols are particularly reactive and so, the recognition and detection of L-Cysteine is more commonly based on the creation of sensors that react with this redox biomolecule by nucleophilic substitution [60]. In terms of supramolecular recognition most of the existing systems are based on metal complexes [3].

3. Receptors for Amino Acids

Supramolecular chemistry has had a fruitful contribution to the development of new, multifunctional organic molecules. In terms of molecular recognition, a variety of different types of receptors have been developed, from polymeric complex structures to both acyclic and cyclic small molecules [61]. Macroyclic and semi-rigid receptors have been gaining importance due to their simplicity and, at the same time, high affinity and selectivity towards a myriad of guests, given not only by the complementary noncovalent interactions that can be formed, but also by the size complementarity between host and guest [61,62].

3.1. Macrocycles

Macrocyclic receptors are a class of molecules that have a well-defined structure for the binding of a target guest. The cavity in these is often more hydrophobic, while the portals usually present relatively polar or charged moieties [62]. In this section, a select group of examples of macrocyclic receptors, which have shown to bind to amino acids, will be presented.

3.1.1. *p*-Sulfonatocalix[n]arenes

p-Sulfonatocalix[n]arene (SC_n) macrocycles (Figure 1) were initially synthesized by Shinkai and co-workers to improve the solubility of calixarenes in water and allow for the binding and detection of not only organic compounds, but also inorganic ions in aqueous conditions [21,63–67]. Owing to their high water solubility (>0.1 M) and low toxicity, SC_n macrocycles have been widely investigated for their pharmaceutical and biological applications [21,68–74].

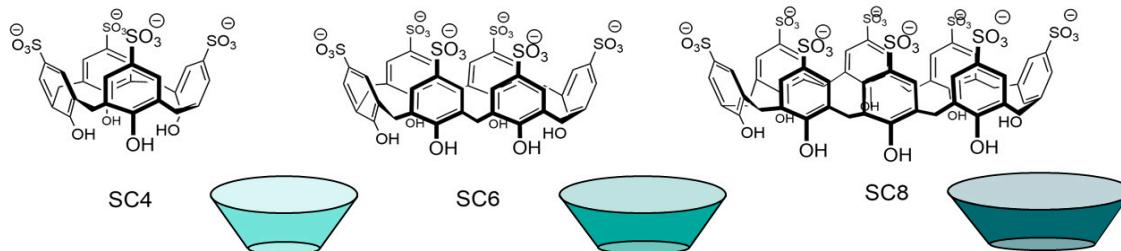


Figure 1. Structures of the *p*-sulfonatocalix[n]arene receptors, SC_n, and corresponding schematizations of the macrocycles.

In early studies by Douteau-Guével et al. and Selkti et al. the complexation of *p*-sulfonatocalix[n]arene ($n = 4, 6$ and 8) with Lys and Arg was explored by NMR spectroscopy and microcalorimetry in solution and by X-ray diffraction in the solid state [21,75,76]. The ^1H NMR studies carried out at different pH values showed that the calixarenes do not interact with these amino acids at pH 13 due to the deprotonation of the amino and guanidino groups, this last one being only partially deprotonated under these conditions ($\text{p}K_{\text{a}} = 13.2$) [75]. At more acidic pH conditions (pH 5 and pH 1), the authors also pointed out the weaker interactions between these amino acids and the larger SC6 and SC8 hosts, attributing the low affinity to the more flexible structure of these hosts. It must be noted, however, that the larger SC8 can form 1:1 and 1:2 host:guest complexes with Arg and Lys. SC4 was investigated in more detail at pH 1 and 5, showing that the receptor displays higher affinity for both amino acids at pH 5 and that, independently of the pH, Arg ($K = 1.7 \times 10^3 \text{ M}^{-1}$ at pH 5, $K = 2.0 \times 10^2 \text{ M}^{-1}$ at pH 1) binds more strongly than Lys ($K = 6.0 \times 10^2 \text{ M}^{-1}$ at pH 5, $K = 1.0 \times 10^2 \text{ M}^{-1}$ at pH 1). It was also noted that the presence of high concentrations of metal cations inhibits the formation of the complexes, a phenomenon which was generalized to other guest molecules in later studies and assigned to the competitive binding of the cations (including hydronium) with SC_n [77–85]. The complexation induced chemical shifts suggests that the terminal amino/guanidino groups in the side chain are included in the cavity of SC4 while the α -amino group remains outside of the cavity, most probably due to repulsion arising from the carboxylate group.

The microcalorimetry studies showed that the binding of both amino acids with SC4 is enthalpy driven ($\Delta H_{\text{Lys}} = -14.4 \text{ kJ/mol}$ and $\Delta H_{\text{Arg}} = -20.3 \text{ kJ/mol}$), with only slight differences in entropy ($T\Delta S_{\text{Lys}} = 2.0 \text{ kJ/mol}$ and $T\Delta S_{\text{Arg}} = -2.1 \text{ kJ/mol}$), yet with two times higher affinity of the receptor towards Arg than Lys (Table 1), being in line with the previous NMR data [21,75]. This study also confirmed that SC6 and SC8 are in fact weaker binders for these two targets but, contrary to what was reported based on NMR titrations, also suggests that the binding affinity of SC4 increases in more acidic conditions.

The thermodynamic information given by the microcalorimetry studies indicates that the more favorable enthalpy of the Arg complex is probably due to the formation of cation– π and π – π interactions between its guanidinium group and the aromatic rings in the cavity of the calixarene (Figure 2). Moreover, this further stabilization can justify the negative value of entropy, seeing as arginine binds deeper in the cavity of the receptor [71] and so, with reduced degrees of freedom [21]. The affinity of this receptor towards L-Histidine (His) was also tested, however, the affinity obtained was two times lower than towards Lysine and near the affinity towards neutral amino acids [28]. In addition to the studies made in solution, the crystallographic structures of the Lys (Figure 3) [76] and Arg [71] complexes further support the recognition of these amino acids through the inclusion of their side chains in the cavity of the SC4 receptor.

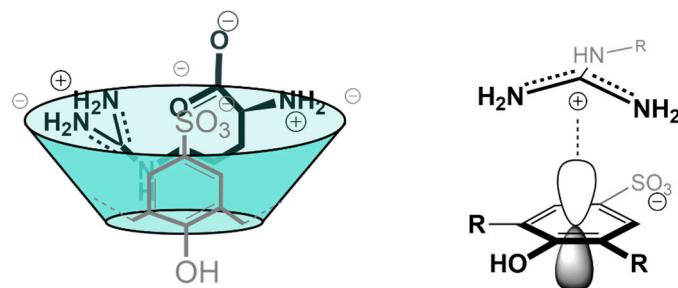


Figure 2. Structure of the complex formed between SC4 and L-Arginine, adapted from [21,71] (left) and a schematization of the cation– π interactions that are established in the former, between the aromatic rings of the receptor and the guanidinium moiety of the analyte.

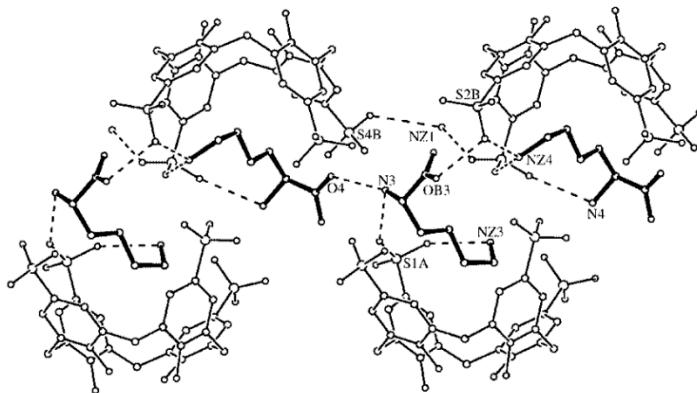


Figure 3. Crystallographic structure of the SC4-L-Lysine: the analyte (in bold) complexes in a parallel position to the calixarenes, as it can be observed by the two top complexes; another possible conformation is perpendicular to the disposition of the calixarenes, interacting not only with the calixarene at the bottom but having its ϵ -amino group, NZ1, interacting externally with the sulfonate groups of the top layer receptors. Reprinted with permission from [76]. Copyright (2000) The Royal Society of Chemistry.

Several more recent studies also measured the affinity of these complexes, using indicator displacement assays, in different conditions, which are summarized in Table 1. Many of these yielded higher affinities of the cationic amino acids towards the receptor

SC4, due to this smaller macrocycle having a more rigid structure in comparison to SC6 and SC8. These studies also indicate that, near physiological conditions, the binding may vary, considering that the presence of other cationic molecules, such as buffer counterions, can interfere extensively with the binding [21,75,86–88].

The higher affinity of SC4 towards methylated lysine is also worth noting, an important post-translational modification. Hof and co-workers observed that the affinity of the 1:1 complexes formed by SC4 with lysine and its methylated derivatives increases with the degree of methylation reaching ca. 70-fold selectivity to trimethyllysine over lysine [89]. Based on NMR titration experiments, the authors observed that while the lysine side chain is accommodated to the SC4 pocket in a “side-on” binding mode, in trimethyllysine the NMe_3^+ group is deeply included in the cavity suggesting an optimal charge and shape complementary between this group and the calixarene receptor. In a following example, the same research group explored the monofunctionalization of trisulfonated calix[4]arenes (Figure 4) to tune their affinity and selectivity towards trimethyllysine. Although most of the different receptors showed lower binding affinities, SC4-Ar, the one with the unsubstituted aromatic panel ($Z = \text{H}$) directly connected to the calixarene rim, was found to display the higher binding affinity towards trimethyllysine, almost 2-fold larger than the one previously observed for SC4 and an improved selectivity of 150-fold over unmethylated lysine (Table 1). This is presumably due to its higher hydrophobic character and additional $\text{CH}-\pi$ interactions between the lysine side chain and the extra aromatic wall [90].

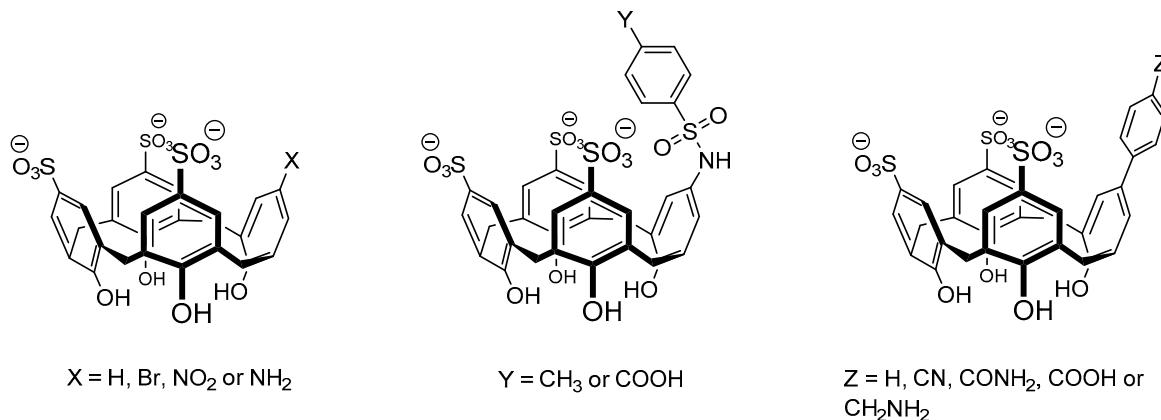


Figure 4. Structures of the upper rim monofunctionalized sulfonatocalix[4]arene receptors synthesized by Hof, where the structure with an added aromatic ring with $Z = \text{H}$, corresponds to SC4-Ar [90].

3.1.2. Pillar[n]arenes

Pillararenes are a relatively new class of macrocycles, that only recently have begun to be applied for the binding of analytes in water [91–93]. Carboxylatopillar [5]arene (CP5) was one of the first water-soluble pillararenes to be synthesized [91] and recently it has been applied as a receptor for cationic organic molecules. Similar to SC n , these molecules present a rigid electron rich cavity lined with anionic groups on both sides of the cavity (Figure 5) [26,94].

In line with its ability to bind positively charged guest molecules, CP5 was demonstrated to selectively bind basic amino acids, i.e., L-lysine, L-arginine and L-histidine, with mM affinity (Table 1). In contrast with SC4, the biomolecules penetrate completely into the cavity, with the amine group of the chiral center being stabilized by ionic interactions at one of the entrances, and the amine in the side group of the amino acid being stabilized by the opposite side. Although these interactions have a bigger role in stabilizing these complexes—as suggested by the selectivity towards cationic amino acids observed in the study—the hydrophobic effect also takes an important role, namely between the CP5 cavity and the aliphatic chains of the amino acids. Furthermore, CP5 has an even higher

affinity towards arginine over the other two cationic amino acids, due to the presence of the guanidinium moiety. This group allows for cation–π interactions with the π orbitals of the cavity of the pillararene, as well as the possibility to form several H-bonds with the receptor due to the conjugation between the two nitrogen atoms [26,94].

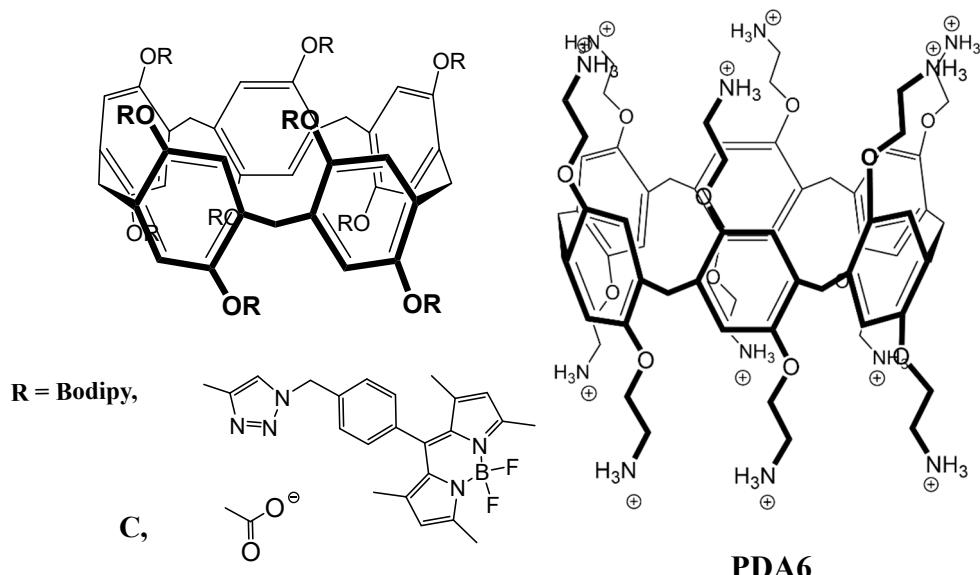


Figure 5. Structure of the Pillar[5]arene receptors, P5-Bodipy [95], carboxylatopillar[5]arene, C-P5 [26], and dodecaamine pillar [6]arene, PDA6 [51].

Another example of a pillararene receptor, is dodecaamine pillar [6]arene, PDA6, which has 6 amine groups at each of its entrances (Figure 5) and has been shown to bind to acidic amino acids [51]. The principles behind the affinity of PDA6 towards these two amino acids is the same as for CP5—the amine groups are positively charged at physiological pH and they can establish ionic interactions with the two carboxylic groups of both Glu and Asp. Although this is the main driving force of the binding (with affinity in the μM range) the complex can be further stabilized by cation–π interactions between the cavity of the pillararene and the amine group of the amino acids. This not only adds attractive interaction between the molecules, but also counterbalance the electrostatic repulsion that could be established between the cationic regions of both receptor and analyte [51].

Pillararenes can be applied for the recognition of neutral amino acids as well. This has been explored by Guler and co-workers [95], who developed a pillar[5]arene-based sensor, P5-Bodipy, which presented high selectivity towards L-Asparagine. L-Asn is an analogue of L-Aspartate, with an amide group in its side chain instead of a carboxylate group. Due to its zwitterionic character and neutral side chain, the influence of electrostatic interactions between receptor and analyte on the receptor's selectivity is negligible, being mostly defined by the hydrogen bonds, hydrophobic and Van der Waals interactions [54,95]. P5-Bodipy is composed by a pillar[5]arene macrocycle, decorated with BODIPY moieties (Figure 5). This receptor detected L-Asn selectively by fluorescence spectroscopy, showing an increase in the Bodipy fluorescence. ¹H-NMR elucidated the types of interactions present, showing that L-Asn is inserted in the pillararene cavity, being stabilized by several types of electrostatic interactions, but with Van der Waals interactions and hydrogen bonds being the most relevant for the selectivity of the receptor.

3.1.3. Cucurbit[n]urils

Cucurbit[n]urils (CB[n]) are macrocycles composed of n glycoluril units, that, owing to their outstanding binding properties, are largely applied for the recognition of several organic molecules and biological analytes (Figure 6) [44,96–100]. Thanks to their hydrophobic

cavity and highly electronegative carbonyl portals, these macrocyclic containers are particularly suitable to bind positively charged and hydrophobic guests. Buschmann et al. [35] performed earlier studies on the complexation of amino acids with CB6 which showed mM affinity towards the Phe, Gly, Ala, and Val, in aqueous formic acid 50%(*v/v*). Nau and co-workers reported the complexation of lysine, arginine and histidine by CB7 with all three amino acids showing moderate affinity in the mM range [45]. Later on, Macartney et al. investigated the recognition of lysine, arginine and their methylated derivatives by CB6 and CB7 [101]. Noteworthy, the affinity increases significantly upon lysine methylation, reaching a 3500-fold selectivity for trimethyllysine over lysine as a result of the good complementarity of CB7 and trimethylalkylammonium motifs. On the other hand, at neutral pH, lysine seems to bind only in the CB6 portal, forming an exclusion complex that, nevertheless, is more stable than the one formed with the larger CB7 ($1.1 \times 10^4 \text{ M}^{-1}$ vs. $5.2 \times 10^2 \text{ M}^{-1}$). A subsequent systematic study on the formation of host–guest complexes between CB7 and different amino acids revealed that their stability increases under acidic conditions [102]. The differences in association constants obtained at different pH values are particularly relevant for lysine and arginine which show ca. 3 order of magnitude increases from $\approx 10^2 \text{ M}^{-1}$ at neutral pH to $\approx 10^5 \text{ M}^{-1}$ at pH 2. This observation puts in evidence the repulsive interactions established between the carbonyl portals of the receptor and carboxylate group of the amino acid. This destabilizing factor is more relevant in these two cases, most probably due to the fact that their side-chain must penetrate deeply into the CB7 cavity for optimal binding interactions, which are counteracted by the negatively charged carboxylate groups at neutral or slightly acid conditions [102].

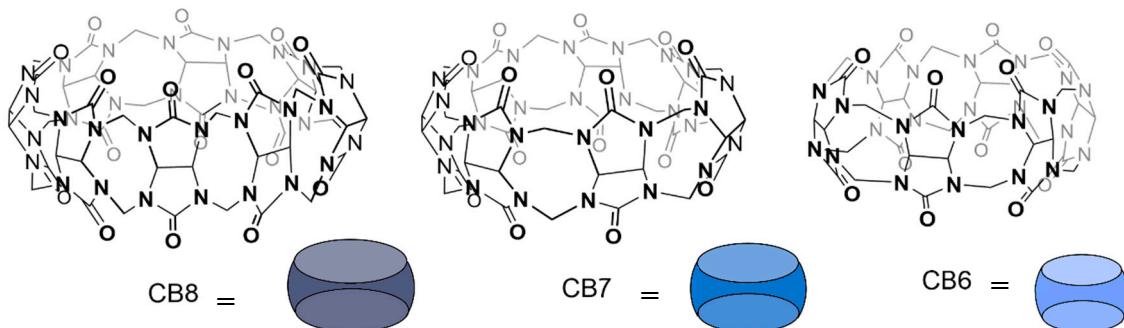


Figure 6. Structure of the Cucurbit[n]uril (CBn) receptors and corresponding schemes that will be used for illustration purposes throughout this review.

Despite being investigated for their affinity towards basic amino acids, CB7, CB8 and other CB[n] derivatives were found to be more successful at targeting aromatic amino acids in biologically relevant conditions [44,102]. Among the different CB[n] homologues, CB6 was shown to display weak affinity towards aromatic amino acids probably due to the small size of its cavity (Figure 7e), while the ones with more than 6 glycoluril units have a large enough cavity to accommodate the aromatic side chain in its interior [36–39,44,45]. This is the case of both CB7 and CB8 that were reported to, respectively, form 1:1 complexes and 1:1/1:2 complexes with aromatic amino acids (Figure 7a,c,d) [36–39,45].

CB7, for example, was shown to be selective for aromatic amino acids at pH 7 (see Table 1) displaying higher affinity for L-phenylalanine with binding constants in the order of 10^5 M^{-1} – 10^6 M^{-1} depending on the medium conditions, in particular on the presence of salts [102–104]. Amongst the different aromatic amino acids, CB7 is selective for Phe by a factor of approximately 10 over Tyr and by a factor of 100 over Trp [102]. Most thermodynamic studies show that the complexation process is enthalpy driven in great part due to the displacement of high energy water molecules from the cavity of the host [102,105–108]. This hydrophobic effect can be complemented by attractive ion-dipole interactions between the protonated amino group of the biomolecules and the carbonyl group of the cucurbituril (Figure 7) [35,44,109]. In the case of the binding of aromatic amino acids by CB7, the process

is also enthalpy driven with unfavorable negative entropy contributions. The enthalpic gain is only slightly higher for phenylalanine with respect to tyrosine and tryptophan, but the entropic loss is very low in first case ($T\Delta S = -0.6 \text{ kJ}\cdot\text{mol}^{-1}$) becoming more important for tyrosine ($T\Delta S = -3.7 \text{ kJ}\cdot\text{mol}^{-1}$) and tryptophan ($T\Delta S = -11.3 \text{ kJ}\cdot\text{mol}^{-1}$) [102]. On the basis of these thermodynamic parameters, the observed selectivity trend for recognition of aromatic amino acids by CB7 was explained by taking into account the higher hydrophobic character of phenylalanine comparatively with tyrosine and the restricted motion of the indole side chain of tryptophan inside the CB7 cavity [102].

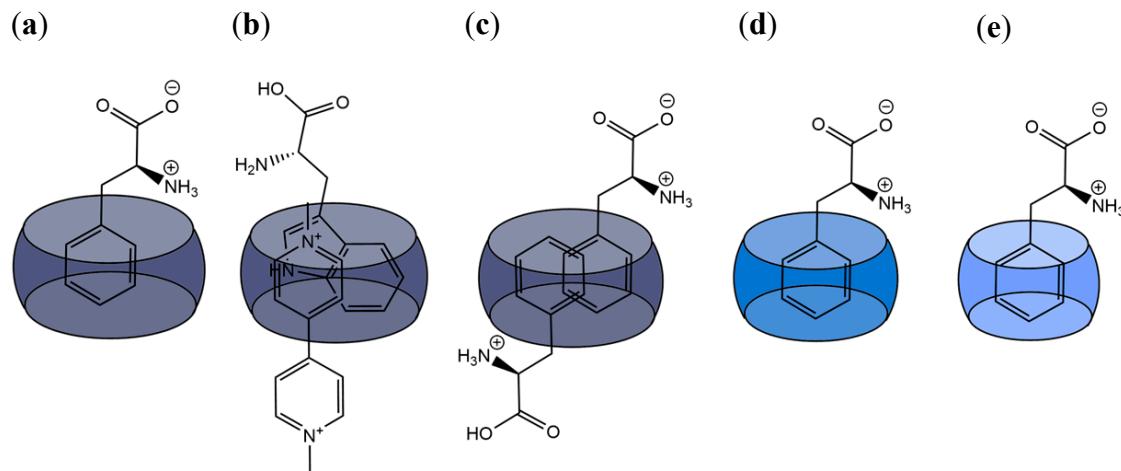


Figure 7. Examples of the CB[n] complexes: (a) CB8's complex with one Phe molecule; (b) CB8's heteroternary complex with Trp and Methyl Viologen; (c) CB8's homoternary complex with two Phe molecules; (d) CB7 complex with Phe; (e) CB6 complex with Phe.

CB8, on the other hand, binds aromatic amino acids with 1:1 (Figure 7a) and 1:2 (Figure 7c) stoichiometry, as mentioned above [110]. The ITC binding studies performed on these systems did not allow the determination of the stepwise binding constants and thermodynamic parameters for the 1:1 and 1:2 complexes in separate which precludes a more in depth thermodynamic analysis. Nevertheless, based on the obtained overall binding constants (i.e., $K_{1:1}K_{1:2}$) one can conclude that CB8 is also more selective for phenylalanine, but in contrast with CB7, tryptophan is complexed with only slight lower affinity than phenylalanine while the affinity of tyrosine seems to be too low to be measured [110].

The larger cavity of CB8 allows the binding of amino acids in the presence of auxiliary guests based on electron deficient organic molecules, such as methyl viologen (MV) [37], 2,7-dimethyl- diazaphenanthrenium (DPT) [38] and tetramethyl benzobis(imidazolium) (MBBI) [39]. The preformed 1:1 CB8:auxiliary guest complex can be viewed as a new receptor capable of forming 1:1:1 heteroternary complexes with electron rich molecules that, in some cases, does not bind significantly to CB8 alone (Figure 7b) [37–39]. Interestingly, the CB8:MV complex was shown to form 1:1:1 complexes only with the aromatic amino acids (except histidine). In the 1:2 complex, phenylalanine establishes strong contacts with the second Phe residue inside the cavity, as well as with CB8 itself [110]. Differently from what is observed for CB8 in the absence of auxiliary guest, the CB8:MV supramolecular receptor shows higher affinity for tryptophan ($4.3 \times 10^4 \text{ M}^{-1}$) than for phenylalanine and tyrosine [110]. This selectivity was attributed to the charge transfer interactions between MV and the indole side chain of this amino acid inside the CB8 cavity. The nature of the auxiliary guest may also influence the recognition of the second guest. CB8:MBBI displays binding affinities for tryptophan ($3.4 \times 10^4 \text{ M}^{-1}$) [39] similar to the one observed for CB8:MV, while CB8:DPT forms 1:1:1 complexes with this amino acid that are one order of magnitude more stable ($4.2 \times 10^5 \text{ M}^{-1}$) [38].

Cucurbiturils can also be modified covalently to enhance their affinity towards amino acids. One example of this is the work done by Isaacs and co-workers, which incorporate

two aromatic rings within the CB6 macrocycle. This study yielded a water-soluble receptor with an elongated cavity, CB6Ar (Figure 8), which is able to form $\pi-\pi$ interactions with included analytes, while maintaining CB[n]’s capacity to form ion-dipole interactions at the macrocycle entrances. The addition of the $\pi-\pi$ interaction renders a receptor with micromolar affinity towards tryptophan ($K_a = 3.2 \times 10^6 \text{ M}^{-1}$) and ca. 100-fold selectivity for this amino acid over phenylalanine and tyrosine [40].

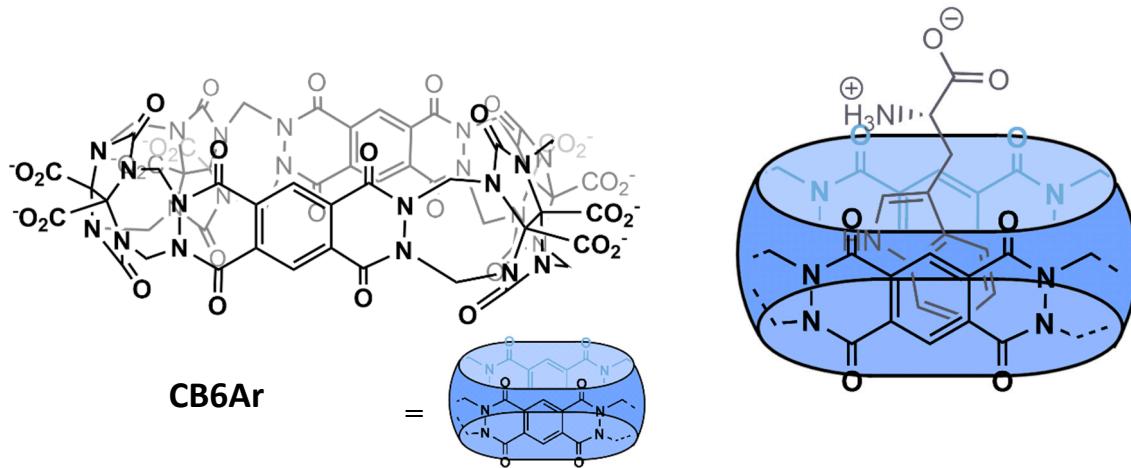


Figure 8. Structure of the modified CB6 receptor, CB6Ar, and a schematization of the complex formed between the receptor and L-Trp, with the establishment of two $\pi-\pi$ interactions [40].

3.1.4. Cyclodextrins

Cyclodextrins (CD) are macrocycles formed by glucose units that can bind a variety of different small molecules. The most common ones are those composed of 6, 7, and 8 glucose units, corresponding respectively to α -CD, β -CD, and γ -CD (Figure 9) [111]. Their host–guest complexes with biomolecules are usually low in specificity, but are slightly more stable with aromatic or small, less polar molecules

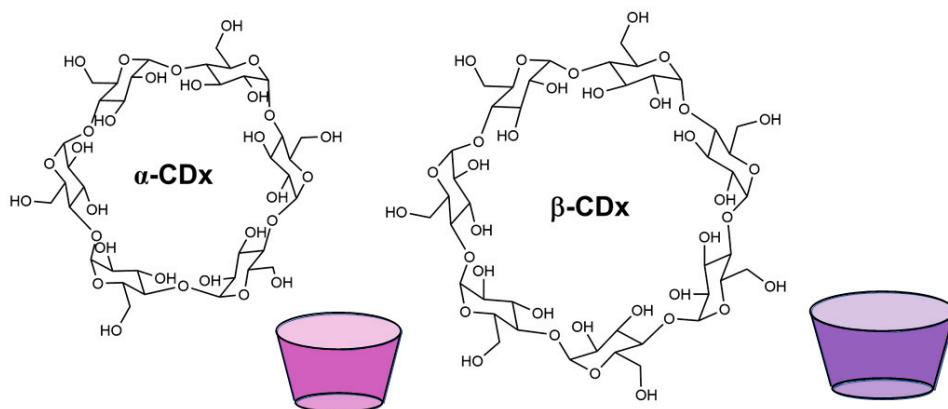


Figure 9. Structure of the receptors α -Cyclodextrin (α -CD) and β -Cyclodextrin (β -CD) and corresponds to Figure 41. Recently synthesized a single substituted analogue of this macrocycle, β -CDU, with high affinity towards L-tryptophan ($K = 5.2 \times 10^4 \text{ M}^{-1}$), thanks to the added “arm” with hydrogen bonding groups as well as a hydrophobic region (Figure 10). Despite the larger cavity of β -cyclodextrin, structural studies showed that there was no inclusion complex, contrary to what is usual for these systems, but instead the complex was formed at the entrance of the receptor. This binding was shown to happen in 2:1 host:guest stoichiometry with the receptor forming a dimer, with the analyte in the created pocket, being stabilized by hydrophobic interactions and hydrogen bonds [41].

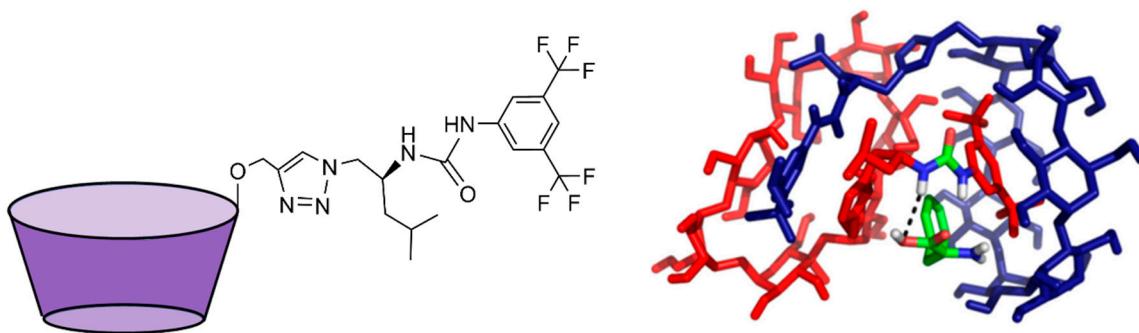


Figure 10. Structure of the receptor β -CDxU (left) and structure of the dimer that is formed upon complexation of L-Phe (right), reprinted with permission from [41]. Copyright (2017) Elsevier.

α -cyclodextrin was shown to bind L-Phe in a 2003 study by Buschmann et al., but the binding is rather unspecific, only driven by hydrophobic interactions, yielding the same affinities for this amino acid and for a series of other less polar amino acids, as well as other non-polar analytes tested [35].

3.1.5. Cavitands

Cavitands is a term that encompasses several types of rigid macrocycles and it was first coined by Cram and co-workers in 1982 [112]. These constitute a class of receptors that present a “enforced cavity” where a small guest can bind, being the most widely studied ones those based on resorcinarenes macrocycles [62,112]. The Dalcanale group has reported several cavitand derivatives (Figure 11) that were found to recognize amino acids and amino acid metabolites [113,114], and were able to obtain the crystallographic structure of the complexes between some of these cavitand receptors and all basic amino acids, Arg, Lys, and His. The 1:2 complexes were shown to be mainly stabilized by direct and water-mediated hydrogen bonds between a guest molecule and two host molecules [114]. These receptors were also shown to be particularly adequate to selectively recognize sarcosine, a potential prostate cancer biomarker, over glycine in water and in biological fluids [113]. The crystal structures of the complexes showed that while glycine methyl ether is localized at the receptor’s portal, sarcosine forms an inclusion complex with the *N*-methyl group deeply included in the cavity. This structural arrangement allows the stabilization of the complex by a combination of non-covalent interactions that include CH- π contacts between the *N*-methyl residue and the aromatic pocket, along with cation-dipole and hydrogen bonding between the ammonium and the phosphonate groups. The selective complexation of sarcosine in solution was confirmed by ITC experiments in methanol, by water:chloroform extraction and solubilization assays. The recognition of sarcosine in water and in urine was achieved by grafting the cavitands to a silicon surface. The authors showed that the receptor held its recognition properties at these interfaces and devised a sensor based on a luminescence indicator displacement assay using (9-anthrylmethyl) methyl ammonium chloride as probe.

More recently, a comprehensive study on recognition of amino acids by tetraphosphonate cavitands was performed to dissect the thermodynamic factors responsible for the observed selectivity of these receptors towards *N*-methyl amino acids [115]. The complexation was investigated both in methanol and in aqueous solutions, showing that the stability of the resulting complexes significantly decreases (almost 2 order of magnitude) in water. The ITC results demonstrated that while the association process is both enthalpy and entropy-driven in methanol, it becomes entropically unfavorable in water. The authors attributed this observation to the non-classical enthalpy driven hydrophobic effect and to the enhanced solvation of the complex in aqueous solution. Despite of the decreased affinity, the selectivity of these receptors toward *N*-methylated amino acids improves in aqueous solution resulting in their exclusive complexation.

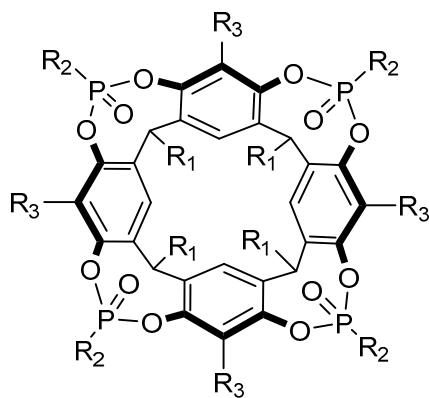


Figure 11. General structure of the Dalcanale cavitand receptors, Cav[R₁,R₂,R₃] [113–115].

3.1.6. Other Macro cyclic Receptors

A variety of different macrocycles have been used for the binding of a myriad of molecules. Some less known examples that are only now being applied in biomolecule recognition or that present more complex molecular structures are described below, showcasing the versatility and diverse design that macrocycles can present.

Early work by Kaifer et al. showed that the tetracationic cyclophane cyclobis(paraquat-p-phenylene), CPQ, commonly known as the blue-box, forms charge transfer complexes with aromatic amino acids in aqueous solution (Figure 12). This receptor is selective for tryptophan ($K = 1.0 \times 10^3 \text{ M}^{-1}$) showing ca. 2 and >10 fold lower affinity for tyrosine and phenylalanine, respectively [116]. Cao and co-workers more recently explored the affinity of two distinct tetracationic cyclophane derivatives towards anionic and aromatic biomolecules. The TeCPh, the cyclophane analogue with a narrower electron-deficient cavity described in this study (Figure 12), showed a high specificity towards tryptophan ($K = 1.2 \times 10^3 \text{ M}^{-1}$) in relation to the other amino acids, forming $\pi-\pi$ bonds with the electron-rich aromatic side chain of the analytes, and possibly being further stabilized by ionic interactions with the carboxylate group of Trp [42].

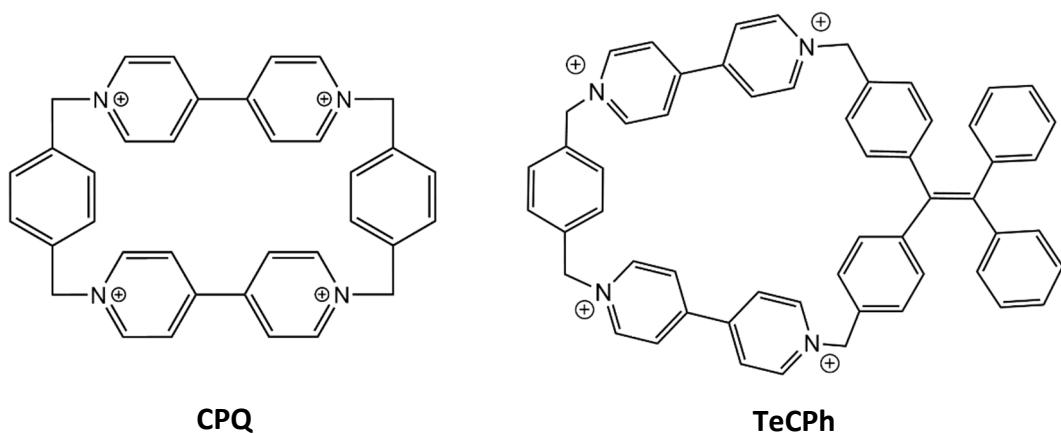


Figure 12. Structure of the cyclophane receptors CPQ [116] and TeCPh [42].

Alfonso et al. synthesized two novel oxaazamacrocycles, which presented enantioselective binding with both aspartate and glutamate—OA (R,R) e OA (S,S,S,S) [53]. Only the latter showed considerable selectivity towards the L isomers of these anionic biomolecules, but with relatively low affinities. The former showed higher affinity towards the D-isomer of aspartate. This specificity is given by the positioning of the positive groups and oxygen on the receptor in relation to the analyte that it has affinity towards (Figure 13). This distinct approach in receptor design allowed for the positioning of its charged and polar moieties in proximity to guests', in a complementary manner. The positive ammonium

groups turned towards the cavity of the receptor and the two oxygens, free to form two hydrogen bonds each, facing in the same general direction, is the ideal conformation for the binding of D-Ac-Asp to OA (R,R), allowing the establishment of ionic interactions and one hydrogen bond without the need for many spatial rearrangements of the receptor [53].

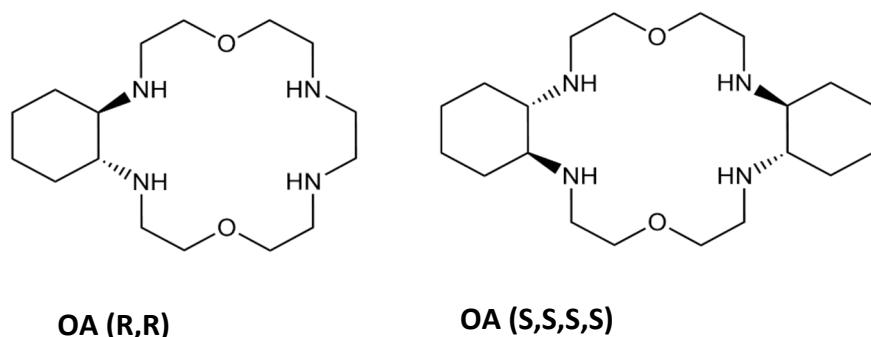


Figure 13. Structure of the oxaazamacrocycles receptors OA (R,R) and OA (S,S,S,S) [53].

Much like the previous examples, calixpyridinium, C4Pyr is a supramolecular system that has not been target of much attention in biomolecule binding applications (Figure 14). However, this macrocycle presents many similarities to calixarenes, with high solubility and with added cationic groups allowing for the creation of a binding site for small anionic biomolecules. Ren and co-workers tested this receptor for its affinity towards anionic amino acids—it showed to be highly selective for these, when compared to the rest of the proteinogenic amino acids, binding aspartate and glutamate ($K \approx 10^3 \text{ M}^{-1}$) solely by ionic interactions [52].

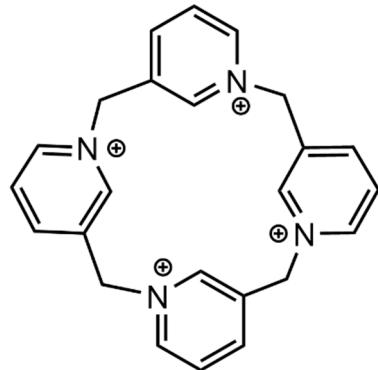
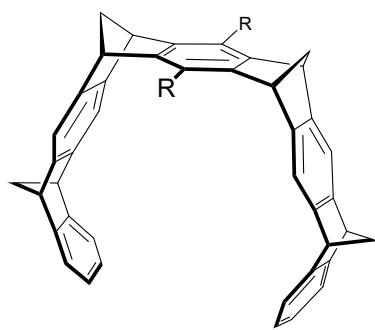


Figure 14. Structure of the receptor calixpyridinium, C4Pyr.

3.2. Open-Chain Receptors

Despite the great variety of macrocycles here presented with affinities reaching the μM range, some open-chain receptors can present rigid binding sites with affinities rivalling those of macrocycles [24]. On the other hand, less rigid structures can bring other advantages, e.g., the flexibility to bind larger guests [117].

Molecular tweezers are open-chain receptors comprising semi-rigid structures that can bind complementary analytes [24,25,118]. Two studies reported rigid molecular tweezers with the ability to bind Arg and Lys, which were designed with nine adjoining six-membered rings and a hydroquinone group at the center. The two hydroxyl groups in the hydroquinone were modified to incorporate different anionic moieties—phosphonate, phosphate, sulfate and carboxylate—creating four differently-substituted receptors, phosphonate (MPnT), phosphate (MPT), sulfate (MST) and carboxylate (MCT), respectively (Figure 15) [24,25]. These receptors showed different affinities towards Arg and Lys, as well as the acylated and O-methylated modified residues, as shown in Table 1.



MPnT: $R = OP(CH_3)O_2^-$

MPT: $R = OPO_3^{2-}$

MST: $R = OSO_3^-$

MCT: $R = OCH_2CO_2^-$

Figure 15. Structure of the Molecular Tweezers reported in Schraddet et al. functionalized with phosphonate (MPnT), phosphate (MPT), sulfate (MST), and carboxylate (MCT).

Table 1. Binding constants (K_a) for the formation of host-guest complexes between amino acids and synthetic supramolecular receptors.

Host	Guest	$K_a (M^{-1})$	Conditions	Method	Ref.
L-Arginine		1.52×10^3	Phosphate Buffer 10 mM pH 8, 20 °C	ITC	[21]
		1.7×10^3	95% water/5% deuterium oxide pH 5		[75]
		2.0×10^2	95% water/5% deuterium oxide pH 1		
		1.55×10^3	95% water/5% deuterium oxide pH 8, 20 °C	¹ H-NMR	[22]
		6.40×10^3	Deuterium oxide unbuffered, 25 °C		[88]
		1.30×10^4	Deuterium oxide unbuffered, 25 °C		[86]
		2.8×10^4	Acetate buffer 5 mM pD 4.5		
SC4		3.30×10^2	Deuterium oxide 40 mM phosphate buffer pD 7, 25 °C		[89]
		7.4×10^2	Phosphate Buffer 10 mM pH 8, 20 °C	ITC	[21]
		6.0×10^2	95% water/5% Deuterium oxide pH 5		[75]
L-Lysine		1.0×10^2	95% water/5% Deuterium oxide pH 1	¹ H-NMR	
		1.36×10^3	95% water/5% Deuterium oxide pH 8, 20 °C		[22]
		$<1 \times 10^3$	Ammonium acetate buffer 10 mM, pH 6, 25 °C	Fluorescence Spectroscopy	[87]
		4.60×10^3	Deuterium oxide unbuffered	¹ H-NMR	[86]
		3.90×10^3	Acetate buffer 5 mM pD 4.5		
		5.20×10^2			
	L-LysMe	4.00×10^3			
	L-LysMe ₂	1.60×10^4	Phosphate buffer 40 mM, pH 7.4, 30 °C	ITC	[89]
	L-LysMe ₃	3.70×10^4			

Table 1. *Cont.*

Host	Guest	K_a (M^{-1})	Conditions	Method	Ref.
SC4-Ar	L-Histidine	5.00×10^2	95% water/5% Deuterium oxide pH 8, 20 °C	$^1\text{H-NMR}$	[22]
	L-Arginine	1.90×10^2			[21]
	L-Lysine	90	Phosphate buffer pH 8, 20 °C		[90]
	L-Arginine	3.50×10^2			[52]
C4Pyr	L-Lysine	4.2×10^2	Phosphate buffer 40 mM pD 7.0, 25 °C	Fluorescence Spectroscopy	[94]
	L-LysMe ₃	6.4×10^4			[26]
P5 Bodipy	L-Lysine	1.40×10^2	Phosphate Buffer pH 8, 20 °C	Water pH 6	[95]
	L-Glutamate	1.76×10^3			[51]
CP5	L-Aspartate	1.16×10^3		Fluorescence Spectroscopy	[94]
	L-Asparagine	2.5×10^5	Dimethyl formamide/water 1:1, 25 °C		[95]
	L-Arginine	5.24×10^3	HEPES buffer pH 7.4, 25 °C		[26]
		5.90×10^3	Deuterium oxide pD 7.2, 25 °C		[52]
PDA6	L-Lysine	1.12×10^3	HEPES buffer pH 7.4, 25 °C	Fluorescence Spectroscopy	[94]
		1.80×10^3			[26]
	L-Histidine	1.50×10^3			[51]
	TriMe-Lys	1.30×10^3	Deuterium oxide pD 7.2, 25 °C		[26]
MPnT	Ac-Lys	1.90×10^2		$^1\text{H-NMR}$	[24]
	L-Glutamate	1.00×10^6	Water pH 6, 25 °C		[51]
	L-Aspartate	9.80×10^5			[24]
	L-Lysine	1.40×10^3	Deuterium oxide, Sodium Dihydrogenophosphate 25 mM pH 4.4		[24]
MPT	Ac-His-OMe	7.00×10^2		Deuterium oxide unbuffered	[24]
	Ac-Lys-OMe	4.40×10^3			[24]
		2.30×10^4			[24]
	L-Lysine	1.14×10^3			[24]
MST	L-Lysine	4.76×10^4	Phosphate buffer 200 mM pH 7.6	Phosphate buffer 10 mM pH 7.6	[25]
	Ac-Arg-OMe	1.67×10^4			[25]
	Ac-Lys-OMe	5.88×10^4			[25]
	Ac-Arg-OMe	5.00×10^4			[25]
MCT	Ac-Lys-OMe	1.11×10^5		Phosphate buffer 10 mM pH 7.2	[25]
	L-Lysine	4.41×10^3	Phosphate buffer 10 mM pH 7.6		[25]
	L-Arginine	1.43×10^3			[25]
	Ac-Arg-OMe	1.30×10^4			[25]
	Ac-Lys-OMe	5.26×10^4		Phosphate buffer 10 mM pH 7.2	[25]
	L-Lysine	8.60×10^2			[25]
	L-Arginine	1.64×10^3	Phosphate buffer 10 mM pH 7.2		[25]
	Ac-Arg-OMe	3.56×10^3			[25]
	Ac-Lys-OMe	1.56×10^3		Phosphate buffer 200 mM pH 7.6	[25]
	Ac-Lys-OMe	4.42×10^3			[25]

Table 1. *Cont.*

Host	Guest	K_a (M^{-1})	Conditions	Method	Ref.
CEAT	Lys-OMe	3.16×10^4	Methanol		
		3.98×10^4	Water with 1.6% Methanol		[118]
		7.0×10^1	HEPES buffer 50 mM pH 7.5 with 1.6% Methanol		
CB6	L-Lysine	1.10×10^4	Deuterium oxide 0.1 M NaCl, 25 °C	1H -NMR	[101]
	Glycine	4.70×10^3			
	L-Alanine	1.00×10^3	Aqueous Formic Acid 50%(v/v), 25 °C		[35]
	L-Valine	1.40×10^3			
	L-Phenylalanine	1.40×10^3		ITC	
CB7	L-Phenylalanine	1.80×10^6	Water, 25 °C		[36]
		1.80×10^5	Phosphate Buffer pH 7, 25 °C		[102]
		1.20×10^6	Phosphoric acid buffer pH 2, 25 °C		
		1.50×10^5	Water, 25 °C	1H -NMR	
		8.20×10^5	nd	UV-Visible Spectroscopy	[44]
		2.30×10^5			
		1.60×10^4	Phosphate Buffer pH 7, 25 °C	ITC	[102]
L-Tyrosine	L-Tyrosine	1.80×10^5	Phosphoric acid buffer pH 2, 25 °C		
		2.40×10^4	Ammonium Acetate Buffer pH 6, 30 °C	Fluorescence Spectroscopy	[45]
		2.20×10^4		ITC	
		3.70×10^5	nd	UV-Visible Spectroscopy	[44]
		1.20×10^3	Phosphate Buffer pH 7, 25 °C	ITC	[102]
L-Tryptophan	L-Tryptophan	7.40×10^3	Phosphoric acid buffer pH 2, 25 °C		
		1.60×10^3	Ammonium Acetate Buffer 10 mM pH 6, 25 °C	Fluorescence Spectroscopy	
		1.90×10^3	Ammonium Acetate Buffer 10 mM pH 6, 30 °C	ITC	
		8.70×10^2	Ammonium Acetate Buffer 10 mM pH 6, 25 °C	Fluorescence Spectroscopy	[45]
		8.00×10^2	Ammonium Acetate Buffer 10 mM pH 6, 30 °C		
L-Lysine	L-Lysine	2.10×10^2	Phosphate Buffer pH 7, 25 °C		
		3.10×10^5	Phosphoric acid buffer pH 2, 25 °C	ITC	[102]
		5.30×10^2			
		1.80×10^3	Sodium Acetate Buffer 50 mM, deuterium oxide pD 4.7, 25 °C	1H -NMR	[101]
		6.00×10^4			
L-Arginine	L-Arginine	1.90×10^6			
		3.10×10^2	Ammonium Acetate Buffer 10 mM pH 6, 25 °C	Fluorescence Spectroscopy	[45]
		1.40×10^5	Phosphoric acid buffer pH 2, 25 °C	ITC	[102]
		3.27×10^2	Ammonium Acetate Buffer 10 mM pH 6, 30 °C		[45]

Table 1. *Cont.*

Host	Guest	K_a (M^{-1})	Conditions	Method	Ref.
CB8	L-Histidine	4.00×10^2	Ammonium Acetate Buffer 10 mM pH 6, 25 °C	Fluorescence Spectroscopy	
		2.40×10^3	Phosphoric acid buffer pH 2, 25 °C	ITC	[102]
	L-Tyrosine ^a	$<1.00 \times 10^3$			
	L-Phenylalanine ^a	1.10×10^8	Sodium Phosphate 10 mM pH 7 CB8-AA ₂ , 27 °C		[110]
	L-Tryptophan ^a	6.90×10^7			
	L-Tyrosine	2.20×10^3		ITC	
	L-Phenylalanine	5.30×10^3	Sodium Phosphate 10 mM pH 7, co-binding with MV, 27 °C		[37]
	L-Tryptophan	4.30×10^4			
	L-Trp-OMe	6.30×10^4			
	L-Tryptophan	4.20×10^5	Co-binding with DPT	UV-Visible Spectroscopy	[38]
CB6Ar		3.40×10^4	Sodium Phosphate 10 mM pH 7, co-binding with MBBI, 27 °C	ITC	[39]
	L-Phenylalanine	4.20×10^4			
	L-Tyrosine	5.70×10^4	Sodium Acetate 50 mM pH 4.74, 22 °C	Fluorescence Spectroscopy	[40]
α -CDx	L-Tryptophan	3.20×10^6			
	Glycine	5.60×10^2			
	L-Alanine	1.12×10^3	Aqueous Formic Acid 50%(v/v), 25 °C	ITC	[35]
	L-Valine	1.62×10^3			
β -CDx	L-Phenylalanine	2.6×10^2			
	L-Tyrosine	5.0×10^1	Phosphate Buffer 0.01 M pH 7.0, 20 °C	Fluorescence Spectroscopy	[119]
	L-Tryptophan	5.24×10^4			
β -CDU	L-Phenylalanine	2.24×10^4			
	L-Tyrosine	1.33×10^4	Phosphate buffer 50 mM, 25 °C	ITC	[120]
	L-Alanine	5.00×10^3			
	L-Serine	1.97×10^4			
CPQ	L-Tryptophan	1.00×10^3			
	L-Tyrosine	4.54×10^2	Phosphate buffer 50 mM pH 7, 25 °C	¹ H-NMR	[116]
	L-Phenylalanine	1.06×10^2			
Cav[(CH ₂) ₂ CH ₃ , CH ₂ CH ₃]	N-Me-Leucine·HCl	1.70×10^5			
	N-Me-Alanine·HCl	1.40×10^5			
	Proline·HCl	1.20×10^4			
	Threonine·HCl	2.00×10^4	Methanol, 20 °C	ITC	
	Alanine·HCl	1.10×10^4			
	Tyrosine·HCl	1.10×10^4			
	Cysteine·HCl	7.30×10^3			

Table 1. *Cont.*

Host	Guest	K_a (M^{-1})	Conditions	Method	Ref.		
<i>N</i> -Me-Lysine-HCl 1:1	<i>N</i> -Me-Lysine-HCl 1:1	1.07×10^6	Methanol, 20 °C (amino acid:cavitand)				
	<i>N</i> -Me-Lysine-HCl 1:2	1.45×10^3					
Cav[$(CH_2)_2$ $CH_3Py^+Cl^-$, CH_2CH_3]	<i>N</i> -Me-Glycine-methyl ester-HCl	6.80×10^4	Methanol, 20 °C				
		3.43×10^3	Water, 20 °C				
	Sarcosine	1.89×10^3	Phosphate Buffer Saline 0.1 M pH 7, 20 °C				
	<i>N</i> -Me-Lysine-HCl	1.02×10^3	Water, 20 °C				
		1.49×10^3					
TeCPh	<i>L</i> -Tryptophan	1.13×10^3	Phosphate Buffer Saline 0.1 M pH 7, 20 °C		[42]		
	<i>N</i> -Ac-D-Aspartate	1.21×10^3	Phosphate Buffer 10 mM pH 7.4, 25 °C				
OA (R,R)-1	<i>N</i> -Ac-D-Glutamate	1.62×10^4	Tetramethylammonium chloride 0.1 M, 20 °C	pH-metric titrations	[53]		
	<i>N</i> -Ac-L-Aspartate	3.31×10^3					
OA (S,S,S,S)-2	<i>N</i> -Ac-L-Glutamate	3.40×10^2					
	<i>N</i> -Ac-L-Alanine	6.30×10^2					
GP3cat	<i>N</i> -Ac-L-Glutamate	2.10×10^3	1:1 90% water/10% Dimethyl sulfoxide	ITC	[59]		
	<i>N</i> -Ac-L-Glutamate	1.52×10^6	2:1 90% water/10% Dimethyl sulfoxide				

^a 1:2 host-guest stoichiometry. Overall binding constant in M^{-2} . The overall binding constant corresponds to the product of the stepwise binding constants for the 1:1 and 1:2 complexes, i.e., $K_{11} \cdot K_{12}$.

The first of these receptors to be tested was the phosphonate tweezer, which showed selectivity towards Lys, with even higher affinity towards its α -N/C-Protected form. NMR and ITC studies elucidated the interactions formed in these complexes where all the analytes tested showed a full incorporation of their side chain near the hydroquinone group (Figure 16). The binding process is aided by cation–π interactions between the positively charged groups of the amino acids and the electron-rich interior region of the tweezer, as well as by hydrophobic interactions between the alkyl chains of the analytes and the aromatic groups. However, the formation of the complexes are mainly driven by ionic interactions between the receptor’s anionic groups and the positive charges in the analyte—explaining the higher affinity towards Lys, seeing as arginine presents a more delocalized charge, diminishing the stability of this latter complex [24].

This was supported by later studies with molecular tweezers comprising different anionic moieties, in which the phosphate and sulfonate containing tweezers present the highest affinities, in contrast to phosphonate and carboxylate which present lower charge density. The lower affinity of most of these tweezers to non-protected forms of the amino acids is most likely due to the presence of repulsive forces between the receptor and the carboxylate group of the amino acids [25].

Another application of this concept was developed by Mandl and König, in which they made a receptor with two modified crown ether moieties, CEAT [118] (Figure 17). These are capable of binding O-methylated Lys, with an affinity of $4.0 \times 10^4 M^{-1}$, namely by the formation of H-bonds between the crown ether ring and the amines in the amino acid. However, these receptors have rather small affinities when in aqueous buffered solutions, while also needing the presence of a small percentage of methanol, which indicates that they are less suitable for physiological applications [118].

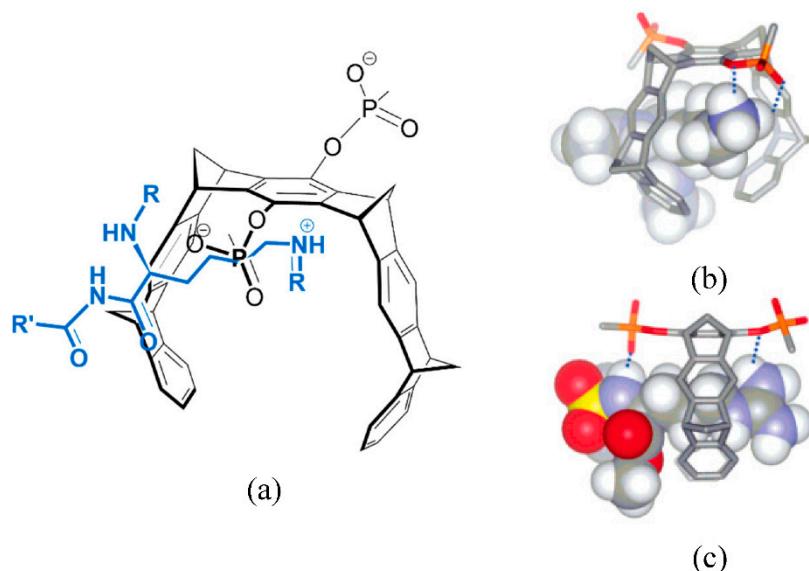


Figure 16. (a) Proposed structure adapted from Fokkens et al. and Monte Carlo simulations of the complexes between MPnT and Ac-Lys-OMe (b) and Ts-Arg-OEt (c)—MacroModel 7.1, Amber*, water, 5000 steps. Reprinted from [24] with permission. Copyright (2005) American Chemical Society.

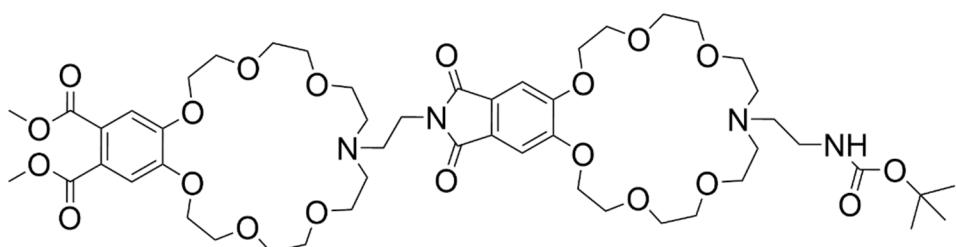


Figure 17. Structure of the molecular tweezer, CEAT, developed by Mandl and König [118].

Schmuck and Geiger also developed an open chain receptor for amino acids, based on a guanidinium pyrrole unit, with 2 added cationic groups, GP3cat (Figure 18). This receptor showed a remarkable increase in affinity towards acetylated L-Ala, in comparison to other mono or di-cationic forms of the receptor. This 1:1 complex is mainly stabilized by ionic interactions between the receptor and the carboxylate of the analyte, as well as by multiple hydrogen bonds that are formed at the same time with the same groups. Despite the extra carboxylate of glutamate, the 1:1 complex is more stable with alanine, although two receptors can bind to 1 molecule of glutamate, which increases the affinity towards this amino acid. Although promising, this system was not studied solely in water, still needing 10% DMSO in order to increase its solubility [59].

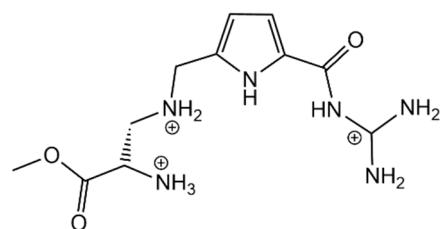


Figure 18. Structure of the guanidinium pyrrole receptor, GP3cat [59].

4. Recognition of Peptides: Taking Inspiration in Nature

One of the main targets for protein detection is the recognition of key elements, most often on the surface of the protein, seeing as this is the region most easily available for binding. For small peptides most of their residues are easily available for binding and is also easier to select a target region for detection. In either case, the best way to achieve this is by the recognition of the peptide side-chains or C and N-terminal amino acid residues [7]. In this section, the focus will be on the previously mentioned receptors, and other, more complex structures, applied in the selective recognition of target amino acid sequences in peptides [22,24–26,35,52,121–123].

4.1. Macroyclic Receptors

From the receptors presented before, there are a few that have received a larger focus and that their affinity towards certain residues and groups is well defined. Cyclodextrins, for example, have been shown to have affinity for hydrophobic amino acids like phenylalanine, but still bind weakly to the peptides studied in most works, while cucurbit[n]urils, which are selective for hydrophobic and basic amino acids form complexes with peptides and proteins that are significantly more stable, reaching nanomolar affinities in specific cases [37,108,124,125]. Charged macrocyclic receptors, as can be expected, bind oppositely charged residues, being this, together with the hydrophobic effect, the main driving forces for peptide binding in aqueous solutions while other non-covalent interactions may contribute in a small extent for overall binding energy [7,9,33]. However, it must be stressed that these secondary interactions may play a decisive role in the overall selectivity and structure of the complexes.

The simplest and non-modified forms of these macrocycles can be used for selective binding of peptides, which is reported in many studies [21,35,41]. Using these as a starting point, many other receptors have been developed, modifying the original structures to suit the type of amino acids and peptides that are being targeted [40,52]. There are also other types of macrocycles that, while less used for biological applications, have been adapted and show promise in the binding and detection of peptides, as was touched upon in the previous chapters [51,94,95].

4.1.1. Calixarenes

p-Sulfonatocalix[4]arene (SC4) binds not only cationic amino acids but also small cationic peptides. Complexes with polycationic peptides are, in general, more stable due to the formation multivalent ionic interactions. One of the first studies with polycationic peptides and SC4 was reported by Morel-Desrosiers and co-workers [22], in which they observe some changes in the way that peptides bind to this receptor, in comparison to the single amino acid. For Lys-Lys, Arg-Arg and Lys-Arg, the *N*-terminal amino acid is enclosed within the cavity of the receptor, with the C-terminal amino acid staying only near the entrance of the macrocycle. NMR studies showed that, between Lys-Lys and Arg-Arg, the latter binds deeper in the SC4 cavity, most likely due to the stabilization of the complex by cation–π interactions. The two other peptides differ slightly in the mode of binding—while Lys-Arg seems to bind with the *N*-terminal amino acid, lysine, in the cavity and arginine remaining outside the macrocycle, the Arg-Lys peptide binds with the C-terminal deeper in the cavity.

The tripeptides tested in this study showed higher stabilization than the remainder of the analytes. This is most likely due to two main factors: the multivalent ionic interactions, due to the presence of an added cationic group, increasing the sites of interaction and, so, increasing the stability of the complex [126]; a greater distance between the C-terminal and the sulfonate groups of the receptor diminishes the electrostatic repulsion between these two negatively charged groups and allows for more stable binding with the *N*-terminal amino acid, deeper in the cavity, and the middle amino acid, at the entrance of the cavity. This study also showed that, in general, the complexes formed with the amino acid L-Arg and Arginine peptides are slightly more stable than their L-Lys counterparts, due to the

presence of the guanidinium moiety that may undergo $\pi-\pi$ interactions with the calixarene cavity, as explained in Section 3.1.1. [22].

Ghale et al. [127] also used *p*-sulfonatocalix[4]arene for the detection of arginine-rich peptides, using an indicator displacement method. In this study, this receptor was applied for the detection of several cationic analytes in liposomes (with more focus on protamine, a natural occurring peptide with 21 arginine residues) and to study their transport across membranes, mediated by the membrane protein OmpF. The association constants presented in Table 2 further establish that, in general, this macrocycle forms more stable complexes with peptides with higher number of positively charged residues due multivalent ionic interactions, as commented above. Later studies show the same effect with an amphiphilic SC4 and its complexation with heptaarginine and other small arginine-rich cell-penetrating peptides [121].

The high affinities towards multivalent cationic peptides make this receptor a promising one, for in vitro study of several types of peptides—antimicrobial peptides [33], protamines [128], and mitochondrial signaling peptides [129]—but also make it a viable tool for applications in sensing of proteins and their modulation. Selected examples of this include the use of SC4 as a promoter of protein crystallization, protein assembly and as a masking agent, to alter proteins' surface properties [130–135].

This is the case of studies with the protein cytochrome c, with characteristic positively charged regions at the surface, composed of several lysine residues, which have a very important role in the protein–protein interactions that Cyt c can establish. McGovern et al. have shown that SC4 binds this protein at three different lysine residues (Figure 19), forming ionic and cation– π interactions with the residue, but being further stabilized by hydrogen bonds with the nearest polar residues and groups, as well as by the hydrophobic effect. The binding of SC4 results in a change in the overall charge at these sites, concealing the lysine residues and altering the protein–protein recognition properties of these biomolecules. Moreover, the changes at the protein surface enhanced the protein's ability to form crystals, indicating possible uses of SC4, not only in physiological conditions, for function modulation purposes, but also as an enhancer of crystallization [131].

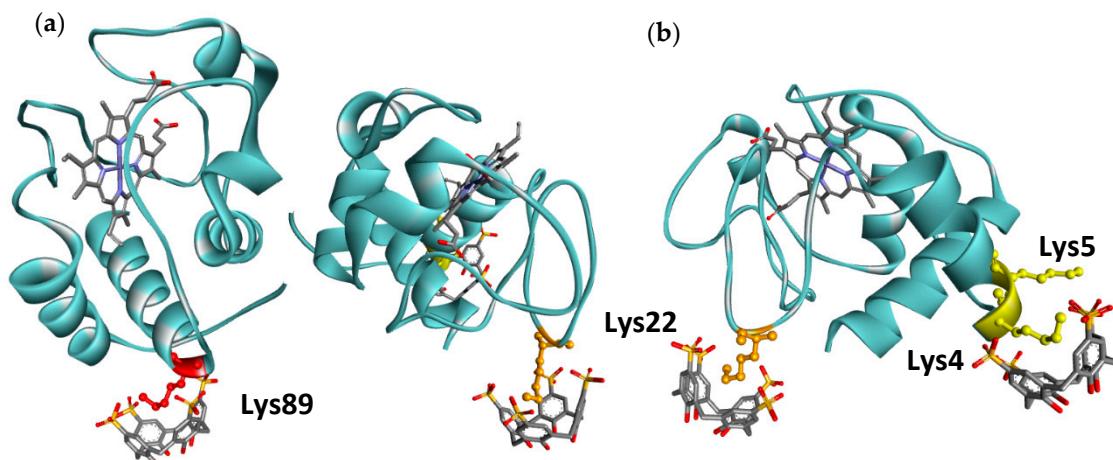


Figure 19. Image of 3TYI (structure obtained by McGovern et al. [131]) created with the software Discovery Studio 2020 Client (BIOVIA, Dassault Systèmes, San Diego: Dassault Systèmes, 2020) (a) Asymmetric unit of the crystal structure of the cytochrome c-SC4 complex, showing three interact Table 89. (red), Lys22 (orange), and Lys4 (yellow), at the surface of the protein, with respective interactions in bright orange. (b) Side view of one of the proteins in the asymmetric unit for better visualization of the Lys4 binding site; Lys5 is also represented in yellow, for it could form complementary interactions with the SC4 molecule in its proximity.

4.1.2. Calixpyridinium

This macrocycle was tested not only with amino acids, as previously mentioned, but with a glutamate dipeptide, to evaluate if the high affinity was maintained or if there were any changes in the binding mode. The same study showed that the calixpyridinium-peptide complex presented very similar stability as the one with only glutamate, which might indicate that the binding occurs mostly in the same way (Table 2) [52]. In the case of the dipeptide it is possible that the *N*-terminal residue is also inside the macrocycle's cavity but that the electrostatic repulsion between the positive *N*-terminal hinders the binding, resulting in equal affinity to that of L-Glu.

Despite the possible use of this system in the detection of peptides, it remains that calixpyridinium presents higher affinity constants towards nucleosides, in particular towards ATP, and so, being more interesting as a receptor for this anionic molecule [136].

4.1.3. Pillararenes

Much like calixpyridinium, pillararenes previously mentioned either were not further tested for their affinity towards peptides or were only tested with a couple of small peptides, yielding very similar results as for the binding of the respective amino acids. Carboxylatopillar[5]arene (CP5) was tested for two peptides—Ala-Arg-Ala and Ala-Lys-Ala—in the same study mentioned before [26]. NMR studies showed that, in this case, the CP5 receptor only encapsulated the side chain of the middle basic amino acids, stabilizing both by ionic interactions and the hydrophobic effect. Similar to the amino acid results, this macrocycle has five times higher affinity for the peptide with Arginine, most likely due to the formation of cation–π interactions, giving it a rather high selectivity for the presence of this amino acid.

4.1.4. Cucurbiturils

The binding of peptides to CB[n] constitutes a remarkable success story in the field of host-guest systems comprising this type of biological molecules. As previously mentioned, CB[n] generally display higher affinity for hydrophobic aromatic amino acids being the release of high energy water molecules from the macrocycle's cavity the main thermodynamic driving force contributing for the stability of the complexes [105]. Ion-dipole and hydrogen bonding interactions between carbonyls portals of the receptor and specific functional groups present in the peptide are believed to contribute in a less extension for the overall driving force but may have a decisive role in the observed selectivity. The high affinity interactions of *N*-terminal aromatic residues, in particular phenylalanine, with CB7 and CB8 constitutes a paradigmatic target for the selective binding behavior of this class of hosts [7,37,108,122,124]. For example, Phe-Gly and CB7 form a 1:1 complex with a $K = 3.0 \times 10^7 \text{ M}^{-1}$ while the complex formed with Gly-Phe is much weaker ($K = 1.3 \times 10^3 \text{ M}^{-1}$), as is summarized in Table 2 [122]. In both cases the phenylalanine residue is included in the CB7 cavity. However, while in Phe-Gly this binding mode favors attractive ion-dipole interactions between the carbonyl portals and the ammonium group and minimizes the repulsion of the portal with the carboxylate group, for Gly-Phe the carboxylate is closer to the carbonyl group precluding optimal inclusion by electrostatic repulsion. Furthermore, the attractive interaction between the ammonium group in the Gly residue with the carbonyl portals of the receptor imposes a significant entropic penalty due to conformational restriction of the guest [122]. Although displaying slightly lower binding constant, both tryptophan- and tyrosine-glycine dipeptides display similar behavior to that described for Phe-Gly/Gly-Phe [122]. It is also worth noting that the smaller CB6, in contrast to CB7 and CB8 (see below), does not significantly bind peptides with terminal aromatic residues, showing much higher affinity for those comprising terminal lysine [122].

Urbach and co-workers [37,108] focus their earlier work on the binding of *N*-terminal amino acids to the larger CB8. As described above for CB7, CB8 also binds preferentially peptides containing aromatic amino acids at the *N*-terminal. These receptor-analyte pairs can form 1:1:1 complexes in the presence of methyl viologen as an auxiliary guest [37], but

can also assist the dimerization *N*-terminal aromatic peptides in the absence of methyl viologen by forming 1:2 host–guest complexes. CB8 shows selectivity towards Trp-Gly-Gly and Phe-Gly-Gly (Figure 20) [108], forming CB8-peptide₂ complexes with overall binding constant up to $3 \times 10^{11} \text{ M}^{-2}$ (Table 2) [108,137,138]. While the arguments described above to rationalize the selective binding of *N*-terminal aromatic residues to CB7 are also likely to hold here, the crystal structures shown in Figure 20, besides showing those interactions, also demonstrate that hydrogen bonding between the amide protons and the oxygen atoms in the carbonyl portals may also contribute to increase the stability of these complexes [108].

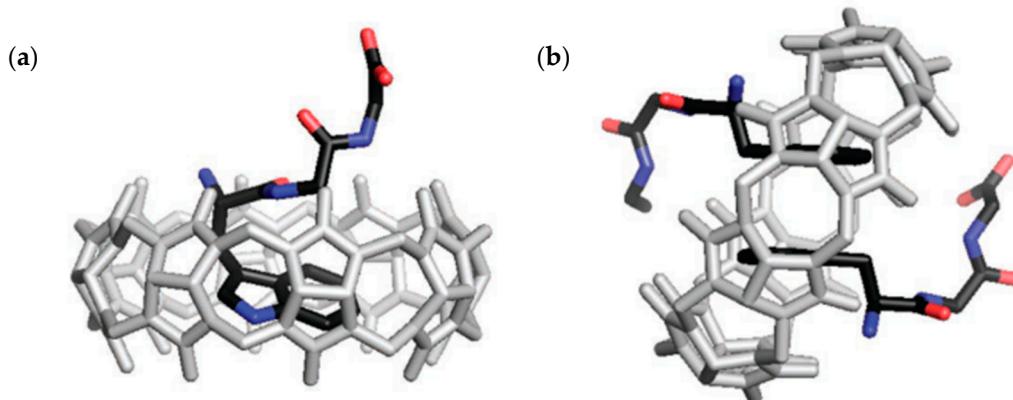


Figure 20. Crystal structure of the complexes with CB8 (structure in white) and the peptides (structures in black)—(a) CB8-(Trp-Gly-Gly) and (b) CB8-(Phe-Gly-Gly)2. Despite the formation of 1:2 complexes for both these peptides, Heitmann et al. were only able to obtain this complex structure for Phe-Gly-Gly. Reprinted with permission from [108]. Copyright (2006) American Chemical Society.

More studies have been done with CB8, screening for any specificity among peptides with *N*-terminal aromatic residues [125]. Most of these were performed with tripeptides, in which the *N*-terminal residue was aromatic, and the two others varied, in order to assess their importance in the binding process. It was discovered that, even though most of the peptides bind in a 2:1 ratio to CB8, the sequence Tyr-Leu-Ala showed high affinity (in the nM range) for this macrocycle. However, it bonded in a 1:1 ratio instead of the 2:1 observed previously. This was not only true for Tyr-Leu-Ala; when tyrosine was the *N*-terminal amino acid, some peptides formed a 1:1 complex with CB8, because the adjacent amino acid could also be encapsulated in the CB8's cavity along with L-Tyr, leaving the C-terminal at the entrance of the macrocycle (Figure 21). These peptides were of the Tyr-X-Ala kind (with X being either Leu, Lys, Phe or Tyr), and did not bind in this way to CB8 when X and Ala were switched. This high affinity towards a Tyr-Leu-Ala peptide is a step towards developing simple supramolecular systems that mimic the specificity displayed by antibody-based receptors.

The same researchers proceeded to do further screening, namely with a library of 144 tripeptides, X1-X2-Ala. X1 consisted of amino acids with large hydrophobic or cationic side chains, seeing as this will be the residue farther inside CB8's cavity and the larger size is needed in order to have 1:1 binding. X2 corresponds to 18 possible amino acids, leaving off only Trp and Cys from the existing 20 human physiological amino acids, due to possible interference in the measurements [123].

This study showed that CB8 had affinity towards peptides with Tyr, Phe, Ile, and Leu at X1 and Tyr, Phe, Leu, and Lys at X2. Sequences with Met, Arg and Lys also showed considerable affinity. The affinity towards methionine at the X1 position was the one explored in this study, seeing as Met is often found on *N*-terminal of proteins before further post-translational modification, being of particular interest to study these PTM mechanisms [139].

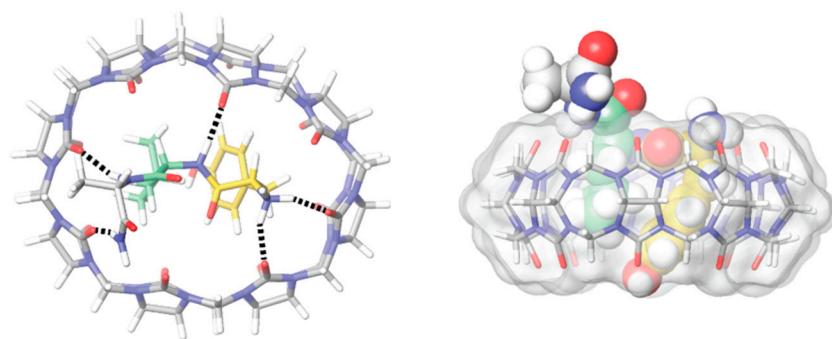


Figure 21. Semiempirical structure obtained by Smith et al. (2015) of the complex CB8-(Tyr-Leu-Ala)—up view (left) with the electrostatic interactions represented by dashed lines, and side view (right), where it is observable the two *N*-terminal residues inserted in the cavity of the macrocycle. Reprinted with permission from [125]. Copyright (2015) American Chemical Society.

Analyzing the binding by ITC studies, showed that the second amino acid should be large and hydrophobic or cationic, but without branching in the beta carbon, while the third amino acid should be small, so as to avoid having stereochemical hindrances [123]. These types of tripeptides show sub micromolar affinities towards CB8 and are also a plausible target for the detection of proteins by their *N*-terminal, seeing as, despite methionine being a common site of excision, large amino acids in the position next to Met can block the binding of the enzymes that catalyze this process, which distinguishes these *N*-terminal sequences from the excised ones [139]. Additionally, this functionality of the macrocycle can be used as a protection against excision of the *N*-terminal, for proteins presenting these sequences [109].

More recent studies focus on different modes of peptide binding to the receptor CB8, with a focus on peptides containing the pairs of residues Tyr-Leu or Phe-Leu. Here it is shown that CB8 can bind to these residues even if farther away from the *N*-terminal, with a diminished affinity by one order of magnitude, and that the binding can occur not only in a 1:1 ratio, but also in a 1:2 ratio for longer peptides with repeated Tyr-Leu sequences (Tyr-Leu-Ala-Gly-Gly-Ala-Leu-Tyr) [140].

Due to the presence of aromatic amino acids in insulin's *N*-terminal, this protein comes up as a possible target for this type of receptors. With this in mind, Chinai et al. evaluated the affinity of this system towards human insulin and a mutant variety without the *N*-terminal aromatic amino acid [124]. This latter variant showed an obvious decrease in affinity towards CB7 when compared to human insulin. The latter showed similar affinity towards the macrocycle as Phe-Gly-Gly. Furthermore, it was observable a slight change in conformation in insulin to allow for this binding, exposing the *N*-terminal to the complexation, but without altering the overall structure of the protein. This implies that even less exposed *N*-terminal amino acids could be targets for this type of sensing without any major interference with the rest of the structure, while also distinguishing them from other aromatic amino acids in a protein's surface or in a peptide's sequence (Figure 22). Additionally, the *N*-terminal is usually a more exposed and flexible region in proteins, being possible that this could be more easily transposed to studies with peptides with a corresponding sequence.

As can be seen by this example, CB[n]'s can be used for the detection of proteins, due to its specificity towards aromatic and cationic residues. Control over protein assembly and function is another route that is being explored for this supramolecular system [9,141–148]. Brunsved and co-workers have used CB8 as a tool to aid in the dimerization of proteins, making use of the target Phe-Gly-Gly motif as an *N*-terminal modification, in order to better control enzyme activity [149] and protein assemblies [142].

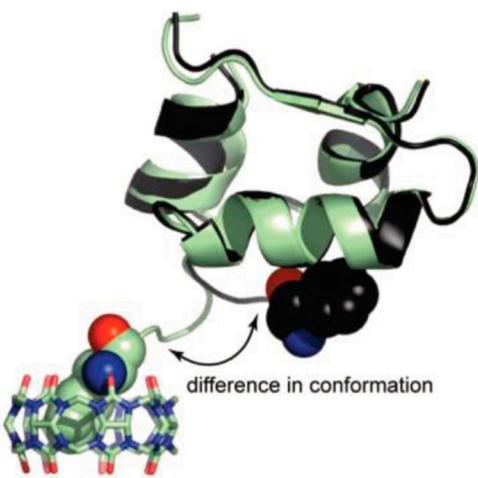


Figure 22. Crystal structure of the complex formed between CB7 and human insulin (light green) and of free insulin (black). Reprinted with permission from [124]. Copyright (2011) American Chemical Society.

The second example consists of the use of a modified peptide, with a 14-3-3 protein binding site at the C-terminal and with an Phe-Gly-Gly moiety at the N-terminal. This structure allows for the binding of this peptide to the 14-3-3 protein, which presents two binding sites, with affinities comparable to those of the physiological targets. CB8 can be added to stabilize this bivalent complex by 14 times, by complexation of the two N-termini of the peptides bound to this protein, serving as a tool to switch the “availability” of the 14-3-3 binding site [142].

A more recent study reported the role of CB8 as a modulator of enzymatic activity using caspase 8, an enzyme that is typically only active upon dimerization. In this work Dang et al., after observing that the Casp-8 N-terminal residues of the two monomers are in close proximity in the dimeric active state, engineered this region of the Casp-8 to have an Phe-Gly-Gly motif [149]. This motif allowed for the use of CB8 to control the dimerization in the inactive mutants and resulted in an easily modulated, dimeric form of Casp-8 active towards a synthetic ligand as well as its physiological substrate, caspase-3.

The CB[n]’s affinity towards basic residues has also been explored in the scope of protein recognition. Namely, Crowley and co-workers tested CB7’s affinity towards a demethylated lysine in a protein, *Ralstonia solanacearum* lectin (RSL). Despite having other methylated lysine residues in its sequence, CB7 recognized selectively the Lys34Me₂, which is located in a more accessible region of RSL, yielding an affinity in the mM range (Table 2). Moreover, CB7 has also been successfully exploited in this context to create novel protein assemblies around CB7 trimers and tetramers, in the crystal structure of the CB7-RSL complexes [150].

4.1.5. Cyclodextrins

As described in the previous sections, the properties of these natural occurring macrocycles translate to higher affinities towards aromatic residues or residues with small, apolar side chains in the context of peptides and proteins [35,41,119,151,152]. The same types of interactions occur with both amino acids and peptides—the binding at the cyclodextrin cavity is generally driven by hydrophobic effects, with some hydrogen bonds formed with the outer shell of the macrocycle. In most cases, these interactions only allow for small affinities, only surpassing mM affinities in 50% methanol solutions [152], or with further modification of the original macrocycle [151].

4.1.6. Crown Ethers

A different approach for detection of aromatic amino acids and small peptides containing these residues was presented by Weißenstein et al. [153]. In their studies they joined

the properties of two different moieties—as readout mechanism they used a perylene bisimide dye (PBI), which has high fluorescence quantum yield and can be quenched by photoinduced electron transfer (PET)—this process can occur when this group forms bonds with electron rich moieties, e.g., aromatic amino acids which can bind to PBI by $\pi-\pi$ stacking of aromatic groups; the other distinct moiety is composed of two crown ethers that serve as a receptor for ammonium cations, e.g., the *N*-terminal of short peptides or amine groups present in cationic amino acids. The presence of two crown ethers allows for the binding of two analytes, while PBI can form two $\pi-\pi$ stacking bonds with the aromatic amino acids (Figure 23).

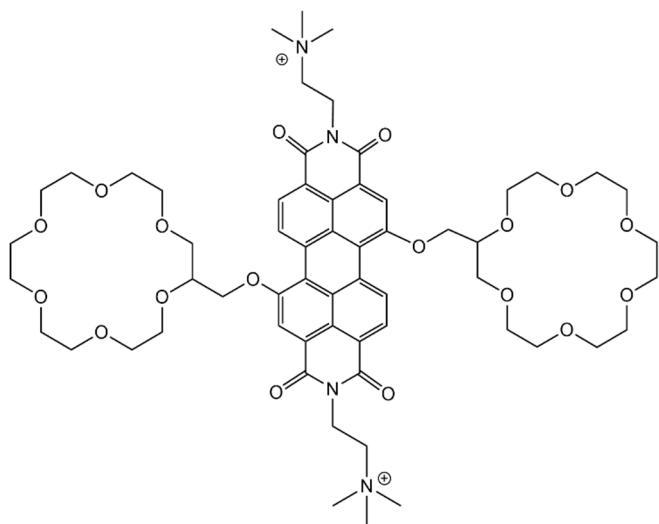


Figure 23. Structure of the crown ether molecular tweezer, perylene bisimide dye (PBI) [153].

The highest affinity of PBI, at least one order of magnitude above that of other peptides and amino acids tested, was towards Ala-Trp. This peptide is believed to have the optimal characteristics for binding to occur due to the separation between the ammonium ion at the *N*-terminal and the aromatic residue. It must be ideal for the formation of interactions between the *N*-terminal ammonium group and the crown ether and between the aromatic side chain and the perylene bisimide unit, at the same time. The downside of this probe is that its studies were not made in water (probably due to poor solubility of the receptor) and so, the conclusions taken, and the overall system are, most likely, not applicable in water and physiological conditions.

Similar systems have been explored before, in water, as is seen as early as 1998 in the work of Hossain and Schneider [154]. They used a crown ether and a trimethylammonium moieties to bind to the peptide's *N* and C-terminals, respectively. These moieties were connected by a hydrophobic linker, with an added fluorescent dansyl unit for the detection of the biomolecules (Figure 24). The synthesized receptor, CENMe₃, reached affinities up to 10^3 M^{-1} , for both Gly-Phe-Gly and Gly-Trp-Gly, at least 1 order of magnitude above the affinities towards other tested peptides. Even though this receptor shows specificity towards the middle positioning of an aromatic residue in this type of peptides, for biological applications, a 10^3 M^{-1} affinity constant is not the most desired value, since physiological concentrations are much lower than those needed for considerable complex formation.

4.1.7. Cavitands

Cavitand receptors have seen multiple applications in recognition of small organic molecules, including peptides and proteins. Some examples comprise the integration of these macrocycles in sensing scaffolds [155,156], nanostructures [157], as well as their use in solution [158–160]. The latter includes a group of self-folding deep cavitands, DCv (Figure 25), which have been applied in the detection of select histone sequences and their

respective post-translational modifications (PTM) in water. These interact with lysine residues, showing different affinities depending on the R₅ group's nature (Figure 25) and the degree of methylation of the guest. The complexes established are stabilized mainly by cation–π interactions but also with the contribution of a variety of other electrostatic interactions [157,159,160]. Furthermore, all three of the deep cavitands schematized have been applied in a dye displacement assay for the screening and monitoring of the PTM on a select library of peptides. These included several methylation sites that are recognized by both histone methyltransferase and demethylase, allowing for an analysis of their activity at several of the specific sites in which they act upon [159].

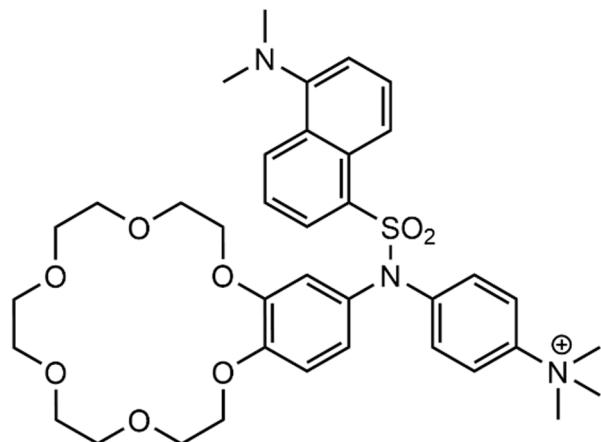


Figure 24. Structure of the crown ether and trimethylammonium sensor, CENMe₃ [154].

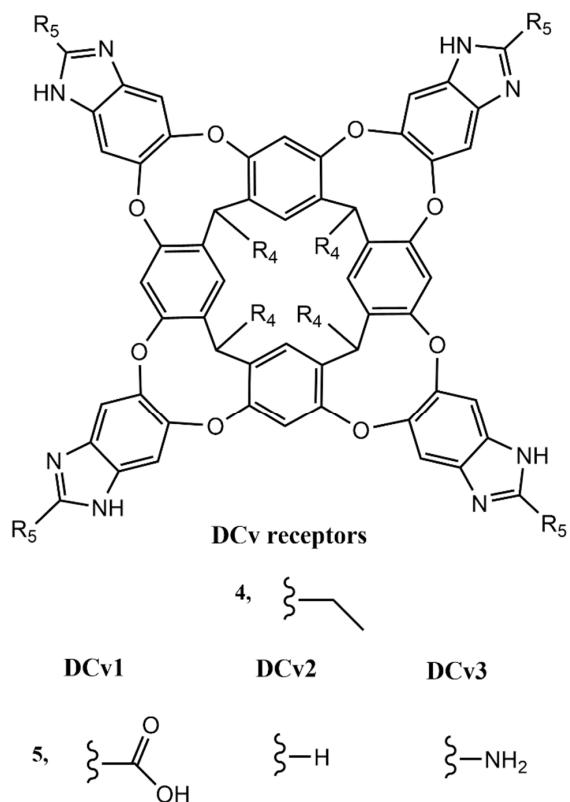


Figure 25. Structure of the deep cavitands, DCv1, DCv2 and DCv3 [158,159].

4.2. Open-Chain Receptors

Besides the obvious interest in peptide and protein sensing, many receptors for peptides have been explored focusing on the role of these biomolecules as targets for the action of several pharmacological agents. The antibiotic vancomycin and its derivatives are amongst the most recognized examples.

Vancomycin binds a specific sequence in the cell wall of gram-positive bacteria, D-Ala-D-Ala-OH, disabling the action of a transpeptidase at this site, which has a very important role in the crosslinking of several peptide strands in the peptidoglycan layer and stabilizing it. Without this crosslinking, the cell envelope is severely destabilized, and lysis of the cell will occur if there are variations in the local osmotic pressure. In drug resistant strains of the same bacteria, this target sequence is mutated to D-Ala-D-Lac-OH. This change in one amino acid breaks a bond between vancomycin and the peptide and adds electrostatic repulsion between the two oxygens indicated below (Figure 26). This diminishes the affinity constant of vancomycin by two orders of magnitude, but does not affect the binding of the bacterial transpeptidase [6].

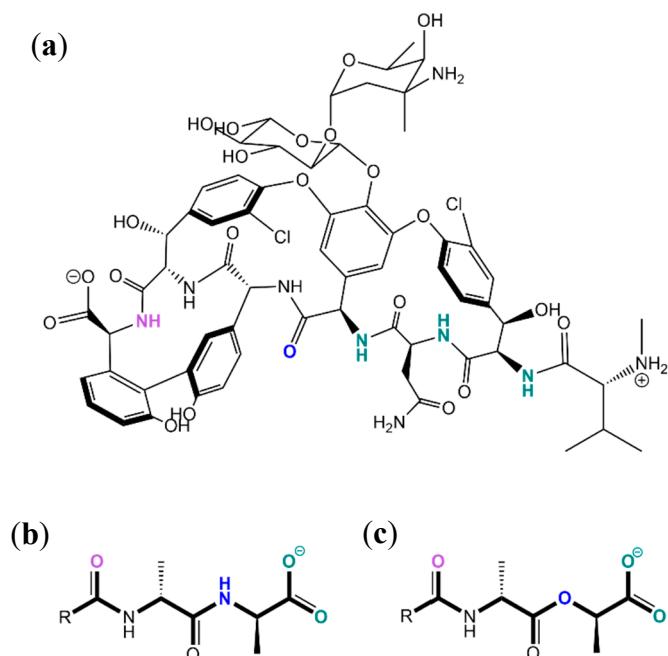


Figure 26. (a) Vancomycin, (b) its target sequence in the cell wall of gram positive bacteria, D-Ala-D-Ala-OH and (c) the mutated sequence, D-Ala-D-Lac-OH. The sites in vancomycin highlighted in pink, blue and green interact with the same colored groups in the sequences. In the mutated sequence, one of the hydrogen bonds is broken and there is instead electrostatic repulsion between the groups in blue [6].

In order to circumvent bacterial mechanisms of resistance, Ellman and co-workers [161] synthesized a library of water soluble receptors, based on vancomycin's structure, but with a peptidyl sequence in the region where the repulsive forces were introduced when a mutation occurred. The peptidyl group has freer rotation than the rigid group of vancomycin, and this will allow for the increase of the distance between receptor and peptide, diminishing the effect of the repulsion when binding occurs. The receptor shows better complementarity to the sequences and, so, the highest affinity was observed for VD1 (10^4 M^{-1}), represented in Figure 27.

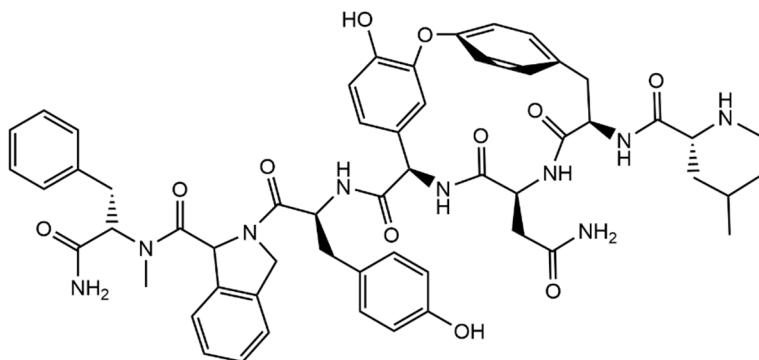


Figure 27. Structure of the peptidyl vancomycin derivative receptor, VD1 [161].

In more recent studies, with other receptors, smaller libraries are being utilized, taking advantage of information obtained from previous studies to focus the design of new receptors towards the properties that are needed [162].

In a study by Schmuck and Heil, with the tetrapeptide *N*-Ac-D-Glu-L-Lys-D-Ala-D-Ala-OH—EKAA—in mind as a target (a bigger fragment of the peptide recognized by vancomycin), a small library of water-soluble receptors was created, with 512 members. They used a similar strategy as the one used for VD1, synthesizing these receptors with a peptidyl chain, with three variable residues, adding the guanidiniocarbonyl pyrrole binding motif (GCP), which was previously shown to bind to the C-terminal of peptides (as mentioned in Section 2.3, relative to the binding of single amino acids) [117]. The GCP binding motif shows similarities with vancomycin in the types of interactions established with the carbonyl group of this sequence with several H-bonds stabilizing both complexes. In the GCP-peptide complex there is the additional ionic bond previously mentioned.

The residues of the peptidyl group varied between Lys, Tyr, Ser, Glu, Phe, Val, Leu, and Trp, being expected that some of these can form more ionic interactions with the analyte and further stabilize the resulting complex. Due to the pyrrole moiety it was possible to monitor the binding of the target peptides with the receptors from the variations in the absorption and fluorescence spectra of the receptors. It was found that the receptor with the sequence Lys-Lys-Phe, GCP-KKF (Figure 28), was the one which presented higher affinity towards both EKAA and the opposite sequence, AAKE, but with markedly higher affinity towards the peptide with Ala at the C-terminal ($K_a = 1.71 \times 10^4 \text{ M}^{-1}$, 4 times higher than that for AAKE), indicating the importance of the C-terminal in the binding of the peptide. Furthermore, the lysine residue also takes part in the binding with the C-terminal, stabilizing the binding to the tetrapeptide further [163].

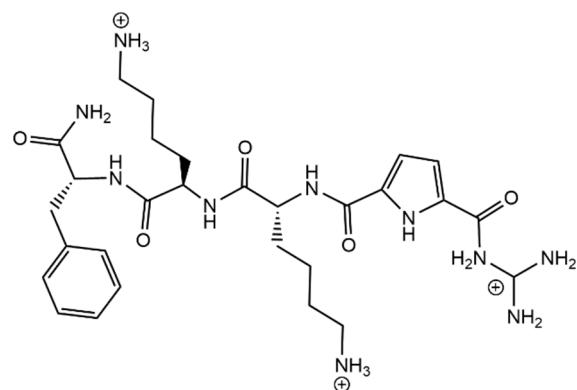


Figure 28. Structure of the peptidyl Guanidiniocarbonyl pyrrole receptor, GCP-KKF [163].

The receptor which showed highest affinity towards EKAA was also tested for its affinity towards other tetrapeptides, with variable residues, originating a library of ~300 target

tetrapeptides, including those with D-Ala and D-Lac in the C-terminal (the first similar to the native sequence of Gram positive bacteria peptidoglycan peptidyl chain and respective mutated sequence, related to vancomycin resistance). The GCP-KKF receptor showed an even higher affinity towards Ac-D-Glu-D-Glu-D-Glu-OH ($K = 2.65 \times 10^4 \text{ M}^{-1}$) and a very accentuated stereoselectivity for D-Ala in D-Glu-D-Glu-D/L-Ala-D-Glu-OH peptide (almost $4\times$ higher than for L-Ala, K_a (L stereoisomer) = $1.40 \times 10^3 \text{ M}^{-1}$ and K_a (D stereoisomer) = $4.50 \times 10^3 \text{ M}^{-1}$) [164].

This shows that focusing on libraries and taking inspiration on known and well-studied host-guest complexes helps to understand the bonds created and how to further stabilize other peptide-receptor complexes [6,163,164].

Molecular Tweezers

Molecular tweezer and clips constitute another example of small supramolecular receptors that have been widely explored for the recognition of peptides. These compounds offer an (often electron-rich aromatic) pocket-like binding site, which is usually decorated with functional groups that may provide additional sites of interaction and water-solubility [165–167].

The most popular molecular tweezers with applications in the recognition of amino acids and peptides are probably those developed by the Klärner group, which bind preferentially basic amino acid lysine and arginine, as mentioned in Section 3.2. [167]. Some of these compounds show a tendency to form self-assembled dimers, a process that may compete with the complexation of selected targets and lead to concentration dependent apparent binding constants [168]. However, anionic tweezers such as the ones shown in Figure 15 (see Section 3.2) are fluorescent, allowing the monitorization of the complexation process and determination of the binding constants at low concentrations [168]. Besides the amino acids tested, a small library of bioactive peptides with basic residues was also investigated, with importance and applications in several fields [167]. The results showed that the stability of the complexes depends on the anionic groups decorating the receptor and increases according to the following trend: O-methylenecarboxylate < phosphonate < sulfate < phosphate. Due to their electron-rich cavities and anionic functional groups this type of receptor fails to bind peptides lacking basic amino acids, showing, in the other hand, high affinity (μM range) for peptides bearing arginine and/or lysine residues, being the last generally more stable probably due to charge delocalization in the guanidinium groups. The highest affinities were reported for the peptides comprising several adjacent lysine residues at the N-terminal which provide more sites to form ionic interactions while minimizing the repulsion of the C-terminal [25].

In general, the complexes seem to be slightly more stable in water than in the methanol, demonstrating the important role of the hydrophobic effect that, nevertheless, is counterbalanced by the increasing strength of the ionic interactions in the organic solvent. Noteworthy, although the binding constant present modest variation upon changing from aqueous to organic media, NMR studies clearly showed that the type of complexes can be dramatically different depending on the solvent: in water the formation of inclusion complexes predominate while methanol favors external complexes [168]. However, it should be stressed that this behavior cannot be generalized and depends on the nature of the anionic groups.

The phosphonate, phosphate and sulfate containing molecular tweezers all complexed peptides with at least one L-Lys residue, while the carboxylate containing tweezer had rather low affinities or did not bind at all to the peptides studied, similarly to their behavior in the binding of the amino acids.

The phosphate molecular tweezer, MPT, has been of great interest, due to its high affinity and specificity towards the positive residues of peptides. Early on, they have been tested as a possible pharmacological agent [25,169–171] and showed positive therapeutic effects in tests with animal models of several diseases, without significant cytotoxicity—they bind to amyloidogenic lysine and arginine-rich proteins (e.g., α -synuclein), inhibiting

their aggregation [25]; MPT also has shown the ability to disrupt viral envelopes (e.g., Zika, Ebola [170], HIV [169,170]), due to their high levels of sphingomyelin and cholesterol. The carboxylate molecular tweezer, MCT, also shows to inhibit viral activity and disrupt the viral envelope, despite not binding strongly with cationic residues—it was hypothesized, in a 2020 study [171], that the MCT and MPT tweezers are able to interact at virus-like lipid rafts, forming inclusion complexes with sphingomyelin, and disrupt them at the interface with the rest of the membrane, by increasing the already high surface tension of these envelopes.

Other studies explored the importance of the linker in these tweezer receptors, showing again that it is necessary to have some rigidity in the linker between the “arms of the tweezer” and a certain angle and distance between the arms to be able to encase the peptide between them [172–174].

4.3. Self-Assembled Coordination Cages

Metal organic capsules (or self-assembled coordination cages) constructed from organic ligands and metal ions often display well defined binding pockets, suitable to accommodate complementary guest molecules in their interior. These have been widely explored as receptors for a variety of molecular targets [175–177]. The positively charged Pd₆L₄ coordination cage (Figure 29) reported by Fujita was shown to bind complementary oligopeptides with high affinity and impressive sequence selectivity [178,179]. The Pd₆L₄ binding pocket was found to be large enough to accommodate up to three amino acid residues in its interior. By analyzing tripeptides with different sequences strong binding was observed for the Ac-Trp-Trp-Ala-NH₂ sequence ($K > 10^6 \text{ M}^{-1}$) while very similar sequences, such as Ac-Trp-Ala-Trp-NH₂, Ac-Ala-Trp-Trp-NH₂, Ac-Trp-Trp-Gly-NH₂, Ac-Trp-Tyr-Ala-NH₂ showed much lower affinity (affinities of the first two sequences can be seen in Table 2) [179]. The results suggest that the cooperative interactions between the two indole rings and the Ala methyl group with the cage and as well intramolecular interactions between the residues (such as $\pi-\pi$ and CH- π) may impart a decisive contribution to the thermodynamic driving force. This seems to put in evidence the potential advantages of using receptors with cavities large enough to accommodate more than one residues in its interior as potential strategies to reach exquisite selectivity [179].

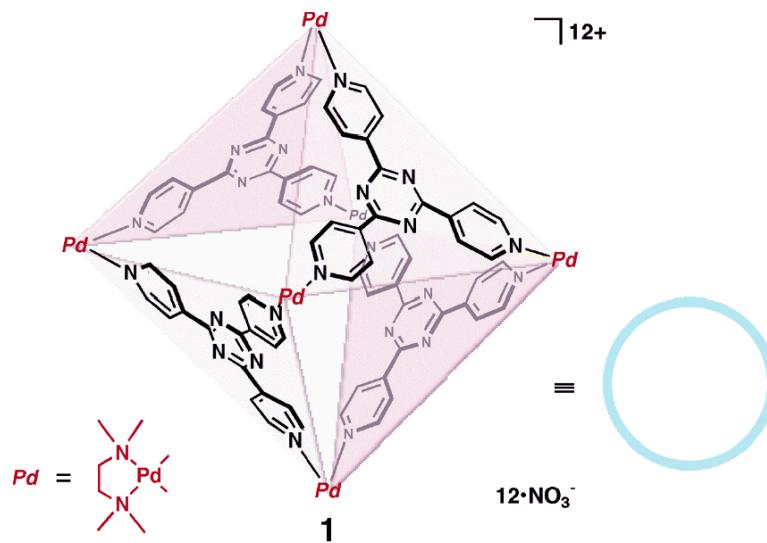


Figure 29. Chemical structure of the Fujita’s Pd₆L₄ coordination cage. Reprinted with permission from [179]. Copyright (2005) American Chemical Society.

More recently, Nitschke and co-workers reported a large Fe₈L₆ cubic coordination cage assembled from Zn-porphyrin based ligands and iron(II) metal ions, FeZnP, able to encapsulate guest molecule containing imidazole functional groups that were anticipated

to interact with the Zn-porphyrin capsule walls [180]. After demonstrating the affinity of the capsule for this type of guests, the authors investigated the ability of the nanocontainer to bind peptides with histidine residues in a 1:1 acetonitrile:water solvent system. Clavanin A, a peptide antibiotic, containing four histidines among its 23 amino acids was found to bind strongly to the capsule with a dissociation constant of 80 nM in a 1:2 binding manner (Table 2). A less hydrophobic peptide containing only three histidines was found to bind 100-fold more weakly, while peptides lacking histidine residues did not interact at all.

Table 2. Binding constants (K_a) for the formation of host-guest complexes between small peptides and synthetic supramolecular receptors.

Receptor	Guest	K_a (M^{-1})	Conditions	Method
SC4	Arg-Arg	7.00×10^3		
	Arg-Lys	3.90×10^3		
	Arg-Arg-Arg	3.30×10^4	90% water/10% Deuterium oxide	^{1}H -NMR
	Lys-Lys	3.40×10^3	Phosphate Buffer 10 mM pH 8, 20 °C	[22]
	Lys-Arg	3.70×10^3		
	Lys-Lys-Lys	3.30×10^4		
	Trp-Lys-Arg-Thr-Leu-Arg-Arg-Leu	1.20×10^6	HEPES buffer 10 mM pH 7, 20 °C	Fluorescence Spectroscopy
	Protamine	1.20×10^9	Phosphate buffer 10 mM pH, 25 °C	[127]
	Cyt C binding site 1	1.20×10^3	20 mM Potassium dihydrogenophosphate, 50 mM NaCl, 1 mM sodium ascorbate, 10% D ₂ O, pH 6, 30 °C	[131]
	Cyt C binding site 2	6.30×10^2		
SC4-C5	(Arg) ₇	7.00×10^7		Fluorescence Spectroscopy
	(Arg) ₇	2.90×10^7	HEPES buffer 10 mM pH 7, 25 °C	[121]
C4Pyr	Trp-Lys-Arg-Thr-Leu-Arg-Arg-Leu	2.80×10^6		
	Glu-Glu	1.7×10^3	Water pH 6	[52]
CP5	Ala-Arg-Ala	4.20×10^3		
	Ala-Lys-Ala	7.50×10^2	Deuterium oxide pD 7.2, 25 °C	^{1}H -NMR
CB7	Phe-Gly	3.00×10^7		
	Gly-Phe	1.30×10^3		
	Tyr-Gly	3.60×10^6	Water, 25 °C	[122]
	Gly-Tyr	2.00×10^2		
	Trp-Gly	5.60×10^5		ITC
	Gly-Trp	2.80×10^2		
	Phe-Gly-Gly	2.80×10^6	Sodium Phosphate 10 mM pH 7, 27 °C	[124]
	Insulin	1.50×10^6		
	RSL-Lys34Me ₂	1.00×10^3	20 mM Potassium Phosphate, 50 mM Sodium Chloride, 1.2 mM α -methyl-L-fucoside, 90% H ₂ O/10% D ₂ O, pH 6, 30 °C	$^{1}H^{15}N$ HSQC
	Trp-Gly-Gly	1.30×10^5		
CB8	Gly-Trp-Gly	2.10×10^4		
	Gly-Gly-Trp	3.1×10^3	Sodium Phosphate 10 mM pH 7, co-binding with MV, 27 °C	ITC
	Gly-Gly-Trp-Gly-Gly	2.50×10^4		[37]

Table 2. Cont.

Receptor	Guest	K_a (M^{-1})	Conditions	Method
CB8	Phe-Gly-Gly ^a	1.50×10^{11}	Sodium Phosphate 10 mM pH 7, CB8-peptide ₂ complex, 27 °C	[108]
	Trp-Gly-Gly ^a	3.60×10^9		
	Tyr-Leu-Ala	1.39×10^8		
	Tyr-Lys-Ala	5.00×10^6		
	Tyr-Phe-Ala	3.44×10^6		
	Tyr-Tyr-Ala	1.43×10^6		
	Met-Phe-Ala	7.14×10^6		
	Met-Tyr-Ala	4.00×10^6		
	Met-Leu-Ala	1.39×10^6		
	Met-Lys-Ala	3.85×10^5		
MPnT	Met-Tyr-Gly-Gly-Tyr	6.25×10^6	Sodium Phosphate 10 mM pH 7, 27 °C Sodium Phosphate 10 mM pH 7, co-binding with peptide-MV conjugates, 27 °C	[123]
	Met-Leu-Gly-Gly-Tyr	3.33×10^6		
	Leu-Met-Gly-Gly-Tyr	6.25×10^6		
	Met-Lys-Gly-Gly-Tyr	2.38×10^6		
	Gly ₂ -Trp-Gly ₂	2.20×10^4		
	(Gly ₂ -Trp-Gly ₂) ₂	5.00×10^5		
	Asp ₂ -(Gly ₂ -Trp- Gly ₂) ₃ -Asp ₂	4.70×10^6		
	Lys-Ala-Ala	1.20×10^3		
		1.10×10^3		
	Lys-Lys-Leu-Val-Phe- Phe	1.41×10^4		
MPT		3.80×10^4		¹ H-NMR [24]
	Lys-Thr-Thr-Lys	5.50×10^3		
	Lys-Thr-Thr-Lys-Ser	4.20×10^3		
	Gly-Arg-Gly-Gly	9.00×10^2		
	Arg-Gly-Asp	1.20×10^3		
		1.00×10^3		
	Lys-Ala-Ala	3.33×10^4		
	Lys-Leu-Val-Phe-Phe	5.00×10^4		
	Lys-Lys-Leu- Val-Phe-Phe	2.50×10^5		
	Lys-Lys-Leu-Val-Phe- -Phe-Ala-Lys	1.43×10^5		
MST	Lys-Lys-Lys-Lys	1×10^5	Phosphate buffer 200 mM pH 7.6	Fluorescence Spectroscopy [25]
	Arg-Gly-Asp	1.16×10^4		
	Lys-Ala-Ala	3.30×10^3		
	Lys-Leu-Val-Phe-Phe	2.63×10^4		
MCT	Lys-Ala-Ala	30	Phosphate buffer 10 mM pH 7.2	
PBI	Ala-Trp ^a	3.72×10^9	Acetonitrile:Methanol (9:1) 2:1 host guest complexes, 23 °C	[153]

Table 2. Cont.

Receptor	Guest	K_a (M^{-1})	Conditions	Method	
CENMe ₃	Gly-Gly-Gly	2.10×10^2	Water, 25 °C	[154]	
	Gly-Trp-Gly	2.15×10^3			
	Gly-Phe-Gly	1.7×10^3			
α -CDx	Gly-Gly-Phe	2.15×10^2	Formic Acid 50%(<i>v/v</i>), 25 °C	ITC	[35]
	Gly-Leu	2.63×10^2			
β -CDx	Gly-Val	2.82×10^2	Phosphate Buffer 0.01 M pH 7.0, 20 °C	Fluorescence Spectroscopy	[119]
	Tyr-Ile-Gly-Ser-Arg	2.24×10^2			
	Tyr-Gly-Gly-Phe-Leu	1.23×10^2	Bicarbonate buffer 0.2 M pH 9, 25 °C		[151]
	cyclic peptide -Asp-Phe-D-Pro-Asp- Phe-D-Pro-	2.20×10^2			
DCv1	H3 (1–21) ^b	2.11×10^5	Tris Buffer 20 mM, pH 7.4, 25 °C	ITC	[159]
	PKP ^c	2.68×10^5			
	MBP ^d	3.00×10^4			
VD1	N-Ac ₂ -L-Lys-D-Ala-D-Ala	1.02×10^5	Water		[161]
	N-Ac ₂ -L-Lys-D-Ala-D-Lac	3.19×10^4			
GP3KKF	(D-Glu) ₄	2.37×10^4	Bis Tris Buffer 1.5 mM pH 6.10		[164]
Pd6L4	Ac-Trp-Trp-Ala-NH ₂	$>1.00 \times 10^6$	Water, 20 °C	UV-Visible Spectroscopy	[179]
	Ac-Trp-Ala-Trp-NH ₂	2.50×10^5			
	Ac-Ala-Trp-Trp-NH ₂	2.10×10^4			
$Fe^{II}8L_6$ cage	Clavanin analogue ^e	1.11×10^5	Acetonitrile:water (1:1)		[180]
	Clavanin A ^{a,f}	1.25×10^7			

^a 1:2 host-guest stoichiometry. Overall binding constant in M^{-2} . The overall binding constant corresponds to the product of the stepwise binding constants for the 1:1 and 1:2 complexes, i.e., $K_{11} \cdot K_{12}$. ^b Ala-Arg-Thr-Lys-Gln-Thr-Ala-Arg-Lys-Ser-Thr-Gly-Gly-Lys-Ala-Pro-Arg-Lys-Gln-Leu-Ala. ^c Gly-Arg-Thr-Gly-Arg-Arg-Asn-Ser-Ile. ^d Ala-Pro-Arg-Thr-Pro-Gly-Gly-Arg-Arg. ^e Sequence of the analogue: Ser-Ser-Trp-Gly-His-Val-Gly-Lys-Tyr-Val-His-Gly-Trp-Ser-His-Val-Ser. ^f Val-Phe-Gln-Phe-Leu-Gly-Lys-Ile-Ile-His-His-Val-Gly-Asn-Phe-Val-His-Gly-Phe-Ser-His-Val-Phe.

5. Conclusions and Outlook

Despite the clear difficulty in the binding and recognition of the biological targets here presented, supramolecular systems begin to approach affinities and selectivities that are observable in nature, as is the case of protein–protein interactions, lock-and-key complementarity and the highly specific antigen–antibody recognition. Furthermore, many of the receptors explored here already possess the ability to directly modify or detect their target analytes, despite the majority of them presenting simple, modular structures. This is the case of *p*-sulfonatocalix[4]arene and cucurbit[n]urils, which can modulate the properties of proteins, such as cytochrome c and Casp-8, respectively. Additionally, some of these receptors present either an intrinsic colorimetric or fluorometric response in the presence of the respective ligands, e.g., the tweezers MPT and MCT, making them especially promising in sensing applications. Moreover, these characteristics are not the only ones that can be explored, seeing as the receptors can also be integrated in more complex structures, as synthetic binding pockets, with the addition of other groups for the creation of new functionalities.

The logical steps going forward are the continuous creation of new receptors, the eventual conjugation of several of these molecules for the multivalent binding of one complex analyte with several binding sights and smaller modifications and optimizations of the preexisting receptors, to enhance their selectivity and affinity towards one particular target ligand.

Funding: This research was funded by the Associate Laboratory for Green Chemistry—LAQV (UIDB/50006/2020), financed by FCT/MCTES.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: N.B. acknowledges the FCT/MCTES for a research contract CEECIND/00466/2017.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Camilli, A.; Bassler, B.L. Bacterial small-molecule signaling pathways. *Science* **2006**, *311*, 1113–1116. [[CrossRef](#)]
2. Sewald, N.; Jakubke, H. *Peptides: Chemistry and Biology*, 2nd ed.; Wiley-VCH: Weinheim, Germany, 2009; ISBN 9783527318674.
3. Maity, D.; Schmuck, C. *Synthetic Receptors for Biomolecules*; Smith, B.D., Ed.; Royal Society of Chemistry: London, UK, 2015; ISBN 978-1-84973-971-9.
4. Broza, Y.Y.; Zhou, X.; Yuan, M.; Qu, D.; Zheng, Y.; Vishinkin, R.; Khatib, M.; Wu, W.; Haick, H. Disease Detection with Molecular Biomarkers: From Chemistry of Body Fluids to Nature-Inspired Chemical Sensors. *Chem. Rev.* **2019**, *119*, 11761–11817. [[CrossRef](#)]
5. Kruse, U.; Bantscheff, M.; Drewes, G.; Hopf, C. Chemical and Pathway Proteomics. *Mol. Cell. Proteom.* **2008**, *7*, 1887–1901. [[CrossRef](#)]
6. Urban, G.; Piletsky, S.A.; Whitcombe, M.J. (Eds.) *Designing Receptors for the Next Generation of Biosensors*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 11, ISBN 978-3-642-28093-1.
7. Kubota, R.; Hamachi, I. Protein recognition using synthetic small-molecular binders toward optical protein sensing in vitro and in live cells. *Chem. Soc. Rev.* **2015**, *44*, 4454–4471. [[CrossRef](#)]
8. Peczuhi, M.W.; Hamilton, A.D. Peptide and Protein Recognition by Designed Molecules. *Chem. Rev.* **2000**, *100*, 2479–2494. [[CrossRef](#)]
9. Van Dun, S.; Ottmann, C.; Milroy, L.G.; Brunsved, L. Supramolecular Chemistry Targeting Proteins. *J. Am. Chem. Soc.* **2017**, *139*, 13960–13968. [[CrossRef](#)]
10. Park, H.S.; Lin, Q.; Hamilton, A.D. Protein Surface Recognition by Synthetic Receptors: A Route to Novel Submicromolar Inhibitors for α -Chymotrypsin. *J. Am. Chem. Soc.* **1999**, *121*, 8–13. [[CrossRef](#)]
11. Hatai, J.; Schmuck, C. Diverse Properties of Guanidinocarbonyl Pyrrole-Based Molecules: Artificial Analogues of Arginine. *Acc. Chem. Res.* **2019**, *52*, 1709–1720. [[CrossRef](#)]
12. Tabet, M.; Labroo, V.; Sheppard, P.; Sasaki, T. Spermine-Induced Conformational Changes of a Synthetic Peptide. *J. Am. Chem. Soc.* **1993**, *115*, 3866–3868. [[CrossRef](#)]
13. Tanaka, A.; Fujiyoshi, S.; Motomura, K.; Hayashida, O.; Hisaeda, Y.; Murakami, Y. Preparation and host-guest interactions of novel cage-type cyclophanes bearing chiral binding sites provided by dipeptide residues. *Tetrahedron* **1998**, *54*, 5187–5206. [[CrossRef](#)]
14. Yoon, S.S.; Still, W.C. An Exceptional Synthetic Receptor for Peptides. *J. Am. Chem. Soc.* **1993**, *115*, 823–824. [[CrossRef](#)]
15. Rick, W. *Methods of Enzymatic Analysis*, 2nd ed.; Bergmeyer, H.U., Ed.; Academic Press: Cambridge, MA, USA, 1974; Volume 2.
16. Whetstine, J.R. *Histone Methylation: Chemically Inert but Chromatin Dynamic*, 2nd ed.; Elsevier Inc.: Amsterdam, The Netherlands, 2010; Volume 3, ISBN 9780123741455.
17. Erce, M.A.; Abeygunawardena, D.; Low, J.K.K.; Hart-Smith, G.; Wilkins, M.R. Interactions affected by arginine methylation in the yeast protein-protein interaction network. *Mol. Cell. Proteom.* **2013**, *12*, 3184–3198. [[CrossRef](#)] [[PubMed](#)]
18. Gravley, B.R.; Maniatis, T. Arginine/Serine-Rich Domains of SR Proteins Can Function as Activators of Pre-mRNA Splicing splicing of all metazoan introns and therefore play a central role in the basic splicing reaction. One function of SR proteins is to promote the binding of U1 snRN. *Mol. Cell* **1998**, *1*, 765–771. [[CrossRef](#)]
19. Klumpp, L.M.; Mackey, A.T.; Farrell, C.M.; Rosenberg, J.M.; Gilbert, S.P. A kinesin switch I arginine to lysine mutation rescues microtubule function. *J. Biol. Chem.* **2003**, *278*, 39059–39067. [[CrossRef](#)]
20. Vazdar, M.; Heyda, J.; Mason, P.E.; Tesei, G.; Allolio, C.; Lund, M.; Jungwirth, P. Arginine “magic”: Guanidinium Like-Charge Ion Pairing from Aqueous Salts to Cell Penetrating Peptides. *Acc. Chem. Res.* **2018**, *51*, 1455–1464. [[CrossRef](#)]
21. Douteau-Guével, N.; Coleman, A.W.; Morel, J.P.; Morel-Desrosiers, N. Complexation of the basic amino acids lysine and arginine by three sulfonatocalix [n] arenes ($n = 4, 6$ and 8) in water: Microcalorimetric determination of the Gibbs energies, enthalpies and entropies of complexation. *J. Chem. Soc. Perkin Trans. 2* **1999**, *629*–633. [[CrossRef](#)]
22. Douteau-Guével, N.; Perret, F.; Coleman, A.W.; Morela, J.P.; Morel-Desrosiers, N. Binding of dipeptides and tripeptides containing lysine or arginine by p-sulfonatocalixarenes in water: NMR and microcalorimetric studies. *J. Chem. Soc. Perkin Trans. 2* **2002**, *524*–532. [[CrossRef](#)]

23. Lazar, A.; Da Silva, E.; Navaza, A.; Barbey, C.; Coleman, A.W. A new packing motif for para-sulfonatocalix[4]arene: The solid state structure of the para-sulfonatocalix[4]arene D-arginine complex. *Chem. Commun.* **2004**, *10*, 2162–2163. [CrossRef]
24. Fokkens, M.; Schrader, T.; Klärner, F.G. A molecular tweezer for lysine and arginine. *J. Am. Chem. Soc.* **2005**, *127*, 14415–14421. [CrossRef]
25. Schrader, T.; Bitan, G.; Klärner, F.G. Molecular tweezers for lysine and arginine—powerful inhibitors of pathologic protein aggregation. *Chem. Commun.* **2016**, *52*, 11318–11334. [CrossRef]
26. Li, C.; Ma, J.; Zhao, L.; Zhang, Y.; Yu, Y.; Shu, X.; Li, J.; Jia, X. Molecular selective binding of basic amino acids by a water-soluble pillar[5]arene. *Chem. Commun.* **2013**, *49*, 1924–1926. [CrossRef] [PubMed]
27. Liao, S.-M.; Du, Q.-S.; Meng, J.-Z.; Pang, Z.-W.; Huang, R.-B. The multiple roles of histidine in protein interactions. *Chem. Cent. J.* **2013**, *7*, 1–12. [CrossRef] [PubMed]
28. Da Silva, E.; Coleman, A.W. Synthesis and complexation properties towards amino acids of mono-substituted p-sulphonato-calix[n]-arenes. *Tetrahedron* **2003**, *59*, 7357–7364. [CrossRef]
29. Bogan, A.A.; Thorn, K.S. Anatomy of hot spots in protein interfaces. *J. Mol. Biol.* **1998**, *280*, 1–9. [CrossRef]
30. Cordes, M.; Köttgen, A.; Jasper, C.; Jacques, O.; Boudebous, H.; Giese, B. Influence of amino acid side chains on long-distance electron transfer in peptides: Electron hopping via “stepping stones”. *Angew. Chem. Int. Ed.* **2008**, *47*, 3461–3463. [CrossRef]
31. Giese, B.; Eckhardt, S.; Lauz, M. Electron Transfer in Peptides and Proteins. *Encycl. Radic. Chem. Biol. Mater.* **2012**. [CrossRef]
32. Li, Q.; Guo, Y.; Shao, S. A BODIPY based fluorescent chemosensor for Cu(II) ions and homocysteine/cysteine. *Sens. Actuators B Chem.* **2012**, *171–172*, 872–877. [CrossRef]
33. Chan, D.I.; Prenner, E.J.; Vogel, H.J. Tryptophan- and arginine-rich antimicrobial peptides: Structures and mechanisms of action. *Biochim. Biophys. Acta Biomembr.* **2006**, *1758*, 1184–1202. [CrossRef]
34. Lakowicz, J.R. *Principles of Fluorescence Spectroscopy*; Springer: New York, NY, USA, 2006.
35. Buschmann, H.J.; Schollmeyer, E.; Mutihac, L. The formation of amino acid and dipeptide complexes with α -cyclodextrin and cucurbit[6]uril in aqueous solutions studied by titration calorimetry. *Thermochim. Acta* **2003**, *399*, 203–208. [CrossRef]
36. Rekharsky, M.V.; Mori, T.; Yang, C.; Young, H.K.; Selvapalam, N.; Kim, H.; Sobransingh, D.; Kaifer, A.E.; Liu, S.; Isaacs, L.; et al. A synthetic host-guest system achieves avidin-biotin affinity by overcoming enthalpy—Entropy compensation. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 20737–20742. [CrossRef]
37. Bush, M.E.; Bouley, N.D.; Urbach, A.R. Charge-Mediated Recognition of N-Terminal Tryptophan in Aqueous Solution by a Synthetic Host. *J. Am. Chem. Soc.* **2005**, *127*, 14511–14517. [CrossRef]
38. Ling, Y.; Wang, W.; Kaifer, A.E. A new cucurbit[8]uril-based fluorescent receptor for indole derivatives. *Chem. Commun.* **2007**, *610–612*. [CrossRef]
39. Biedermann, F.; Rauwald, U.; Cziferszky, M.; Williams, K.A.; Gann, L.D.; Guo, B.Y.; Urbach, A.R.; Bielawski, C.W.; Scherman, O.A. Benzobis(imidazolium)-cucurbit[8]uril complexes for binding and sensing aromatic compounds in aqueous solution. *Chem. A Eur. J.* **2010**, *16*, 13716–13722. [CrossRef]
40. Lagona, J.; Wagner, B.D.; Isaacs, L. Molecular-recognition properties of a water-soluble cucurbit[6]uril analogue. *J. Org. Chem.* **2006**, *71*, 1181–1190. [CrossRef]
41. Stepniaik, P.; Lainer, B.; Chmurski, K.; Jurczak, J. The effect of urea moiety in amino acid binding by β -cyclodextrin derivatives: A 1000-fold increase in efficacy comparing to native β -cyclodextrin. *Carbohydr. Polym.* **2017**, *164*, 233–241. [CrossRef]
42. Cheng, L.; Zhang, H.; Dong, Y.; Zhao, Y.; Yu, Y.; Cao, L. Tetraphenylethene-based tetracationic cyclophanes and their selective recognition for amino acids and adenosine derivatives in water. *Chem. Commun.* **2019**, *55*, 2372–2375. [CrossRef]
43. Kyne, G.M.; Light, M.E.; Hursthouse, M.B.; De Mendoza, J.; Kilburn, J.D. Enantioselective amino acid recognition using acyclic thiourea receptors. *J. Chem. Soc. Perkin Trans. 1* **2001**, *1258–1263*. [CrossRef]
44. Urbach, A.R.; Ramalingam, V. Molecular recognition of amino acids, peptides, and proteins by cucurbit[n] uril receptors. *Isr. J. Chem.* **2011**, *51*, 664–678. [CrossRef]
45. Bailey, D.M.; Hennig, A.; Uzunova, V.D.; Nau, W.M. Supramolecular tandem enzyme assays for multiparameter sensor arrays and enantiomeric excess determination of amino acids. *Chem. A Eur. J.* **2008**, *14*, 6069–6077. [CrossRef]
46. Petroff, O.A.C. GABA and glutamate in the human brain. *Neuroscientist* **2002**, *8*, 562–573. [CrossRef]
47. Grabowska, A.; Nowicki, M.; Kwinta, J. Glutamate dehydrogenase of the germinating triticale seeds: Gene expression, activity distribution and kinetic characteristics. *Acta Physiol. Plant* **2011**, *33*, 1981–1990. [CrossRef]
48. Barmore, W.; Stone, W.L. Physiology, Urea Cycle. 2019. Available online: <https://europemc.org/article/NBK/NBK513323> (accessed on 20 July 2020).
49. Digerness, S.B.; Reddy, W.J. The malate-aspartate shuttle in heart mitochondria. *J. Mol. Cell. Cardiol.* **1976**, *8*, 779–785. [CrossRef]
50. Cox, M.; Nelson, D.L. *Lehninger Principles of Biochemistry*; W.H. Freeman & Company: New York, NY, USA, 2008; ISBN 0716743396.
51. Duan, Q.; Zhao, W.; Lu, K. Synthesis of a water-soluble pillar[6]arene dodecaamine and its selective binding of acidic amino acids in water. *Tetrahedron Lett.* **2017**, *58*, 4403–4406. [CrossRef]
52. Wang, K.; Cui, J.H.; Xing, S.Y.; Ren, X.W. Selective Recognition of Acidic Amino Acids in Water by Calixpyridinium. *Asian J. Org. Chem.* **2017**, *6*, 1385–1389. [CrossRef]
53. Alfonso, I.; Rebollo, F.; Gotor, V. Optically active dioxatetraazamacrocycles: Chemoenzymatic syntheses and applications in chiral anion recognition. *Chem. A Eur. J.* **2000**, *6*, 3331–3338. [CrossRef]
54. Alfonso, I.; Solà, J. Molecular Recognition of Zwitterions with Artificial Receptors. *Chem. Asian J.* **2020**, *15*, 986–994. [CrossRef]

55. Brockhausen, I.; Stanley, P. O-GalNAc Glycans. In *Essentials of Glycobiology*; Varki, A., Cummings, R.D., Esko, J.D., Stanley, P., Hart, G.W., Aebi, M., Darvill, A.G., Kinoshita, T., Packer, N.H., Prestegard, J.H., et al., Eds.; Cold Spring Harbor: New York, NY, USA, 2017.
56. Stanley, P.; Taniguchi, N.; Aebi, M. N-Glycans. In *Essentials of Glycobiology*; Varki, A., Cummings, R.D., Esko, J.D., Stanley, P., Hart, G.W., Aebi, M., Darvill, A.G., Kinoshita, T., Packer, N.H., Prestegard, J.H., et al., Eds.; Cold Spring Harbor: New York, NY, USA, 2017.
57. Guyton, A.C.; Hall, J.E. *Textbook of Medical Physiology*, 11th ed.; Guyton, A.C., Hall, J.E., Eds.; Elsevier Saunders: Philadelphia, PA, USA, 2006.
58. Mailoux, R.J.; Jin, X.; Willmore, W.G. Redox regulation of mitochondrial function with emphasis on cysteine oxidation reactions. *Redox Biol.* **2014**, *2*, 123–139. [[CrossRef](#)]
59. Schmuck, C.; Geiger, L. Efficient complexation of N-acetyl amino acid carboxylates in water by an artificial receptor: Unexpected cooperativity in the binding of glutamate but not aspartate. *J. Am. Chem. Soc.* **2005**, *127*, 10486–10487. [[CrossRef](#)]
60. Liu, Y.; Lv, X.; Hou, M.; Shi, Y.; Guo, W. Selective Fluorescence Detection of Cysteine over Homocysteine and Glutathione Based on a Cysteine-Triggered Dual Michael Addition/Retro-aza-aldo Cascade Reaction. *Anal. Chem.* **2015**, *87*, 11475–11483. [[CrossRef](#)]
61. Peck, E.M.; Bradley, D. Smith Applications of Synthetic Receptors for Biomolecules. In *Synthetic Receptors for Biomolecules*; Smith, B.D., Ed.; Royal Society of Chemistry: Cambridge, UK, 2015; pp. 1–34.
62. Pinalli, R.; Pedrini, A.; Dalcanale, E. Biochemical sensing with macrocyclic receptors. *Chem. Soc. Rev.* **2018**, *47*, 7006–7026. [[CrossRef](#)]
63. Shinkai, S.; Araki, K.; Matsuda, T.; Nishiyama, N.; Ikeda, H.; Takasu, I.; Iwamoto, M. NMR and crystallographic studies of a p-sulfonatocalix[4]arene-guest complex. *J. Am. Chem. Soc.* **1990**, *112*, 9053–9058. [[CrossRef](#)]
64. Shinkai, S.; Araki, K.; Manabe, O. NMR determination of association constants for calixarene complexes. Evidence for the formation of a 1:2 complex with calix[8]arene. *J. Am. Chem. Soc.* **1988**, *110*, 7214–7215. [[CrossRef](#)]
65. Shinkai, S.; Mori, S.; Koreishi, H.; Tsubaki, T.; Manabe, O. Hexasulfonated calix[6]arene derivatives: A new class of catalysts, surfactants, and host molecules. *J. Am. Chem. Soc.* **1986**, *108*, 2409–2416. [[CrossRef](#)] [[PubMed](#)]
66. Shinkai, S. Calixarenes as new functionalized host molecules. *Pure Appl. Chem.* **1986**, *58*, 1523–1528. [[CrossRef](#)]
67. Dalgarno, S.J.; Atwood, J.L.; Raston, C.L. Sulfonatocalixarenes: Molecular capsule and ‘Russian doll’ arrays to structures mimicking viral geometry. *Chem. Commun.* **2006**, *4567–4574*. [[CrossRef](#)]
68. Shinkai, S.; Mori, S.; Tsubaki, T.; Sone, T.; Manabe, O. New water-soluble host molecules derived from calix[6]arene. *Tetrahedron Lett.* **1984**, *25*, 5315–5318. [[CrossRef](#)]
69. Coleman, A.W.; Jebors, S.; Cecillon, S.; Perret, P.; Garin, D.; Marti-Battle, D.; Moulin, M. Toxicity and biodistribution of para-sulfonato-calix[4]arene in mice. *New J. Chem.* **2008**, *32*, 780. [[CrossRef](#)]
70. Perret, F.; Coleman, A.W. Biochemistry of anionic calix [n] arenes. *Chem. Commun.* **2011**, *47*, 7303. [[CrossRef](#)]
71. Perret, F.; Lazar, A.N.; Coleman, A.W. Biochemistry of the para-sulfonato-calix [n] arenes. *Chem. Commun.* **2006**, *2425–2438*. [[CrossRef](#)]
72. Hof, F. Host-guest chemistry that directly targets lysine methylation: Synthetic host molecules as alternatives to bio-reagents. *Chem. Commun.* **2016**, *52*, 10093–10108. [[CrossRef](#)]
73. Guo, D.-S.; Liu, Y. Supramolecular Chemistry of p-Sulfonatocalix [n] arenes and Its Biological Applications. *Acc. Chem. Res.* **2014**, *47*, 1925–1934. [[CrossRef](#)]
74. Ghale, G.; Nau, W.M. Dynamically Analyte-Responsive Macro cyclic Host-Fluorophore Systems. *Acc. Chem. Res.* **2014**, *47*, 2150–2159. [[CrossRef](#)]
75. Douteau-Guével, N.; Coleman, A.W.; Morel, J.; Morel-Desrosiers, N. Complexation of basic amino acids by water-soluble calixarene sulphonates as a study of the possible mechanisms of recognition of calixarene sulphonates by proteins. *J. Phys. Org. Chem.* **1998**, *11*, 693–696. [[CrossRef](#)]
76. Selkti, M.; Coleman, A.W.; Nicolis, I.; Douteau-Guével, N.; Villain, F.; Tomas, A.; Rango, C. The first example of a substrate spanning the calix[4]arene bilayer: The solid state complex of p-sulfonatocalix[4]arene with L-lysine. *Chem. Commun.* **1999**, *161–162*. [[CrossRef](#)]
77. Bakirci, H.; Koner, A.L.; Nau, W.M. Binding of inorganic cations by p-sulfonatocalix[4]arene monitored through competitive fluorophore displacement in aqueous solution. *Chem. Commun.* **2005**, *5411–5413*. [[CrossRef](#)]
78. Bakirci, H.; Koner, A.L.; Schwarzlose, T.; Nau, W.M. Analysis of Host-Assisted Guest Protonation Exemplified for p-Sulfonatocalix[4]arene—Towards Enzyme-Mimetic pKa Shifts. *Chem. Eur. J.* **2006**, *12*, 4799–4807. [[CrossRef](#)]
79. Francisco, V.; Piñeiro, A.; Nau, W.M.; García-Río, L. The “True” Affinities of Metal Cations to p-Sulfonatocalix[4]arene: A Thermodynamic Study at Neutral pH Reveals a Pitfall Due to Salt Effects in Microcalorimetry. *Chem. A Eur. J.* **2013**, *19*, 17809–17820. [[CrossRef](#)]
80. Garcia-Rio, L.; Basilio, N.; Francisco, V. Counterion effect on sulfonatocalix [n] arene recognition. *Pure Appl. Chem.* **2019**. [[CrossRef](#)]
81. Lavande, N.; Acuña, A.; Basilio, N.; Francisco, V.; Malkhede, D.D.; Garcia-Rio, L. A journey from calix[4]arene to calix[6]and calix[8]arene reveals more than a matter of size. Receptor concentration affects the stability and stoichiometric nature of the complexes. *Phys. Chem. Chem. Phys.* **2017**, *19*, 13640–13649. [[CrossRef](#)]
82. Pessêgo, M.; Basilio, N.; Muñiz, M.C.; García-Río, L. Competitive counterion complexation allows the true host: Guest binding constants from a single titration by ionic receptors. *Org. Biomol. Chem.* **2016**, *14*, 6442–6448. [[CrossRef](#)]

83. Francisco, V.; Basílio, N.; García-Río, L. Ionic Exchange in p-Sulfonatocalix[4]arene-Mediated Formation of Metal–Ligand Complexes. *J. Phys. Chem. B* **2014**, *118*, 4710–4716. [[CrossRef](#)]
84. Basilio, N.; García-Río, L.; Martín-Pastor, M. Counterion Binding in Solutions of p-Sulfonatocalix[4]arene. *J. Phys. Chem. B* **2010**, *114*, 7201–7206. [[CrossRef](#)]
85. Francisco, V.; Basilio, N.; García-Río, L. Counterion exchange as a decisive factor in the formation of host: Guest complexes by p-sulfonatocalix[4]arene. *J. Phys. Chem. B* **2012**, *116*, 5308–5315. [[CrossRef](#)]
86. Romero, M.A.; Mateus, P.; Matos, B.; Acuña, Á.; García-Río, L.; Arteaga, J.F.; Pischel, U.; Basílio, N. Binding of Flavylium Ions to Sulfonatocalix[4]arene and Implication in the Photorelease of Biologically Relevant Guests in Water. *J. Org. Chem.* **2019**, *84*, 10852–10859. [[CrossRef](#)]
87. Guo, D.S.; Uzunova, V.D.; Su, X.; Liu, Y.; Nau, W.M. Operational calixarene-based fluorescent sensing systems for choline and acetylcholine and their application to enzymatic reactions. *Chem. Sci.* **2011**, *2*, 1722–1734. [[CrossRef](#)]
88. Nau, W.M.; Ghale, G.; Hennig, A.; Bakirci, H.; Bailey, D.M. Substrate-selective supramolecular tandem assays: Monitoring enzyme inhibition of arginase and diamine oxidase by fluorescent dye displacement from calixarene and cucurbituril macrocycles. *J. Am. Chem. Soc.* **2009**, *131*, 11558–11570. [[CrossRef](#)]
89. Beshara, C.S.; Jones, C.E.; Daze, K.D.; Lilgert, B.J.; Hof, F. A simple calixarene recognizes post-translationally methylated lysine. *ChemBioChem* **2010**, *11*, 63–66. [[CrossRef](#)]
90. Daze, K.D.; Ma, M.C.F.; Pineux, F.; Hof, F. Synthesis of New Trisulfonated Calix[4]arenes Functionalized at the Upper Rim, and Their Complexation with the Trimethyllysine Epigenetic Mark. *Org. Lett.* **2012**, *14*, 1512–1515. [[CrossRef](#)]
91. Ogoshi, T.; Hashizume, M.; Yamagishi, T.A.; Nakamoto, Y. Synthesis, conformational and host-guest properties of water-soluble pillar[5]arene. *Chem. Commun.* **2010**, *46*, 3708–3710. [[CrossRef](#)]
92. Murray, J.; Kim, K.; Ogoshi, T.; Yao, W.; Gibb, B.C. The aqueous supramolecular chemistry of cucurbit [n] urils, pillar [n] arenes and deep-cavity cavitands. *Chem. Soc. Rev.* **2017**, *46*, 2479–2496. [[CrossRef](#)]
93. Ogoshi, T.; Yamagishi, T.; Nakamoto, Y. Pillar-Shaped Macroyclic Hosts Pillar [n] arenes: New Key Players for Supramolecular Chemistry. *Chem. Rev.* **2016**, *116*, 7937–8002. [[CrossRef](#)] [[PubMed](#)]
94. Bojtár, M.; Paudics, A.; Hessz, D.; Kubinyi, M.; Bitter, I. Amino acid recognition by fine tuning the association constants: Tailored naphthalimides in pillar[5]arene-based indicator displacement assays. *RSC Adv.* **2016**, *6*, 86269–86275. [[CrossRef](#)]
95. Bastug, E.; Kursunlu, A.N.; Guler, E. A fluorescent clever macrocycle: Deca-bodipy bearing a pillar[5]arene and its selective binding of asparagine in half-aqueous medium. *J. Lumin.* **2020**, *225*, 117343. [[CrossRef](#)]
96. Assaf, K.I.; Nau, W.M. Cucurbiturils: From synthesis to high-affinity binding and catalysis. *Chem. Soc. Rev.* **2015**, *44*, 394–418. [[CrossRef](#)] [[PubMed](#)]
97. Barrow, S.J.; Kasera, S.; Rowland, M.J.; del Barrio, J.; Scherman, O.A. Cucurbituril-Based Molecular Recognition. *Chem. Rev.* **2015**, *115*, 12320–12406. [[CrossRef](#)] [[PubMed](#)]
98. Masson, E.; Ling, X.; Joseph, R.; Kyeremeh-Mensah, L.; Lu, X. Cucurbituril chemistry: A tale of supramolecular success. *RSC Adv.* **2012**, *2*, 1213–1247. [[CrossRef](#)]
99. Deng, C.-L.; Murkli, S.L.; Isaacs, L.D. Supramolecular hosts as in vivo sequestration agents for pharmaceuticals and toxins. *Chem. Soc. Rev.* **2020**, *49*, 7516–7532. [[CrossRef](#)]
100. Shetty, D.; Khedkar, J.K.; Park, K.M.; Kim, K. Can we beat the biotin–avidin pair? Cucurbit[7]uril-based ultrahigh affinity host–guest complexes and their applications. *Chem. Soc. Rev.* **2015**, *44*, 8747–8761. [[CrossRef](#)]
101. Gamal-Eldin, M.A.; MacArtney, D.H. Selective molecular recognition of methylated lysines and arginines by cucurbit[6]uril and cucurbit[7]uril in aqueous solution. *Org. Biomol. Chem.* **2013**, *11*, 488–495. [[CrossRef](#)]
102. Lee, J.W.; Lee, H.H.L.; Ko, Y.H.; Kim, K.; Kim, H.I. Deciphering the Specific High-Affinity Binding of Cucurbit[7]uril to Amino Acids in Water. *J. Phys. Chem. B* **2015**, *119*, 4628–4636. [[CrossRef](#)]
103. Liu, S.; Ruspic, C.; Mukhopadhyay, P.; Chakrabarti, S.; Zavalij, P.Y.; Isaacs, L. The Cucurbit [n] uril Family: Prime Components for Self-Sorting Systems. *J. Am. Chem. Soc.* **2005**, *127*, 15959–15967. [[CrossRef](#)] [[PubMed](#)]
104. Romero, M.A.; Basílio, N.; Moro, A.J.; Domingues, M.; González-Delgado, J.A.; Arteaga, J.F.; Pischel, U. Photocaged Competitor Guests: A General Approach Toward Light-Activated Cargo Release From Cucurbiturils. *Chem. A Eur. J.* **2017**, *23*, 13105–13111. [[CrossRef](#)]
105. Biedermann, F.; Uzunova, V.D.; Scherman, O.A.; Nau, W.M.; De Simone, A. Release of High-Energy Water as an Essential Driving Force for the High-Affinity Binding of Cucurbit [n] urils. *J. Am. Chem. Soc.* **2012**, *134*, 15318–15323. [[CrossRef](#)]
106. Biedermann, F.; Vendruscolo, M.; Scherman, O.A.; De Simone, A.; Nau, W.M. Cucurbit[8]uril and blue-box: High-energy water release overwhelms electrostatic interactions. *J. Am. Chem. Soc.* **2013**, *135*, 14879–14888. [[CrossRef](#)]
107. Biedermann, F.; Nau, W.M.; Schneider, H.-J. The Hydrophobic Effect Revisited—Studies with Supramolecular Complexes Imply High-Energy Water as a Noncovalent Driving Force. *Angew. Chem. Int. Ed.* **2014**, *53*, 11158–11171. [[CrossRef](#)]
108. Heitmann, L.M.; Taylor, A.B.; Hart, P.J.; Urbach, A.R. Sequence-specific recognition and cooperative dimerization of N-terminal aromatic peptides in aqueous solution by a synthetic host. *J. Am. Chem. Soc.* **2006**, *128*, 12574–12581. [[CrossRef](#)]
109. Liu, Y.-H.; Zhang, Y.-M.; Yu, H.-J.; Liu, Y. Cucurbituril-Based Biomacromolecular Assemblies. *Angew. Chem. Int. Ed.* **2020**, *1*–11. [[CrossRef](#)]
110. Rajgariah, P.; Urbach, A.R. Scope of amino acid recognition by cucurbit[8]uril. *J. Incl. Phenom. Macrocycl. Chem.* **2008**, *62*, 251–254. [[CrossRef](#)]

111. Saenger, W.; Steiner, T. Cyclodextrin Inclusion Complexes: Host-Guest Interactions and Hydrogen-Bonding Networks. *Acta Crystallogr. Sect. A Found. Crystallogr.* **1998**, *54*, 798–805. [[CrossRef](#)]
112. Moran, J.R.; Karbach, S.; Cram, D.J. Cavitands: Synthetic Molecular Vessels. *J. Am. Chem. Soc.* **1982**, *104*, 5826–5828. [[CrossRef](#)]
113. Biavardi, E.; Tudisco, C.; Maffei, F.; Motta, A.; Massera, C.; Condorelli, G.G.; Dalcanale, E. Exclusive recognition of sarcosine in water and urine by a cavitand-functionalized silicon surface. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 2263–2268. [[CrossRef](#)]
114. Brancatelli, G.; Dalcanale, E.; Pinalli, R.; Geremia, S. Probing the structural determinants of amino acid recognition: X-ray studies of crystalline ditopic host-guest complexes of the positively charged amino acids, ARg, Lys, and His with a cavitand molecule. *Molecules* **2018**, *23*, 3368. [[CrossRef](#)]
115. Pinalli, R.; Brancatelli, G.; Pedrini, A.; Menozzi, D.; Hernández, D.; Ballester, P.; Geremia, S.; Dalcanale, E. The Origin of Selectivity in the Complexation of N-Methyl Amino Acids by Tetraphosphonate Cavitands. *J. Am. Chem. Soc.* **2016**, *138*, 8569–8580. [[CrossRef](#)]
116. Goodnow, T.T.; Reddington, M.V.; Stoddart, J.F.; Kaifer, A.E. Cyclobis(paraquat-p-phenylene): A Novel Synthetic Receptor for Amino Acids with Electron-Rich Aromatic Moieties. *J. Am. Chem. Soc.* **1991**, *113*, 4335–4337. [[CrossRef](#)]
117. Schmuck, C. Carboxylate Binding by 2-(Guanidinocarbonyl) pyrrole Receptors in Aqueous Solvents: Improving the Binding Properties of Guanidinium Cations through Additional Hydrogen Bonds. *Chem. A Eur. J.* **2000**, *6*, 709–718. [[CrossRef](#)]
118. Mandl, C.P.; König, B. Luminescent crown ether amino acids: Selective binding to N-terminal lysine in peptides. *J. Org. Chem.* **2005**, *70*, 670–674. [[CrossRef](#)]
119. Bekos, E.J.; Gardella, J.A.; Bright, F.V. The binding of free oligopeptides to cyclodextrins: The role of the tyrosine group. *J. Incl. Phenom. Mol. Recognit. Chem.* **1996**, *26*, 185–195. [[CrossRef](#)]
120. Stepniak, P.; Lainer, B.; Chmurski, K.; Jurczak, J. pH-Controlled recognition of amino acids by urea derivatives of β -cyclodextrin. *RSC Adv.* **2017**, *7*, 15742–15746. [[CrossRef](#)]
121. Barba-Bon, A.; Pan, Y.; Biedermann, F.; Guo, D.; Nau, W.M.; Hennig, A. Fluorescence Monitoring of Peptide Transport Pathways into Large and Giant Vesicles by Supramolecular Host–Dye Reporter Pairs. *J. Am. Chem. Soc.* **2019**, *141*, 20137–20145. [[CrossRef](#)]
122. Rekharsky, M.V.; Yamamura, H.; Ko, Y.H.; Selvapalam, N.; Kim, K.; Inoue, Y. Sequence recognition and self-sorting of a dipeptide by cucurbit[6]uril and cucurbit[7]uril. *Chem. Commun.* **2008**, 2236–2238. [[CrossRef](#)]
123. Hirani, Z.; Taylor, H.F.; Babcock, E.F.; Bockus, A.T.; Varnado, C.D.; Bielawski, C.W.; Urbach, A.R. Molecular Recognition of Methionine-Terminated Peptides by Cucurbit[8]uril. *J. Am. Chem. Soc.* **2018**, *140*, 12263–12269. [[CrossRef](#)]
124. Chinai, J.M.; Taylor, A.B.; Ryno, L.M.; Hargreaves, N.D.; Morris, C.A.; Hart, P.J.; Urbach, A.R. Molecular recognition of insulin by a synthetic receptor. *J. Am. Chem. Soc.* **2011**, *133*, 8810–8813. [[CrossRef](#)]
125. Smith, L.C.; Leach, D.G.; Blaylock, B.E.; Ali, O.A.; Urbach, A.R. Sequence-Specific, Nanomolar Peptide Binding via Cucurbit[8]uril-Induced Folding and Inclusion of Neighboring Side Chains. *J. Am. Chem. Soc.* **2015**, *137*, 3663–3669. [[CrossRef](#)]
126. Schneider, H.J. Binding mechanisms in supramolecular complexes. *Angew. Chem. Int. Ed.* **2009**, *48*, 3924–3977. [[CrossRef](#)]
127. Ghale, G.; Lanctôt, A.G.; Kreissl, H.T.; Jacob, M.H.; Weingart, H.; Winterhalter, M.; Nau, W.M. Chemosensing ensembles for monitoring biomembrane transport in real time. *Angew. Chem. Int. Ed.* **2014**, *53*, 2762–2765. [[CrossRef](#)]
128. Balhorn, R. The protamine family of sperm nuclear proteins. *Genome Biol.* **2007**, *8*. [[CrossRef](#)]
129. Voigt, A.; Jelinek, H.F. Humanin: A mitochondrial signaling peptide as a biomarker for impaired fasting glucose-related oxidative stress. *Physiol. Rep.* **2016**, *4*, 1–5. [[CrossRef](#)]
130. Memmi, L.; Lazar, A.; Brioude, A.; Ball, V.; Coleman, A.W. Protein-calixarene interactions: Complexation of Bovine Serum Albumin by sulfonatocalix [n] arenes. *Chem. Commun.* **2001**, *2*, 2474–2475. [[CrossRef](#)]
131. McGovern, R.E.; Fernandes, H.; Khan, A.R.; Power, N.P.; Crowley, P.B. Protein camouflage in cytochrome c-calixarene complexes. *Nat. Chem.* **2012**, *4*, 527–533. [[CrossRef](#)]
132. McGovern, R.E.; McCarthy, A.A.; Crowley, P.B. Protein assembly mediated by sulfonatocalix[4]arene. *Chem. Commun.* **2014**, *50*, 10412–10415. [[CrossRef](#)]
133. Rennie, M.L.; Fox, G.C.; Pérez, J.; Crowley, P.B. Auto-regulated Protein Assembly on a Supramolecular Scaffold. *Angew. Chem. Int. Ed.* **2018**, *57*, 13764–13769. [[CrossRef](#)]
134. Rennie, M.L.; Crowley, P.B. A Thermodynamic Model of Auto-regulated Protein Assembly by a Supramolecular Scaffold. *ChemPhysChem* **2019**, *20*, 1011–1017. [[CrossRef](#)]
135. Bhasikuttan, A.C.; Mohanty, J. Detection, inhibition and disintegration of amyloid fibrils: The role of optical probes and macrocyclic receptors. *Chem. Commun.* **2017**, *53*, 2789–2809. [[CrossRef](#)]
136. Nguyen, B.T.; Anslyn, E.V. Indicator—Displacement assays. *Coord. Chem. Rev.* **2006**, *250*, 3118–3127. [[CrossRef](#)]
137. Cavatorta, E.; Jonkheijm, P.; Huskens, J. Assessment of Cooperativity in Ternary Peptide-Cucurbit[8]uril Complexes. *Chem. A Eur. J.* **2017**, *23*, 4046–4050. [[CrossRef](#)]
138. Romero, M.A.; Fernandes, R.J.; Moro, A.J.; Basílio, N.; Pischel, U. Light-induced cargo release from a cucurbit[8]uril host by means of a sequential logic operation. *Chem. Commun.* **2018**, *54*, 13335–13338. [[CrossRef](#)]
139. Giglione, C.; Boularot, A.; Meinnel, T. Protein N-terminal methionine excision. *Cell. Mol. Life Sci.* **2004**, *61*, 1455–1474. [[CrossRef](#)]
140. Wu, G.; Clarke, D.E.; Wu, C.; Scherman, O.A. Oligopeptide-CB[8]complexation with switchable binding pathways. *Org. Biomol. Chem.* **2019**. [[CrossRef](#)]
141. Lee, J.W.; Park, M.H.; Ju, J.T.; Choi, Y.S.; Hwang, S.M.; Jung, D.J.; Kim, H.I. Stabilization of compact protein structures by macrocyclic hosts cucurbit [n] urils in the gas phase. *Mass Spectrom. Lett.* **2016**, *7*, 16–20. [[CrossRef](#)]

142. De Vink, P.J.; Briels, J.M.; Schrader, T.; Milroy, L.-G.; Brunsved, L.; Ottmann, C. A Binary Bivalent Supramolecular Assembly Platform Based on Cucurbit[8]uril and Dimeric Adapter Protein 14-3-3. *Angew. Chem. Int. Ed.* **2017**, *56*, 8998–9002. [CrossRef]
143. Lee, H.H.; Choi, T.S.; Lee, S.J.C.; Lee, J.W.; Park, J.; Ko, Y.H.; Kim, W.J.; Kim, K.; Kim, H.I. Supramolecular inhibition of amyloid fibrillation by cucurbit[7]uril. *Angew. Chem. Int. Ed.* **2014**, *53*, 7461–7465. [CrossRef]
144. Webber, M.J.; Appel, E.A.; Vinciguerra, B.; Cortinas, A.B.; Thapa, L.S.; Jhunjhunwala, S.; Isaacs, L.; Langer, R.; Anderson, D.G. Supramolecular PEGylation of biopharmaceuticals. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 14189–14194. [CrossRef]
145. Maikawa, C.L.; Smith, A.A.A.; Zou, L.; Roth, G.A.; Gale, E.C.; Stapleton, L.M.; Baker, S.W.; Mann, J.L.; Yu, A.C.; Correa, S.; et al. A co-formulation of supramolecularly stabilized insulin and pramlintide enhances mealtime glucagon suppression in diabetic pigs. *Nat. Biomed. Eng.* **2020**, *4*, 507–517. [CrossRef]
146. Wang, R.; Qiao, S.; Zhao, L.; Hou, C.; Li, X.; Liu, Y.; Luo, Q.; Xu, J.; Li, H.; Liu, J. Dynamic protein self-assembly driven by host–guest chemistry and the folding–unfolding feature of a mutually exclusive protein. *Chem. Commun.* **2017**, *53*, 10532–10535. [CrossRef]
147. Hou, C.; Huang, Z.; Fang, Y.; Liu, J. Construction of protein assemblies by host–guest interactions with cucurbiturils. *Org. Biomol. Chem.* **2017**, *15*, 4272–4281. [CrossRef]
148. Hou, C.; Zeng, X.; Gao, Y.; Qiao, S.; Zhang, X.; Xu, J.; Liu, J. Cucurbituril As A Versatile Tool to Tune the Functions of Proteins. *Isr. J. Chem.* **2017**, *215163*, 1–11. [CrossRef]
149. Dang, D.T.; Van Onzen, A.H.A.M.; Dorland, Y.L.; Brunsved, L. Reactivation of an Inactivated Caspase-8 Mutant Reveals Differentiated Enzymatic Substrate Processing. *ChemBioChem* **2020**, *19*. [CrossRef]
150. Guagnini, F.; Antonik, P.M.; Rennie, M.L.; O’Byrne, P.; Khan, A.R.; Pinalli, R.; Dalcanale, E.; Crowley, P.B. Cucurbit[7]uril-Dimethyllysine Recognition in a Model Protein. *Angew. Chem. Int. Ed.* **2018**, *57*, 7126–7130. [CrossRef]
151. Breslow, R.; Yang, Z.; Ching, R.; Trojandt, G.; Odobel, F. Sequence Selective Binding of Peptides by Artificial Receptors in Aqueous Solution. *J. Am. Chem. Soc.* **1998**, *120*, 3536–3537. [CrossRef]
152. Fukuhara, G.; Inoue, Y. Peptide chirality sensing by a cyclodextrin-polythiophene conjugate. *Chem. A Eur. J.* **2012**, *18*, 11459–11464. [CrossRef]
153. Weissenstein, A.; Saha-Möller, C.R.; Würthner, F. Optical Sensing of Aromatic Amino Acids and Dipeptides by a Crown-Ether-Functionalized Perylene Bisimide Fluorophore. *Chem. A Eur. J.* **2018**, *24*, 8009–8016. [CrossRef] [PubMed]
154. Hossain, M.A.; Schneider, H.J. Sequence-selective evaluation of peptide side-chain interaction. New artificial receptors for selective recognition in water. *J. Am. Chem. Soc.* **1998**, *120*, 11208–11209. [CrossRef]
155. Liu, Y.; Liao, P.; Cheng, Q.; Hooley, R.J. Protein and small molecule recognition properties of deep cavitands in a supported lipid membrane determined by calcination-enhanced SPR spectroscopy. *J. Am. Chem. Soc.* **2010**, *132*, 10383–10390. [CrossRef] [PubMed]
156. Ghang, Y.J.; Lloyd, J.J.; Moehlig, M.P.; Arguelles, J.K.; Mettry, M.; Zhang, X.; Julian, R.R.; Cheng, Q.; Hooley, R.J. Labeled protein recognition at a membrane bilayer interface by embedded synthetic receptors. *Langmuir* **2014**, *30*, 10161–10166. [CrossRef]
157. Bontempi, N.; Biavardi, E.; Bordiga, D.; Candiani, G.; Alessandri, I.; Bergese, P.; Dalcanale, E. Probing lysine mono-methylation in histone H3 tail peptides with an abiotic receptor coupled to a non-plasmonic resonator. *Nanoscale* **2017**, *9*, 8639–8646. [CrossRef]
158. Liu, Y.; Perez, L.; Mettry, M.; Easley, C.J.; Hooley, R.J.; Zhong, W. Self-Aggregating Deep Cavitand Acts as a Fluorescence Displacement Sensor for Lysine Methylation. *J. Am. Chem. Soc.* **2016**, *138*, 10746–10749. [CrossRef]
159. Liu, Y.; Perez, L.; Gill, A.D.; Mettry, M.; Li, L.; Wang, Y.; Hooley, R.J.; Zhong, W. Site-Selective Sensing of Histone Methylation Enzyme Activity via an Arrayed Supramolecular Tandem Assay. *J. Am. Chem. Soc.* **2017**, *139*, 10964–10967. [CrossRef]
160. Biros, S.M.; Ullrich, E.C.; Hof, F.; Trembleau, L.; Rebek, J. Kinetically Stable Complexes in Water: The Role of Hydration and Hydrophobicity. *J. Am. Chem. Soc.* **2004**, *126*, 2870–2876. [CrossRef]
161. Xu, R.; Greiveldinger, G.; Marenus, L.E.; Cooper, A.; Ellman, J.A. Combinatorial library approach for the identification of synthetic receptors targeting vancomycin-resistant bacteria[4]. *J. Am. Chem. Soc.* **1999**, *121*, 4898–4899. [CrossRef]
162. Chen, G.; Zheng, S.; Luo, X.; Shen, J.; Zhu, W.; Liu, H.; Gui, C.; Zhang, J.; Zheng, M.; Chum, M.P.; et al. Focused combinatorial library design based on structural diversity, druglikeness and binding affinity score. *J. Comb. Chem.* **2005**, *7*, 398–406. [CrossRef]
163. Schmuck, C.; Heil, M. One-armed artificial receptors for the binding of polar tetrapeptides in water: Probing the substrate selectivity of a combinatorial receptor library. *Chem. A Eur. J.* **2006**, *12*, 1339–1348. [CrossRef] [PubMed]
164. Schmuck, C.; Wich, P. Sequence-dependent stereoselectivity in the binding of tetrapeptides in water by a flexible artificial receptor. *Angew. Chem. Int. Ed.* **2006**, *45*, 4277–4281. [CrossRef] [PubMed]
165. Whitlock, H.W., Jr.; Chen, C.-W. Molecular Tweezers: A Simple Model of Bifunctional Intercalation. *J. Am. Chem. Soc.* **1978**, *100*, 4921–4922. [CrossRef]
166. Leblond, J.; Petitjean, A. Molecular Tweezers: Concepts and Applications. *ChemPhysChem* **2011**, *12*, 1043–1051. [CrossRef] [PubMed]
167. Klärner, F.-G.; Schrader, T. Aromatic Interactions by Molecular Tweezers and Clips in Chemical and Biological Systems. *Acc. Chem. Res.* **2013**, *46*, 967–978. [CrossRef] [PubMed]
168. Dutt, S.; Wilch, C.; Gersthagen, T.; Talbiersky, P.; Bravo-Rodriguez, K.; Hanni, M.; Sánchez-García, E.; Ochsenfeld, C.; Klärner, F.-G.; Schrader, T. Molecular Tweezers with Varying Anions: A Comparative Study. *J. Org. Chem.* **2013**, *78*, 6721–6734. [CrossRef]
169. Lump, E.; Castellano, L.M.; Meier, C.; Seeliger, J.; Erwin, N.; Sperlich, B.; Stürzel, C.M.; Usmani, S.; Hammond, R.M.; Von Einem, J.; et al. A molecular tweezer antagonizes seminal amyloids and HIV infection. *Elife* **2015**, *4*, 1–33. [CrossRef]

170. Röcker, A.E.; Müller, J.A.; Dietzel, E.; Harms, M.; Krüger, F.; Heid, C.; Sowislok, A.; Riber, C.F.; Kupke, A.; Lippold, S.; et al. The molecular tweezer CLR01 inhibits Ebola and Zika virus infection. *Antivir. Res.* **2018**, *152*, 26–35. [[CrossRef](#)]
171. Weil, T.; Groß, R.; Röcker, A.; Bravo-Rodriguez, K.; Heid, C.; Sowislok, A.; Le, M.-H.; Erwin, N.; Dwivedi, M.; Bart, S.; et al. Supramolecular Mechanism of Viral Envelope Disruption by Molecular Tweezers. *J. Am. Chem. Soc.* **2020**. [[CrossRef](#)]
172. Schmuck, C.; Kuchelmeister, H. An efficient synthesis of an orthogonally protected aromatic diamine as scaffold for tweezer receptors with two different arms. *Eur. J. Org. Chem.* **2009**, 4480–4485.
173. Löwik, D.; Weingarten, D.; Broekema, M.; Brouwer, A.; Still, W.; Liskamp, R. Tweezers with different bite: Increasing the affinity of synthetic receptors by varying the hinge part. *Angew. Chem. Int. Ed.* **1998**, *37*, 1846–1850. [[CrossRef](#)]
174. Monnee, M.C.F.; Brouwer, A.J.; Liskamp, R.M.J. Synthesis, screening and evaluation of a combined library of tweezer- and tripodal synthetic receptors. *QSAR Comb. Sci.* **2004**, *23*, 546–559. [[CrossRef](#)]
175. Zarra, S.; Wood, D.M.; Roberts, D.A.; Nitschke, J.R. Molecular containers in complex chemical systems. *Chem. Soc. Rev.* **2015**, *44*, 419–432. [[CrossRef](#)]
176. Brown, C.J.; Toste, F.D.; Bergman, R.G.; Raymond, K.N. Supramolecular Catalysis in Metal–Ligand Cluster Hosts. *Chem. Rev.* **2015**, *115*, 3012–3035. [[CrossRef](#)]
177. Chakrabarty, R.; Mukherjee, P.S.; Stang, P.J. Supramolecular Coordination: Self-Assembly of Finite Two- and Three-Dimensional Ensembles. *Chem. Rev.* **2011**, *111*, 6810–6918. [[CrossRef](#)]
178. Fujita, M.; Tominaga, M.; Hori, A.; Therrien, B. Coordination Assemblies from a Pd(II)-Cornered Square Complex. *Acc. Chem. Res.* **2005**, *38*, 369–378. [[CrossRef](#)]
179. Tashiro, S.; Tominaga, M.; Kawano, M.; Therrien, B.; Ozeki, T.; Fujita, M. Sequence-Selective Recognition of Peptides within the Single Binding Pocket of a Self-Assembled Coordination Cage. *J. Am. Chem. Soc.* **2005**, *127*, 4546–4547. [[CrossRef](#)]
180. Mosquera, J.; Szyszko, B.; Ho, S.K.Y.; Nitschke, J.R. Sequence-selective encapsulation and protection of long peptides by a self-assembled Fe II 8 L 6 cubic cage. *Nat. Commun.* **2017**, *8*, 6–11. [[CrossRef](#)]
181. Reczek, J.J.; Kennedy, A.A.; Halbert, B.T.; Urbach, A.R. Multivalent recognition of peptides by modular self-assembled receptors. *J. Am. Chem. Soc.* **2009**, *131*, 2408–2415. [[CrossRef](#)]