Molecular recognition of zwitterions with artificial receptors

Ignacio Alfonso* and Jordi Solà*

Dedicated to Prof. Carsten Schmuck. In memoriam.

[a] Dr. I. Alfonso, Dr. J. Solà.
Department of Biological Chemistry
Institute for Advanced Chemistry of Catalonia
Jordi Girona 18-26, 08034, Barcelona. Spain.
E-mail: ignacio.alfonso@iqac.csic.es, jordi.sola@iqac.csic.es

Abstract: Many biomolecules exist as internal ion pairs or zwitterions within a biologically relevant pH range. Despite their importance, the molecular recognition of this type of systems is specially challenging due to their strong solvation in aqueous media, and their trend to form folded or self-assembled structures by pairing of charges of different sign. In this mini-review, we will discuss the molecular recognition of zwitterions using non-natural, synthetic receptors. This contribution does not intend to make a full in-depth revision of the existing research in the field, but a personal overview with selected representative examples from the recent literature.

1. Introduction

Zwitterions are important chemical entities that are ubiquitous in biological processes. For example, amino acids, most of them zwitterionic at biological pH, are present in all living organisms playing fundamental roles in biochemical processes such as signal transmission or cellular structure building. Accordingly, the binding and sensing of zwitterions by artificial receptors has long attracted chemists' attention. Unfortunately, the design of artificial receptors to bind zwitterions present several challenges. As these guests are strongly solvated species their recognition and encapsulation in aqueous or protic solvents has to pay a strong energetic cost. Moreover, in order to recognize both a cationic and an anionic moiety, a sophisticated design of the receptor is necessary involving multiple recognition motifs. Thus, ditopic receptors with two different binding sites have been proposed for the interaction with the positively and negatively charged groups of zwitterions. This design, however, presents some drawbacks: both sites display complementary features and these fragments can interact with each other giving rise to intramolecularly folded or cyclic, dimeric or oligomeric species that will prevent the association with the intended guests. These processes were extensively studied by Schmuck in a series of zwitterionic compounds based on guanidiniocarbonyl pyrroles.^[1] Taking these considerations into account, in addition to optimize each of the binding sites for its planned function, chemists must pay attention to the fragment or linker between the two complementary sites of the receptor, ensuring a proper distance and relative orientation between them in order to favor the interaction with the guests and to disfavor all the competing pathways. Several binding motifs have been used to the purpose of zwitterionic recognition. In this minireview, we will highlight some of these motifs classified by the structure of the receptor and paying special attention to the progress made in the last years.

2. Hosts based on calixarenes and related compounds.

The macrocyclic structure of calixarenes consisting of a polyaromatic pocket with a conical shape is particularly well suited for the interaction with cationic and/or hydrophobic moieties leading to an inclusion complex.^[2] Since amino acids have an amino function, the ability of several types of calixarenes to bind amino acids and their derivatives has extensively been studied.^[3] Functionalization of the rims can provide additional interactions with the functional groups of the guests providing a particularly well suited scaffold for zwitterion recognition. This was well demonstrated by Rebek in his seminal work with amine-functionalized deep cavitands (compound **1**, figure 1) which provided a binding constant with carnitine in DMSO in the order of $15 \cdot 10^3 \text{ M}^{-1}$ (measured by NMR at 298 K).^[4] According to the NMR data and molecular modelling studies, an inclusion complex with the tetraalkylammonium group inside the cavity is formed. Interestingly, when the carboxylic acid of carnitine is protonated the binding decreases indicating an interaction between the rim and the negatively charged moiety. More recently, the same group reported a bifunctional cavitand incorporating a Zn-salen complex which showed a synergetic binding towards phosphocholine esters.^[5]

In 2006, Ballester reported a hybrid cavitand-resorcin[4]arene receptor (**1b**, figure 1) with a pH-modulated binding affinity towards trimethylalkylammonium ions, able to work as a supramolecular sensor for choline.^[6] The binding affinities towards different cations were established in methanol as a solvent by a competitive displacement assay, using a pyrene-modified N-methylpyridinium cation as a fluorescent indicator and gave a binding constant of $3.3 \pm 0.8 \cdot 10^5 \text{ M}^{-1}$ for choline. Later, Inoue and Mattay^[7] reported the binding of carnitine, betaine and their hydrochlorides by a pyrogallol[4]arene in ethanolic solutions obtaining K_a ranging from 1.5 to $6.1 \cdot 10^3 \text{ M}^{-1}$ for most substrates (compound **2a**, figure 1B). The cage, however, was particularly well suited for the encapsulation of L-carnitine which gave a binding affinity 3-12 times larger than other guests ($18 \cdot \pm 1 \cdot 10^3 \text{ M}^{-1}$). Interestingly, the X-ray structure of the complex formed by L-carnitine and resorcin[4]arene (compound **2b**, figure 1C) was published by Aoki,^[8] and shows the presence of cation- π interactions between the quaternary trimethylammonium cation and the aromatic rings of the macrocyclic receptor which mimics the carnitine binding sites in acyltransferases or a carnitine-transporter.^[9]

[a] Dr. I. Alfonso, Dr. J. Solà.
Department of Biological Chemistry
Institute for Advanced Chemistry of Catalonia
Jordi Girona 18-26, 08034, Barcelona. Spain.
E-mail: ignacio.alfonso@iqac.csic.es, jordi.sola@iqac.csic.es

Figure 1. A) Molecular receptor for carnitine and choline synthesised by Rebek **1a** and hybrid cavitand-r resorcin[4]arene **1b** described by Ballester. B) Pyrogallol[4]arene **2a** reported by Mattay. C) Crystal structure of tetramethylated resorcin[4]arene **2b** and carnitine complex described by Aoki. Solvent molecules and other receptor/guests presents in the crystallographic cell are omitted for clarity

More recently, Jabin and co-workers described two homologous bis-calixarenes that behave as a heteroditopic receptor and can bind mainly cationic or zwitterionic guests (compounds **3-5**, figure 2). The D_{3h} symmetrical bis-calix[6]arenes **3** were constructed using tri-urea or thiourea bridges and both calixarene moieties adopt a flattened cone conformation. Thus, quaternary ammonium groups are especially suited for this receptors through the establishment of the typical π -cationic and CH- π interactions into the cavity. Thiourea **3a** can bind 1,2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPC) or dodecylphosphocholine (DPC) in chloroform ($K_a > 2.8 \times 10^5 \text{ M}^{-1}$ and $K_a > 1.8 \times 10^5 \text{ M}^{-1}$ respectively). Remarkably, in the presence of methanol (1:50 CD₃OD:CDCl₃) a selective binding of DOPC over DPC is observed. NMR data and theoretical calculations confirmed the presence of the ammonium moieties inside the calixarene while thiourea groups establish H-bonding interactions with the phosphate group.^[10] Zwitterionic guests such as trimethylglycine, β -alanine betaine, or deoxycarnitine were shown to bind better with the related urea cage **3b**. The interaction pattern is similar as the one described above with the quaternary ammonium group deeply buried inside one of the calixarene units while carboxylate group of the guest is then bound to two tri-urea arms through three H-bonding interactions.^[11] In comparison, calix[6]trisureas **4** present binding affinities 2 orders of magnitude lower than the corresponding bis-calix[6]arene. A compound related to **4** but containing pyrenil units as aromatic moieties demonstrated chemosensing capacities of as it is able to selectively interact with phosphatidylcholine-type lipids in organic media.^{11d} Interestingly, the related calix[6]-cryptothiureas **5** do not bind DOPC but they can encapsulate different zwitterions in a protic environment (CDCl₃:CD₃OD 1:1); **5a** displays a remarkable selectivity for β -alanine betaine thanks to a high complementarity between host and guest.^[12]

Figure 2. Bis-calix[6]arenes (**3**), calix[6]cryptothiureas (**4**), and calix[6]trisureas (**5**) developed by Jabin.

3. Crown ethers and polyaza macrocycles.

Crown ethers efficiently interact with cations and, depending on the structure, may interact selectively with ammonium ions, particularly with protonated primary amines. Therefore, receptors containing a crown ether fragment plus an additional site for anion recognition have been a useful platform to construct receptors for zwitterionic species.^[13]

Ignacio Alfonso studied Chemistry at the University of Oviedo (Spain), obtaining his PhD at the same University in 1999. After two post-doctoral studies at The Scripps Research Institute in La Jolla (USA, 2000–2002) and back to the University of Oviedo (2002–2004), he obtained the Ramón y Cajal contract and moved to Universitat Jaume I (Castellón). Since 2007 he has a permanent position at IQAC-CSIC in Barcelona, where he leads the Supramolecular Chemistry group. His research interests are all fields related to Supramolecular Chemistry, especially the molecular recognition of biomolecules and the study of dynamic complex systems..

Jordi Solà received a PhD in Chemistry from the University of Barcelona in 2006. After, he moved to U.K. where he performed postdoctoral studies with Prof. J. Clayden (University of Manchester) and later with Prof. D. A. Leigh (University of Edinburgh). In 2012 he received Ramon y Cajal Fellowship and joined the Institute for Advanced Chemistry of Catalonia as independent researcher. Since 2017 he is Científico Titular at the same institute belonging to the CSIC. His research interest interests are Supramolecular Chemistry and Dynamic Chemical Systems

Several functional groups able to interact with the anionic moiety of switterions have been used in order to synthesize ditopic receptors. Among them, the ammonium group the simplest

one.^[14] Importantly, receptors based on primary ammonium groups are pH sensitive and intramolecular interactions between this group and the crown ether will compete with the binding of the guest; thus quaternary ammonium groups are best suited. An early example is compound **6** in figure 3 developed by Schmidtchen which works as a receptor for ω -amino acids.^[15] A similar design using a crown ether and ammonium group separated by a linker was studied by Schneider in the recognition of tripeptides in water.^[16] Charged guanidinium (compound **7**, figure 3)^[17] or heteroaromatic cations^[18] have also been employed targeting single amino acids or small peptides. Moreover, the combination of crown ethers and cyclen macrocycle were studied by Delgado, these receptors show a high dependency towards pH and show moderate binding constants in water at neutral pH, mainly, because of the competition of the solvent with the recognition units (compound **8**, figure 3).^[19]

Figure 3. Ditopic receptors containing crown ethers developed by Schmidtchen (**6**), König (**7**) and Delgado (**8**).

Neutral ditopic receptors containing crown ethers have also been described. Phenylboronic acid group was used by Reetz in combination with a crown ether to act as transporters of amino acids through organic solvents.^[20] The group of Costero described ligands containing crown ethers as cation binding sites and thiourea or amide groups for anion recognition that allow the solid-liquid extraction of ω -amino acids into DMSO solutions.^[21] More recently, the same group has prepared a fluorescent ditopic receptor composed of an 18-azacrown-6 moiety as a cation binder and a meso-octamethylcalix[4]pyrrole as a ligand for the anionic part (**9** in figure 4). The two fragments are connected through a boron–dipyrrromethene (BODIPY) spacer which also acts as a fluorophore. The compound can act as ditopic sensor for zwitterionic amino acids and analyzing the response in the emission/absorption optical properties allows the discrimination of several species.^[22] The distance between recognition sites is adequate for complexation with GABA but for example too long for the recognition of glycine. A different design combining aza-crown ethers with hydrogen bond donors such as xanthone or chromane has been used by Moran and Muñiz for selective amino acid extraction from water (**10** in figure 4). While the first is selective for phenylalanine the latter is selective for alanine in front other amino acids.^[23] Finally, the related cyclic polyamines containing pyrazole units **11** have been studied by the group of García-España. It has been proved that these compounds act as ditopic or polytopic receptors and they are able to interact in water at physiological pH with dopamine ($\log K = 3$ -6) or glutamate ($K_{\text{eff}} = 2.04 \cdot 10^4 \text{ M}^{-1}$).^[24] The interaction with dopamine is stronger using the copper complex of the ligand (Cu₂L) as the hydroxo groups of dopamine deprotonate to bind the Cu₂⁺ ions at pH around 8.

Figure 4. Neutral ditopic receptors containing crown ethers developed by Costero (**9**), Muñiz (**10**) and pyrazol-containing cyclic polyamines described by García-España (**11**).

4. Cyclodextrins.

Native unmodified cyclodextrins show low ability to recognize zwitterionic species like aminoacids.^[25] Thus, α -Cyclodextrin shows complexes with K_a in the order of 250 M⁻¹ (for Gly) to 1600 M⁻¹ (for Val) in water while β -cyclodextrin binds Trp with a K_a of 13 M⁻¹. In both cases the process is mainly driven by entropic change. Modification of cyclodextrins with carboxylic and ammonium functionalities leads to very modest improvements. Therefore, the introduction of additional hydrogen bond donors or acceptors into the CD rims has been studied including different functional groups such as esters, sulfonic acids or sulfonylamines.^[26] Recently, Jurczack described the modification of β -cyclodextrin with a urea moiety that provides an extra hydrogen bond donor and increases the binding by a 1000-fold ratio (compound **12**, figure 5). Interestingly, the presence of an isopropyl functionality allows the formation of ‘active aggregates’ which allow the urea functionality to establish interactions isolated from the bulk solvent.^[27] Based on this design, the same group studied a series of compounds and found binding affinities for non-polar amino acids from 2300 M⁻¹ for alanine to 54 800 M⁻¹ for tryptophan. Importantly, carboxylates lacking the amino group remained not bound.^[28] The binding strength showed to be highly pH dependent with good affinities at pH 8 but almost no binding at pH 6.

Figure 5. Urea modified β -cyclodextrins described by Jurczack (**12**) for the recognition of amino acids and bis- β -cyclodextrins developed in the early 90's by Breslow (**13**).

The interaction of CDs with zwitterionic molecules can be enhanced by the formation of metal complexes. Marchelli and Rizzanelli prepared a histidine-modified CD; thus the corresponding Cu(II) complex can form ternary complexes with amino acids in water. Enantioselective discrimination (with D-aa showing better binding) is achieved thanks to the chiral environment of the cavity.^[29] The introduction of a dansyl fluorophore linked through a Phe bridge (D or L) results in different

behavior as according to NMR, ESI-MS and time-resolved fluorescence measurements. The DNS chromophore is buried inside the cavity when the host contains L-Phe but not in the case of the one containing D-Phe. Thus, the ability of the corresponding copper complexes to discriminate between enantiomeric amino acids is greatly affected and the chiral discrimination is better for the compound having L-Phe.^[30]

Formation of bis-CDs receptors has the advantage of building polyfunctional hosts. One of the first examples was described by Breslow's group (**13** in figure 5). Two CDs were linked through a bipyridine moiety allowing for the recognition of a hydrophobic dipeptide with high binding constants ($2.5 \times 10^6 \text{ M}^{-1}$ for Trp-Trp). The addition of Zn(II) increases the binding one order of magnitude.^[31] Other receptors based on bis-CDs for the recognition of di or tri-peptides have been reported with^[32] or without the aid of metal centers.^[33]

5. Cucurbit[n]urils.

Cucurbit[n]urils (CBn) **14** are synthetic macrocycles that have widespread applications in molecular recognition and are able to bind organic amines with affinities ranging from 10^3 M^{-1} to 10^{12} M^{-1} in aqueous solutions.^[34] These molecules present a hydrophobic cavity while the two openings are lined with ureido-carbonyl oxygens (figure 6). The size of the CBn can be controlled by the number of monomers present in the structure. Thanks to their particular structure, CBn receptors can bind guests that contain cationic and nonpolar groups via ion-dipole interactions with carbonyl groups and hydrophobic interactions within the cavity. Their recognition properties towards amino acids were first studied by Buschmann^[35] with cucurbit[6]uril leading binding constants in the order of 1×10^3 to $5 \times 10^3 \text{ M}^{-1}$ for hydrophobic amino acids such as Gly, Val, Phe and Ala in 50% v/v formic acid.^[25b] Macartney studied the binding of CB7 with several guests in water at neutral pH and found that this macrocycle forms strong complexes with a series of cationic cholines and some zwitterions like carnitine or phosphorylated choline.^[36] The same compound can also form stable complexes with amino acids, especially with those having aromatic non-polar side chains.^[37] Taking advantage of this, a displacement assay using a fluorophore allowed to design tandem enzyme assays for multiparameter sensor arrays and for the determination of the enantiomeric excess of amino acids. The molecular recognition properties of CB7 towards phenylalanine and several of its derivatives and their peptides was described by Urbach.^[38] ITC measurements showed 4-*tert*-butyl (tBuPhe) and 4-aminomethyl (AMPhe) derivatives of this Phe increased the binding constant 20-30 fold as compared to Phe. Placement on the latter on the N-terminus of model tripeptides (Xxx-Gly-Gly) revealed that AMPhe-Gly-Gly binds to CB7 with a dissociation constant (K_d) of 0.95 nM at neutral pH which represents a 500-fold respect Phe-Gly-Gly. Different functional groups were shown to cooperatively interact to achieve this stability: ion dipole forces between the ammonium groups of the side chain and terminal groups with the carbonyl rim, H-bonding interactions with the peptide backbone and the burial of the aromatic amino acid inside the hydrophobic cleft.

Figure 6. Chemical structure of cucurbit[n]urils and schematic stabilizing forces for the formation of inclusion complexes.

The complexation of CB7 with amino acids having different side chains has also been investigated by ESI-MS techniques. In the gas phase, hydrogen bonding interactions and ion-dipole attraction between the ammonium and the ureido portal are the dominant stabilizing interactions. Thus, the higher affinity for aromatic residues in aqueous solutions is not observed in the gas phase.^[39]

The binding properties towards amino acids, peptides and proteins of the larger CB8 has been extensively studied and reviewed by Urbach.^[40] CB8 binds Trp and Phe in phosphate buffer at pH 7 and, with a lower affinity, Tyr. The rest 16 amino acids didn't show inclusion complexes as evidenced by ITC measurements.^[41] CB8 can include two aromatic residues in its interior thanks to the expanded cavity, thus they found 1 : 2 host : guest stoichiometry with Trp and Phe. The introduction of methyl viologen (MV) as an auxiliary guest to cucurbit[8]uril forms the corresponding [CB₈·MV] complex which then can selectively encapsulate Trp at neutral pH $K_a = 4.3 \times 10^4 \text{ M}^{-1}$. This is due to the ability of CB8 to bind two aromatic rings in its cavity in a cooperative manner. Remarkably, [CB·MV] can bind tripeptides containing Trp but the position of this amino acid deeply affects the stability of the corresponding complexes.^[42]

6. Metal complexes.

It is well known that metal ions bind simple α -amino acids thanks to the formation of 5 membered N-O chelates with the carboxylate and amino groups functions, also with the participation of side chain donor atoms when possible. Nitrogenated ligands are usually employed to form metal complexes used as receptors and the use of chiral environment allows for enantioselective discrimination. Since the early work by Chin^[43] with a pyridine and Co(III) complex (compound **15**, figure 7) a plethora of metals have been used including Cu(II), Pd(II), Pt(II), Rh(III), Sc(III) lanthanides, being the Zn(II) cation one of the most commonly used metallic centers in organometallic receptors as its complexes are inert to redox reactions and suitable for NMR spectroscopic studies owing to their diamagnetic properties.^[44] The use of metal complexes for the detection of amino acids through fluorescence, luminescence or colorimetric techniques has extensively been reviewed by different authors and it is an extensive field of study.^[45] Selected recent examples are shown in Figure 7. Kwong and Lee described a macrocyclic receptor based on the Zn(II) complex of a C₂-terpyridine and a crown that

allowed the recognition of the zwitterionic form of L-aspartate ($K = 4.5 \times 10^4 \text{ M}^{-1}$) and L-cysteine ($K = 2.5 \times 10^4 \text{ M}^{-1}$) over other proteogenic amino acids.^[46] The authors found that the binding affinity of receptor **16** towards L-amino acids depends on the coordinative properties of the functional groups in the side chains. Yu and Pu described the sensing of histidine by a crown ether–terpyridine–Zn(II) complex that shows enhanced fluorescence in the presence of His but not with other amino acids or imidazoles and was caused by the formation of rigid polymeric structures.^[47] Also based in Zn(II) complex Martínez and Dutasta reported a hemicyptophane cage combining a cyclotrimeratrylene unit and a Zn(II) metal center holding fluorescence properties (compound **17** in figure 7). The fluorescent hemicyptophane can be used as fluorescent sensor for choline phosphate.^[48]

Liu and Cui described the self-assembly of pyridine functionalized salen ligands and Zn(II) complexes to form Zn_6L_6 metallacycles with an inner cavity able to recognize different molecules and capable of enantioselective discrimination of amino acids. In accordance with the Stern–Völmer (S–V) equation, the measured absorbance I/I_0 at 465 nm varied as a function of concentration (mM) in a linear relationship suggesting 1:1 stoichiometry. The $K_{\text{sv}}(\text{R-1})/K_{\text{sv}}(\text{S-1})$ values were especially good for L-valine (4.86 ± 0.37), L-leucine (2.99 ± 0.23) and L-phenylalanine (2.54 ± 0.19).^[49] More recently, Kim et al. synthesized binaphthol aldehydes with contiguous dipicolylamine which in the presence of amino acids and Zn(II) form the corresponding Schiff base complexes. The complexes work as chiral shift reagents and discrimination between diastereomeric complexes allows for the detection of the amino acid chirality by $^1\text{H-NMR}$.^[50]

Figure 7. Metal complexes for the recognition of amino acids: Cobalt complex described by Chin (**15**), terpyridine-crown macrocycle Zn(II) complex prepared by Kwong and Lee (**16**) and Zn(II) complex hemicyptophane synthesized by Martínez and Dutasta (**17**).

A particular case of study is that of metal porphyrin complexes which have been frequently used as receptors for amino acids. Again, Zn(II) is the preferred metal center but other metals such as Cu or Co have been commonly employed. Nevertheless, the sole interaction with the metal center is not enough for an effective binding so this structure is usually decorated with extra binding sites. For example, Imai included trimethylammonium groups able to bind the carboxylate which results in good enantiomeric discrimination in water.^[51] Another strategy was developed by Schneider's group which included the synthesis water-soluble host that contained three pyridinium units crown ether moiety in

the meso positions of a porphyrin.^[52] The functionalized porphyrin is able to bind small peptides (2–4 amino acids) even in the absence of a metal complex. Metalation leads to small changes with binding affinities above 10^5 M^{-1} in water.

7. Peptidic and pseudopeptidic compounds.

Since zwitterions are important biomolecules involved in several biological processes they are recognized by several peptides and proteins. Usually, the specific recognition takes place through electrostatic and H-bonding contacts. Taking Nature of a source of inspiration, several peptides and peptidomimetics have been developed as artificial receptors for zwitterionic species. The group of Schmuck worked extensively in this area,^[53] for example, they used a combinatorial approach to find receptors for tetrapeptides in water. The main focus was the preparation of a receptor for the highly polar D-Glu-L-Lys-D-Ala-D-Ala. The designed library was built on the general formula Gua-Xxx1-Xxx2-Xxx3 involving a guanidinium pyrrol unit (Gua) for the binding of carboxylate. This motif is an excellent binder for carboxylates including the C-terminus of peptides. The molecular diversity came from the combination of amino acids (a set of eight) in the three contiguous positions (figure 8). Thus, a quantitative on-bead binding assay was performed with the *N*-dansyl fluorescent labelled substrate peptide reaching to affinities up to $K = 17100 \text{ M}^{-1}$. The results were confirmed by UV titration that rendered a $K = 15400 \text{ M}^{-1}$ for the best binder, the Gua-Lys-Lys-Phe peptide. Moreover, the same group used this combinatorial approach to assess substrate selectivity. The binding assays revealed that affinity to the inverse peptidic sequence is smaller than in the initial substrate. In addition, 320 substrates were evaluated finding that closely related structures are effectively discriminated by the receptor. Furthermore, the Gua-Lys-Lys-Phe receptor is able to discriminate between D/L isomers in some tetrapeptides. This stereoselectivity is highly dependent on the position of the stereocenter, which is explained by the necessity of a multipoint binding with well-defined structures.^[54]

Figure 8. Synthesis of a solid-supported combinatorial library for the selective interaction with D-Glu-L-Lys-D-Ala-D-Ala.

Later on, the same group used a rational design to synthesize receptors for the Arg-Gly-Asp peptide which plays an important role in bacterial cell-wall maturation. The receptor design involves two independent recognition motifs for the guanidinium and carboxylate side chains separated by a covalent linker. A *m*-xylylene bisphosphonate bisanion and a guanidiniocarbonyl

pyrrole cation were chosen for guanidinium and carboxylate recognition respectively. The choice of an appropriate spacer was found to be crucial to achieve a strong binding; as stated earlier, flexible linkers lead to intermolecular dimerization or intramolecular folding so a rigid aromatic spacer was needed.^[55] Intriguingly, a higher binding constant of $K = 4700 \text{ M}^{-1}$ was obtained when they used the corresponding bis-phosphonate tetramethyl ester, which lacks the negative charges for guanidinium recognition. Since this is a more hydrophobic receptor the authors concluded that most likely this improvement is due to an easier desolvation process prior to the formation of hydrophobic contacts.

8. Miscellaneous receptors.

The former categories are quite flexible and some of the compounds could easily be considered in more than one. However, some of the structures reported for zwitterion recognition do not fit well in any of them. Some of the different compounds used for the recognition are described here. For example, Schrader and co-workers prepared a molecular tweezer **19** for the recognition of Arginine and Lysine (figure 9).^[56] The design of the receptor features an aromatic cleft-like cavity adorned with two peripheral anionic phosphonate groups for the interaction with positively charged groups. Binding studies in buffered D₂O revealed remarkable binding constants for basic amino acids ($K = 1400 \text{ M}^{-1}$ for Lys). However, the association constants for protected amino acids were higher than those of the zwitterionic guests. An extension of this work was published later investigating different anionic sites for binding (phosphate, methanephosphonate, sulfate, or OCH₂-carboxylate groups). While they all form stable host-guest complexes with various amino lysine or arginine and some peptides containing these amino acids, the receptors containing carboxylate functions are significantly weaker hosts as they direct the amino acid side chain outside of the cavity losing CH- π and hydrophobic interactions.^[57] After evaluating different solvents the authors reach the conclusion that the hydrophobic effect in water is the major force driving the guest molecule inside the tweezer.

A macrocyclic receptor based on 1,1'-binaphthyl imidazolium was synthesized by You and Lan. This compound allowed the selective recognition of tryptophan (Trp) among the other amino acids in aqueous environment. Multiple hydrogen bonding and electrostatic interactions are involved in the recognition of the guest.^[58] As described earlier, the group of Dutasta designed and synthesized several hemicyptophanes that function as ditopic receptors. In a first design a hemicyptophane formed from a cyclotrimeratrylene (CTV) unit and a triamide subunit was studied. While the CTV can recognize ammonium ions the triamide can stabilize negative charges by hydrogen bonding. Thus, the group studied the binding with several zwitterions using DOSY experiments in CD₃CN/D₂O 90:10. In these conditions, the hemicyptophane was able to interact strongly with taurine $K_a = 14000 \text{ M}^{-1}$ and to a lesser extent with β -alanine ($K_a = 8500 \text{ M}^{-1}$).^[59] In order to avoid possible electrostatic repulsions between the anions and electron-rich aromatic rings of the linkers a new design was performed. In the new receptor **20** an electron-deficient system replaced the aforementioned linker improving the

binding properties of the system by the establishment of anion- π interactions. The number of hydrogen bond-donors was also increased by adding three amide functions.^[60] Thus, the binding constant raised up to three orders of magnitude giving $K_a = 5.0 \pm 0.4 \times 10^5 \text{ M}^{-1}$ for taurine and $2.3 \pm 0.1 \times 10^5 \text{ M}^{-1}$ for GABA in CD₃CN/D₂O 80:20.

More recently, a related structure was described by Hutton using a cyclic hexapeptide attached to the CTV moiety through the side-chains of three Tyr residues; modest recognition constants and chiral discriminations were obtained using carnitine as guest.^[61] Lately different scaffolds used in zwitterion recognition have been described: Costa published a receptor based on bis-squaramides;^[62] the group of Moran has published different molecular receptors for the formation of non-polar host-guest complexes including a system based in a benzofuran skeleton and two pyridine molecules attached as sulfonamide and carboxamide groups^[63] or a receptor based on a chiral chromane.^[64] A choline phosphate receptor based on a tripodal hexa-urea scaffold functionalized with aromatic terminal groups **21** was described by Yang and Wu.^[65] The complex is stabilized thanks to the dual binding mode where the phosphate is coordinated by the urea groups and the quaternary ammonium moiety is bound in a 'composite aromatic box' through cation- π and hydrogen bonding interactions.

A binding constants of $1.2 \times 10^6 \text{ M}^{-1}$ in 2% D₂O-DMSO-*d*₆ is obtained for this receptor and carnitine.

Figure 9. Miscellaneous receptors: a supramolecular molecular tweezer **19** described by Schrader; hemicyptophane **20** reported by Dutasta and tripodal urea receptor **21** prepared by Yang and Wu.

Finally, molecular recognition followed by covalent trapping of a zwitterion has been observed to shift the equilibria in dynamic covalent libraries. Solà and Alfonso reported how the recognition of zwitterionic cysteine and its reaction by the free sulfide functional group leads to simplification of potentially complex mixtures.^[66] The authors later used this specificity to design a modified library that gives a fluorescent response in the presence of cysteine (or its oxidized dimer, cystine) and allows to detect this amino acid in aqueous environment and even in urine.^[67]

Conclusions and Outlook

In this minireview we summarized some of the strategies for zwitterion recognition with artificial hosts. A variety of receptors including crown ethers, calixarenes, cucurbiturils or metal

complexes has been discussed. The selectivity and efficiency of the different receptors strongly depends on the molecular design and complementarity with the guest molecules and thus different scaffolds are needed. Despite remarkable examples, most known receptors are mainly addressed to amino acid molecules or short peptides. In addition, efficient binding of zwitterionic molecules in aqueous or protic media is still a challenge due to the limitations of solubility for organic receptors and the cost of desolvation. Therefore, investigation towards recognition of other guests, preparation of water-soluble receptors and the progress concerning practical applications should be addressed in the following years.

Acknowledgements

Financial Support from the Ministry of Science, Innovation and Universities (RTI2018-096182-B-I00, MCIU/AEI/FEDER,EU), and AGAUR (2017SGR208) are gratefully acknowledged..

Keywords: keyword 1 • keyword 2 • keyword 3 • keyword 4 • keyword 5

- [1] a) C. Schmuck, *Eur. J. Org. Chem.* **1999**, 1999, 2397-2403; b) C. Schmuck, *J. Org. Chem.* **2000**, 65, 2432-2437; c) S. Carsten and G. Lars, *Current Organic Chemistry* **2003**, 7, 1485-1502; d) C. Schmuck and W. Wienand, *J. Am. Chem. Soc.* **2003**, 125, 452-459; e) C. Schmuck and J. Dudaczek, *Eur. J. Org. Chem.* **2007**, 2007, 3326-3330.
- [2] a) H. J. Schneider, D. Guettes and U. Schneider, *J. Am. Chem. Soc.* **1988**, 110, 6449-6454; b) T. Lippmann, H. Wilde, M. Pink, A. Schäfer, M. Hesse and G. Mann, *Angew. Chem. Int. Ed. Engl.* **1993**, 32, 1195-1197; *Angew. Chem.* **1993**, 105, 1258-1260; c) S. M. Biros, E. C. Ullrich, F. Hof, L. Trembleau and J. Rebek, *J. Am. Chem. Soc.* **2004**, 126, 2870-2876; d) S. M. Biros and J. Rebek, *Chem. Soc. Rev.* **2007**, 36, 93-104.
- [3] L. Muthiac, J. H. Lee, J. S. Kim and J. Vicens, *Chem. Soc. Rev.* **2011**, 40, 2777-2796.
- [4] P. Ballester, A. Shivanyuk, A. R. Far and J. Rebek, *J. Am. Chem. Soc.* **2002**, 124, 14014-14016.
- [5] F. Hof, L. Trembleau, E. C. Ullrich and J. Rebek, Julius, *Angew. Chem. Int. Ed.* **2003**, 42, 3150-3153; *Angew. Chem.* **2003**, 115, 3258-3261.
- [6] P. Ballester and M. A. Sarmentero, *Org. Lett.* **2006**, 8, 3477-3480.
- [7] B. Schnatwinkel, M. V. Rekharsky, V. V. Borovkov, Y. Inoue and J. Mattay, *Tetrahedron Lett.* **2009**, 50, 1374-1376.
- [8] F. Ikuhide, T. Daiiki, K. Ryo, M. Kazutaka and A. Katsuyuki, *Bull. Chem. Soc. Jpn.* **2011**, 84, 1133-1135.
- [9] L. Tang, L. Bai, W.-h. Wang and T. Jiang, *Nature Structural & Molecular Biology* **2010**, 17, 492-496.
- [10] S. Moerkerke, J. Wouters and I. Jabin, *J. Org. Chem.* **2015**, 80, 8720-8726.
- [11] a) S. Moerkerke, V. Malytskyi, L. Marcelis, J. Wouters and I. Jabin, *Org. Biomol. Chem.* **2017**, 15, 8967-8974. b) E. Brunetti, S. Moerkerke, J. Wouters, K. Bartik and I. Jabin, *Org. Biomol. Chem.* **2016**, 14, 10201-10207.
- [12] D. Cornut, S. Moerkerke, J. Wouters, G. Bruylants and I. Jabin, *Chem. Asian J.* **2015**, 10, 440-446.
- [13] Q. He, G. I. Vargas-Zúñiga, S. H. Kim, S. K. Kim and J. L. Sessler, *Chem. Rev.* **2019**, 119, 9753-9835.
- [14] M. Barboiu, C. T. Supuran, A. Scozzafava, F. Briganti, C. Luca, G. Popescu, L. Cot and N. Hovnanian, *Liebigs Annalen* **1997**, 1997, 1853-1859.
- [15] F. P. Schmidtchen, *J. Org. Chem.* **1986**, 51, 5161-5168.
- [16] M. A. Hossain and H.-J. Schneider, *J. Am. Chem. Soc.* **1998**, 120, 11208-11209.
- [17] a) A. Galan, D. Andreu, A. M. Echavarren, P. Prados and J. De Mendoza, *J. Am. Chem. Soc.* **1992**, 114, 1511-1512; b) P. Breccia, M. Van Gool, R. Pérez-Fernández, S. Martín-Santamaría, F. Gago, P. Prados and J. de Mendoza, *J. Am. Chem. Soc.* **2003**, 125, 8270-8284; c) A. Späth and B. König, *Tetrahedron* **2010**, 66, 1859-1873.
- [18] a) A. Moghimi, M. F. Rastegar, M. Ghandi, M. Taghizadeh, A. Yari, M. Shamsipur, G. P. A. Yap and H. Rahbarnoochi, *J. Org. Chem.* **2002**, 67, 2065-2074; b) A. Moghimi, B. Maddah, A. Yari, M. Shamsipur, M. Boostani, M. Fall Rastegar and A. R. Ghaderi, *Journal of Molecular Structure* **2005**, 752, 68-77.
- [19] N. Bernier, C. V. Esteves and R. Delgado, *Tetrahedron* **2012**, 68, 4860-4868.
- [20] M. T. Reetz, J. Huff, J. Rudolph, K. Toellner, A. Deege and R. Goddard, *J. Am. Chem. Soc.* **1994**, 116, 11588-11589.
- [21] A. M. Costero, G. M. Rodríguez-Muñiz, S. Gil, S. Peransi and P. Gaviña, *Tetrahedron* **2008**, 64, 110-116.
- [22] R. Gotor, A. M. Costero, S. Gil, P. Gaviña and K. Rurack, *Eur. J. Org. Chem.* **2014**, 2014, 4005-4013.
- [23] a) J. V. Hernández, F. M. Muñoz, A. I. Oliva, L. Simón, E. Pérez and J. n. R. Morán, *Tetrahedron Lett.* **2003**, 44, 6983-6985; b) O. H. Rubio, R. Taouil, F. M. Muñoz, L. M. Monleón, L. Simón, F. Sanz and J. R. Morán, *Org. Biomol. Chem.* **2017**, 15, 477-485.
- [24] a) L. Lamarque, P. Navarro, C. Miranda, V. J. Arán, C. Ochoa, F. Escartí, E. García-España, J. Latorre, S. V. Luis and J. F. Miravet, *J. Am. Chem. Soc.* **2001**, 123, 10560-10570; b) C. Miranda, F. Escartí, L. Lamarque, M. J. R. Yunta, P. Navarro, E. García-España and M. L. Jimeno, *J. Am. Chem. Soc.* **2004**, 126, 823-833.
- [25] a) I. Tabushi, Y. Kuroda and T. Mizutani, *J. Am. Chem. Soc.* **1986**, 108, 4514-4518; b) H. J. Buschmann, E. Schollmeyer and L. Muthiac, *Thermochim. Acta* **2003**, 399, 203-208.
- [26] a) C. Kahle, R. Deubner, C. Schollmayer, J. Scheiber, K. Baumann and U. Holzgrabe, *Eur. J. Org. Chem.* **2005**, 2005, 1578-1589; b) V. P. Kumar, I. Suryanarayana, Y. V. D. Nageswar and K. R. Rao, *Helv. Chim. Acta* **2008**, 91, 753-758.
- [27] P. Stepniak, B. Lainer, K. Chmurski and J. Jurczak, *Carbohydr. Polym.* **2017**, 164, 233-241.
- [28] P. Stepniak, B. Lainer, K. Chmurski and J. Jurczak, *RSC Advances* **2017**, 7, 15742-15746.
- [29] R. Corradini, A. Dossena, G. Impellizzeri, G. Maccarrone, R. Marchelli, E. Rizzarelli, G. Sartor and G. Vecchio, *J. Am. Chem. Soc.* **1994**, 116, 10267-10274.
- [30] S. Pagliari, R. Corradini, G. Galaverna, S. Sforza, A. Dossena, M. Montalti, L. Prodi, N. Zaccaroni and R. Marchelli, *Chem. Eur. J.* **2004**, 10, 2749-2758.
- [31] R. Breslow and B. Zhang, *J. Am. Chem. Soc.* **1992**, 114, 5882-5883.
- [32] Y. Liu, Y.-L. Zhao, Y. Chen, F. Ding and G.-S. Chen, *Bioconjugate Chem.* **2004**, 15, 1236-1245.
- [33] Y. Liu, Y.-W. Yang, Y. Chen and F. Ding, *Biorg. Med. Chem.* **2005**, 13, 963-971.
- [34] a) W. L. Mock and N. Y. Shih, *J. Am. Chem. Soc.* **1988**, 110, 4706-4710; b) J. Lagona, P. Mukhopadhyay, S. Chakrabarti and L. Isaacs, *Angew. Chem. Int. Ed.* **2005**, 44, 4844-4870; *Angew. Chem.* **2005**, 117, 4922-4949.
- [35] H. J. Buschmann, K. Jansen and E. Schollmeyer, *Thermochim. Acta* **1998**, 317, 95-98.
- [36] I. W. Wyman and D. H. Macartney, *Org. Biomol. Chem.* **2010**, 8, 253-260.
- [37] D. M. Bailey, A. Hennig, V. D. Uzunova and W. M. Nau, *Chem. Eur. J.* **2008**, 14, 6069-6077.
- [38] L. A. Logsdon, C. L. Schardon, V. Ramalingam, S. K. Kwee and A. R. Urbach, *J. Am. Chem. Soc.* **2011**, 133, 17087-17092.
- [39] E. Kovalenko, M. Vilaseca, M. Díaz-Lobo, A. N. Masliy, C. Vicent and V. P. Fedin, *J. Am. Soc. Mass. Spectrom.* **2016**, 27, 265-276.
- [40] A. R. Urbach and V. Ramalingam, *Isr. J. Chem.* **2011**, 51, 664-678.
- [41] P. Rajgariah and A. R. Urbach, *J. Incl. Phenom. Macrocycl. Chem.* **2008**, 62, 251-254.

-
- [42] L. M. Heitmann, A. B. Taylor, P. J. Hart and A. R. Urbach, *J. Am. Chem. Soc.* **2006**, *128*, 12574-12581.
- [43] J. Chin, S. S. Lee, K. J. Lee, S. Park and D. H. Kim, *Nature* **1999**, *401*, 254-257.
- [44] Y. Shimazaki, M. Takani and O. Yamauchi, *Dalton Transactions* **2009**, 7854-7869.
- [45] a) Y. Zhou and J. Yoon, *Chem. Soc. Rev.* **2012**, *41*, 52-67; b) J. Wang, H.-B. Liu, Z. Tong and C.-S. Ha, *Coord. Chem. Rev.* **2015**, *303*, 139-184.
- [46] H.-L. Kwong, W.-L. Wong, C.-S. Lee, C.-T. Yeung and P.-F. Teng, *Inorg. Chem. Commun.* **2009**, *12*, 815-818.
- [47] J. Du, Z. Huang, X.-Q. Yu and L. Pu, *Chem. Commun.* **2013**, *49*, 5399-5401.
- [48] D. Zhang, G. Gao, L. Guy, V. Robert, J.-P. Dutasta and A. Martinez, *Chem. Commun.* **2015**, *51*, 2679-2682.
- [49] J. Dong, C. Tan, K. Zhang, Y. Liu, P. J. Low, J. Jiang and Y. Cui, *J. Am. Chem. Soc.* **2017**, *139*, 1554-1564.
- [50] M. E. Shirbhate, R. Nandhakumar, Y. Kim, S.-J. Kim, S. K. Kim and K. M. Kim, *Eur. J. Org. Chem.* **2018**, *2018*, 4959-4964.
- [51] H. Imai, H. Munakata, Y. Uemori and N. Sakura, *Inorg. Chem.* **2004**, *43*, 1211-1213.
- [52] M. Sirish, V. A. Chertkov and H.-J. Schneider, *Chem. Eur. J.* **2002**, *8*, 1181-1188.
- [53] D. Maity and C. Schmuck in *Chapter 8 Synthetic Receptors for Amino Acids and Peptides*, The Royal Society of Chemistry, **2015**, pp. 326-368.
- [54] C. Schmuck and P. Wich, *Angew. Chem. Int. Ed.* **2006**, *45*, 4277-4281; *Angew. Chem.* **2006**, *118*, 4383-4387.
- [55] C. Schmuck, D. Rupprecht, M. Junkers and T. Schrader, *Chem. Eur. J.* **2007**, *13*, 6864-6873.
- [56] M. Fokkens, T. Schrader and F.-G. Klärner, *J. Am. Chem. Soc.* **2005**, *127*, 14415-14421.
- [57] S. Dutt, C. Wilch, T. Gersthagen, P. Talbiersky, K. Bravo-Rodriguez, M. Hanni, E. Sánchez-García, C. Ochsenfeld, F.-G. Klärner and T. Schrader, *J. Org. Chem.* **2013**, *78*, 6721-6734.
- [58] L. Yang, S. Qin, X. Su, F. Yang, J. You, C. Hu, R. Xie and J. Lan, *Org. Biomol. Chem.* **2010**, *8*, 339-348.
- [59] O. Perraud, V. Robert, A. Martinez and J.-P. Dutasta, *Chem. Eur. J.* **2011**, *17*, 13405-13408.
- [60] O. Perraud, V. Robert, H. Gornitzka, A. Martinez and J.-P. Dutasta, *Angew. Chem. Int. Ed.* **2012**, *51*, 504-508. *Angew. Chem.* **2012**, *124*, 519-523.
- [61] J. R. Cochrane, A. Schmitt, U. Wille and C. A. Hutton, *Chem. Commun.* **2013**, *49*, 8504-8506.
- [62] C. Lopez, E. Sanna, L. Carreras, M. Vega, C. Rotger and A. Costa, *Chem. Asian J.* **2013**, *8*, 84-87.
- [63] F. G. Herrero, O. H. Rubio, L. M. Monleón, Á. L. Fuentes de Arriba, L. S. Rubio and J. R. Morán, *Org. Biomol. Chem.* **2016**, *14*, 3906-3912.
- [64] Á. G. Temprano, L. M. Monleón, O. H. Rubio, L. S. Rubio, A. B. Pérez, F. Sanz and J. R. Morán, *Org. Biomol. Chem.* **2016**, *14*, 1325-1331.
- [65] W. Zuo, C. Jia, H. Zhang, Y. Zhao, X.-J. Yang and B. Wu, *Chem. Sci.* **2019**, *10*, 2483-2488.
- [66] M. Lafuente, J. Atcher, J. Sola and I. Alfonso, *Chem. Eur. J.* **2015**, *21*, 17002-17009.
- [67] M. Lafuente, J. Sola and I. Alfonso, *Angew. Chem. Int. Ed.* **2018**, *57*, 8421-8424; *Angew. Chem.* **2018**, *130*, 8557-8561.

Entry for the Table of Contents (Please choose one layout)

Layout 1:

MINIREVIEW

Molecular receptors developed in the last years for recognition of zwitterionic species are reviewed.

Ignacio Alfonso and Jordi Solà**

Page No. – Page No.

**Molecular recognition of zwitterions
with artificial receptors**

Layout 2:

MINIREVIEW

((Insert TOC Graphic here))

*Author(s), Corresponding Author(s)**

Page No. – Page No.

Title

Text for Table of Contents