# Learning from Proteins and Drugs: Receptors That Mimic Biomedically Important Binding Motifs

**Fraser Hof and Thomas Pinter**

**Abstract** Proteins are highly evolved, sophisticated machines which function together to maintain homeostasis in their hosts. While merely a collection of amino acids covalently bonded in a specific sequence, their wide variety of functions is truly remarkable. Of course these covalent sequences are essential for proper function, but equally important for proper function are *weak* interactions: protein folding, enzyme-substrate interaction, and protein–protein communication are all controlled by forces weaker than covalent bonds and understanding these forces is fundamental in medicinal chemistry and drug design. Many inhibitory drugs mimic natural substrates for protein binding sites but inhibition of the substrate by mimicking the binding site is also possible. This mimicry and the biological consequences are under investigation.

**Keywords** Aromatic interactions, Cation-pi interactions, Host-guest chemistry, Hydrogen bonds, Molecular Recognition

## Contents



F. Hof (✉) and T. Pinter
Department of Chemistry, University of Victoria, Victoria, BC, Canada V8W 3V6
e-mail: fhof@uvic.ca

S.A. Piletsky and M.J. Whitcombe (eds.), *Designing Receptors for the Next Generation of Biosensors*, Springer Series on Chemical Sensors and Biosensors (2013) 12: 33–52
DOI 10.1007/5346_2012_15, 
Published online: 27 June 2012

## 1 Introduction: Weak Interactions

Weak interactions determine protein size and shape and are therefore an essential part of normal protein function. Discreet binding pockets and motifs that have evolved to be highly selective for only a very particular class of substrates come about as a result of a myriad of these non-covalent interactions. The helical shape of DNA is so because of an intricate combination of H-bond donor and acceptor pairs, stacking between the bases and solvation effects. Weak hydrogen bonding, electrostatic, and hydrophobic interactions play roles in all protein–substrate and protein–protein interactions.

The most commonly observed weak interactions and arguably most important for normal protein function are hydrogen bonds. Ubiquitous in complex natural systems, almost all biological processes involve hydrogen bonding in some form or another and these interactions have been the topic of extensive study for decades [1–6]. Proteins which act on anionic substrates universally contain highly evolved hydrogen bond networks in their active sites. Figure 1a depicts the pore of a ClC chloride channel whose crystal structure was solved in 2002 [7]. The ion is coaxed into the pore by four key hydrogen bond donating residues. These attractive interactions pull the chloride in close proximity to an aspartic acid residue which is displaced, thus opening the ion channel. During drug design, medicinal chemists often seek to emulate the natural substrate of a biological target and attempt to preserve all attractive forces in the host–guest complex. Replacement of the phosphate linker in natural RNA with an acylsulfonamide in a simple dinucleoside mimic (**2**) preserved a key H-bond with His119 and resulted in inhibition of RNase A (Fig. 1b, c) [8].

Other essential, yet not as well understood non-covalent interactions present in biological systems are those involving aromatic residues [9–12] and the hydrophobic effect [13–16]. The former is often considered to be a result of the latter, as hydrophobic aromatic residues pack close together in the core of the protein while hydrophilic residues are more often observed near the protein surface [16]. Although disfavored on the basis of configurational entropy, as a folded protein loses huge numbers of degrees of freedom relative to its unfolded state, the favored enthalpic gains of water molecules able to hydrogen bond to each other causes the overall energy of the system to be favorable. As a direct result, hydrophobic aromatic residues are often seen stacked on one another. Doubly mutating two stacked tyrosine residues in bacterial ribonuclease (barnase) resulted in a 4.6 kcal/mol decrease in protein stability (Fig. 2a) [9]. Because many hydrophilic residues reside near protein surfaces, salt bridges and cation–pi interactions are often seen at protein–protein interfaces [3]. The latter is sometimes observed as a quaternary ammonium residue bound inside an electron-rich pocket heavily populated with aromatic residues. These binding sites, or "hotspots," contain highly preorganized tryptophan, phenylalanine, and tyrosine residues which stabilize the incoming cation through their electron-rich pi-clouds. An important interaction of this type is between methylated lysine and the CBX class of proteins, when methylation of lysine is misregulated, disease often follows [17–20]. Figure 2b depicts the co-

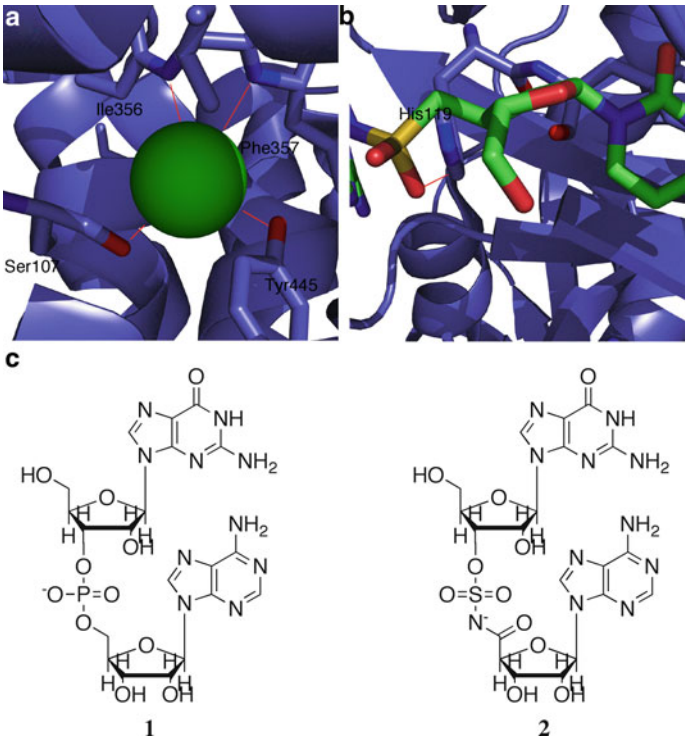


**Fig. 1** (**a**) Crystal structure of chloride (*green sphere*) bound in the pore of a ClC chloride channel (PDB 1KPL). Key hydrogen bond contacts are observed with surrounding Ile, Ser, Tyr, and Phe residues. (**b**) *N*-acyl sulfonamide linked dinucleoside mimic bound to RNase A. A key H-bond between the sulfone of the inhibitor and a nearby histidine is observed in the crystal structure (PDB 2XOI). H-bonds are shown as *red lines*. (**c**) Natural dimeric RNA fragment (*left*) and *N*-acyl sulfonamide functionalized mimic (*right*). Both compounds are deprotonated at physiological pH and the mimic displays moderate inhibitory activity against RNase A

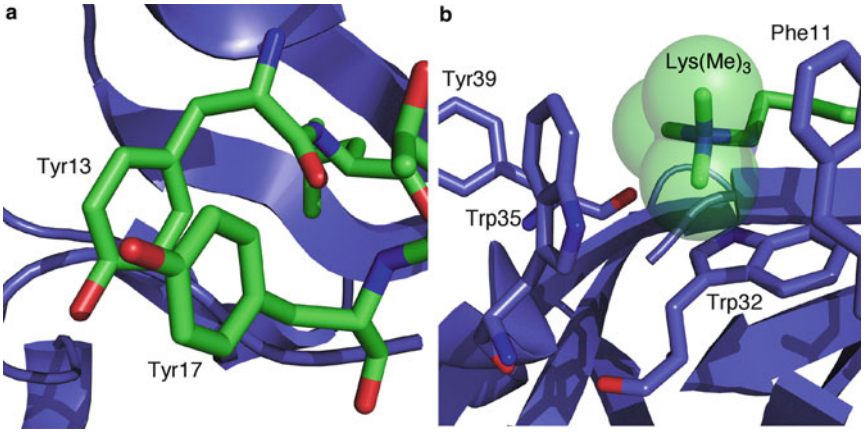


**Fig. 2** (**a**) Aromatic–aromatic stacking in barnase stabilizes the protein (PDB 1A2P). (**b**) Trimethylated lysine reisdue is recognized by an "aromatic cage" binding motif in the protein CBX6 (PDB 3I90)

structure of a peptide fragment with trimethylated lysine binding in the aromatic cage pocket of CBX6 [21].

Development of compounds which emulate a natural substrate, be it for protein inhibition or detection if a fluorescent tag is attached, is commonplace. Is it possible to develop sensors which do the opposite—that is, mimic the binding domain of the protein itself and sequester the substrate before it enters? Our group and others are actively exploring this avenue of research.

## 2 Anion Recognition by Carboxylic Acid Bioisosteres

Since the advent of crown ethers in the 1960s, cation coordination chemistry has been thoroughly studied and rational design of synthetic cation receptors has matured a great deal. Anion recognition chemistry, however, had received less attention until the 1980s. The importance of anionic species to living systems is critical. Anions are ubiquitous in biological systems: careful regulation of intra- and extracellular charge gradients is necessary to maintain homeostasis, and the majority of enzyme substrates and cofactors carry a negative charge. DNA owes its helical shape to well-defined hydrogen bond networks between complimentary base pairs, phosphates provide the energy source crucial to all biochemical processes, transport channels for small anions such as chloride and sulfate regulate the flow of nutrients and osmotic pressure in and out the cell. Misregulation of these certain chloride channels have been proven to cause the debilitating respiratory illness cystic fibrosis [22], along with degenerative renal ailments Dent's disease [23] and Bartter's syndrome [24]. Many pollutants, be it from agricultural runoff (lake eutrophication from excess phosphate), nuclear waste (radioactive pertechnetate discarded into the ocean), are a cause of growing environmental concern and are anionic in nature. It is not surprising then that much attention has been focused on creating potent receptors that are selective for anionic species of interest [25–27].

In biological systems, the most prominent anion present is carboxylate. Negatively charged amino acids are often present at cation-binding hotspots to offer favorable electrostatic attractions. They are also one half of the common salt-bridge binding motif and are present in myriad enzyme substrates and cofactors. Development of drugs and sensors that mimic these substrates often requires the creation of esterified analogs that upon passage through a cell membrane are hydrolyzed to the corresponding acid derivative by native esterases. An alternate strategy in drug development is replacement of the carboxylic acid group with a functionally similar moiety, one for which the body lacks the evolved metabolic pathways to affect its degradation. These are known as bioisosteres of carboxylic acid. Common acid bioisosteres include aryl sulfonamides, acyl sulfonamides, and tetrazoles (Fig. 3). Utilizing these functional groups in drug development often leads to improved oral availability, metabolic stability, and potency relative to carboxylate bearing analogs [28–30]. It follows that a vast number of small molecule therapeutic targets contain tetrazole [31, 32], aryl sulfonamide [33, 34], and acyl sulfonamide functionality [35–37].

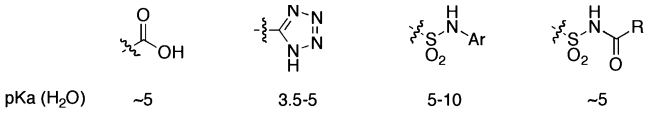


**Fig. 3** Some common carboxylic acid bioisosteres

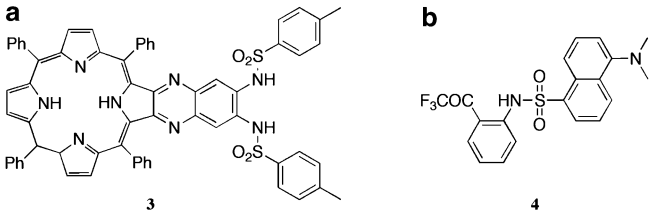


**Fig. 4** (**a**) Fluoride selective sensor developed by Starnes et al. (**b**) Cyanide selective receptor developed by Ahn et al.

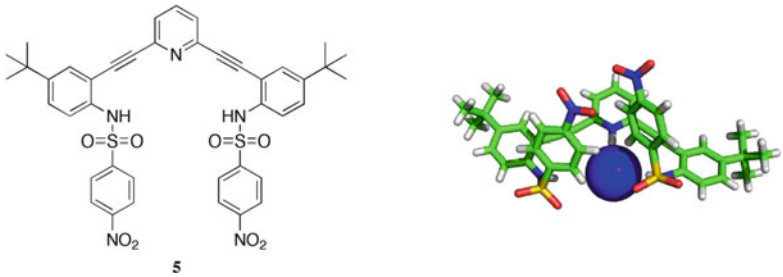


**Fig. 5** Chloride selective fluorescent sensor developed by Johnson et al. (*left*) and its crystal structure complexed with chloride (*right*)

These groups have two possible modes of molecular recognition: when protonated, the acidic hydrogen atoms should make excellent donors useful for anion binding. When deprotonated, the anionic conjugate bases of these functional groups are potential cation binders. Our group and others have investigated the use of deprotonated tetrazolates for cation binding [38–43]. While aryl sulfonamides have been used as anion binders, the affinities of tetrazoles and acyl sulfonamides in neutral, protonated form have been largely unexplored.

Starnes and co-workers appended bis-aryl sulfonamide functionality to a well-known porphyrin scaffold (**3**) and observed changes in the UV–vis absorption spectra when titrating with various anions (Fig. 4a) [44], while Ahn et al. observed fluorescence enhancement and selectivity towards cyanide with their naphthalene-based aryl sulfonamide receptor (**4**) (Fig. 4b) [45]. Recently, Johnson and co-workers have been developing aryl ethynyl scaffolds as anion sensors and have shown remarkable fluorescence imaging of chloride in vitro [46]. They have also fabricated a scaffold decorated with aryl sulfonamide functionality (**5**) that is also a successful anion binder [47] (Fig. 5). Against this backdrop of favorable results, it is


**Fig. 6** Amide and sulfonamide functionalized anion receptors exhibit different selectivities

**Table 1** Anion affinities of aryl sulfonamide functionalized hosts **8** and **9**, acyl sulfonamide functionalized hosts **10** and **11**, and tetrazole functionalized hosts **12** and **13**

| $K_{assoc}(M^{-1})$ in $CD_3CN$ | | | | | |
|---|---|---|---|---|---|
| Host | $Bu_4N^+Cl^-$ | $Bu_4N^+NBr^-$ | $Bu_4N^+OTs^-$ | $Bu_4N^+NO_3^-$ | $Bu_4N^+HSO_4^-$ |
| **10** | 600(10) | 110(10) | 60(5) | 40(40) | 50(5) |
| **11** | 1,030(5) | 320(5) | 250(5) | 100(5) | 180(10) |
| **12** | 100(50) | 20(10) | 100(100) | 80(5) | 70(20) |
| **13** | 390(10) | 90(5) | 120(5) | 60(5) | 100(5) |
| **14** | 8,450(10) | 720(10) | 1,400(20) | 5,800(30) | 660(50) |
| **15** | 3,560(40) | 800(10) | 520(10) | 330(5) | 340(5) |

Errors in brackets are standard deviations of 2–3 replicate titrations

interesting to note that the original report of isophthalamides as anion binders by Crabtree also reports aryl sulfonamide analogs that have varying affinities and selectivities relative to amides (Fig. 6) [48]. Compound **6** exhibits threefold more potent binding for chloride relative to **7** while **7** binds fluoride approximately twice as strongly as **6**.

We recently explored anion binding by all three bioisosteres—aryl sulfonamides, acyl sulfonamides, and tetrazoles—affixed to a common calixarene scaffold. Hosts **8**–**13** were prepared and their binding with several biologically important halides and oxyanions was determined. The results are summarized in Table 1. These studies revealed that although the N–H proton acidities of these three classes of compounds are similar (representative N–H p$K_a$ values are 4.6, 8.5, and 5.2 for tetrazole [49], aryl sulfonamide [50], and acyl sulfonamide [36] moieties, respectively), tetrazoles proved to be superior anion-binding elements relative to their sulfonamide analogs within this structural context. The expected trends based on N–H proton acidities were in fact observed with aryl sulfonamides, as **8** bound all anions studied less strongly than did electron poor **9**. Notably, acyl sulfonamide functionalized hosts **10** and **11**, although known to be more acidic than simple aryl sulfonamides were less competent hosts. These findings implied that there were more forces at work in these systems than simply hydrogen bond donation (Fig. 7).

We hypothesized that their varying conformational preferences, largely ignored in their simple classification as interchangeable replacements for carboxylic acids, might play a role in determining their anion-binding affinities. Molecular modeling studies were carried out to investigate the host–guest complexes. We identified geometries for each host–guest complex in which the calix[4]arene is in a perfect "cone" conformation and all four hydrogen bond donors are engaging the central

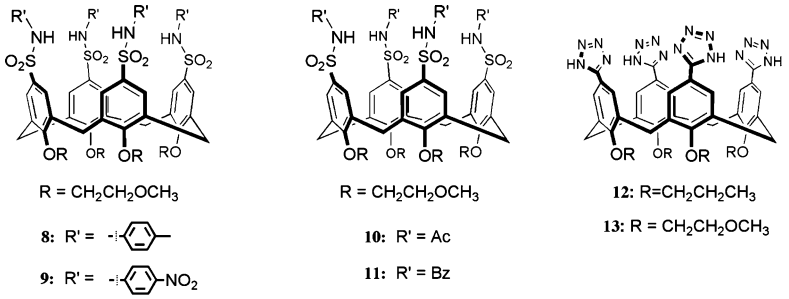


**Fig. 7** New class of carboxylic acid bioisostere functionalized calix[4]arene recognition elements

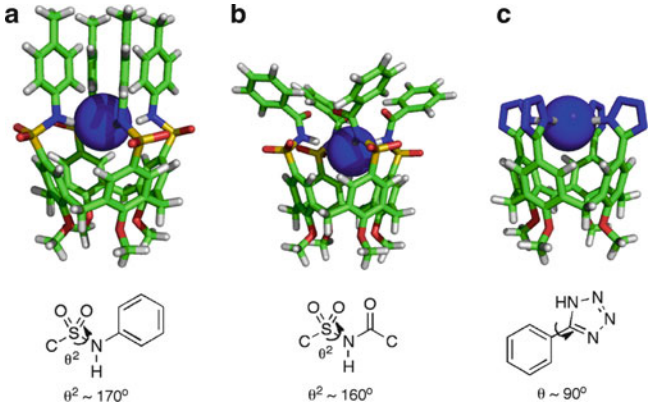


**Fig. 8** Local minima that involve the maximum four host–guest hydrogen bonds for representative host–guest complex with key dihedral angles calculated (HF/6-31+G*). Lower rim substituents have been omitted. (**a**) Aryl sulfonamide functionalized host complexed with $Cl^-$. (**b**) Acyl sulfonamide functionalized host complexed with $Cl^-$. (**c**) Tetrazole functionalized host complexed with $Cl^-$

anion symmetrically. These provided values for the key torsion angles of each functional group that would be optimal for binding (Fig. 8). The actual inherent conformational preferences for each functional group were then determined by dihedral driving calculations and CSD surveys (not shown) and compared to the values required for anion binding in this structural context.

The combined computational analyses for the key dihedral angles that define the inherent shapes of these moieties are presented in Fig. 9.

These comparisons showed that the tetrazole hosts **12** and **13** must adopt a nonideal dihedral that takes the tetrazole out of conjugation with its neighboring ring to bind an anion, and that it paid ~13 kJ/mol in order to do so (Fig. 9c). The shapes of the aryl and acyl sulfonamides are defined by three important dihedral angles. The conformations of the rotatable carbon–sulfur bonds for aryl sulfone type functionalities ($\theta_1$, Fig. 9d) and are known to be similar ($\theta_1 = \sim 90°$) [51].

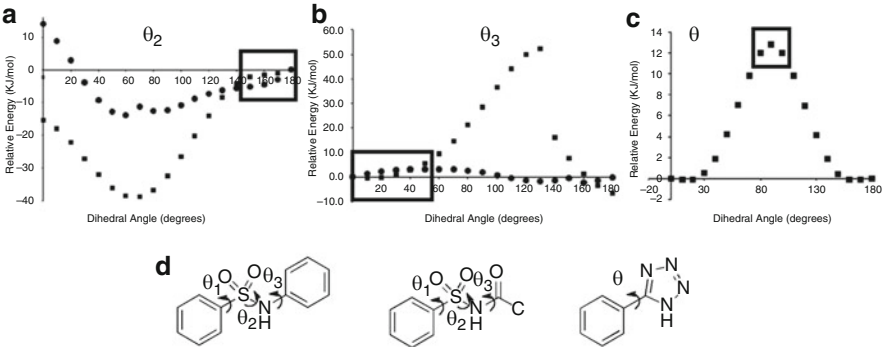


**Fig. 9** Energy diagrams calculated at the HF/6-31+G* level of theory when driving key dihedral angles from 0° to 180° for simplified fragments of aryl (*square*) and acyl (*circle*) sulfonamides (**a**, **b**) and for a simplified tetrazole fragment (**c**). (**d**) Labeling of key dihedral angles studied computationally. The approximate angles required to fully engage a central guest in a fourfold symmetrical manner are highlighted with *boxes*

Focusing on the dihedral for rotation about the sulfonamide S–N bond, ($\theta_2$; Fig. 9a) and the amide/aniline dihedrals that define rotation about the neighboring N–C bonds ($\theta_3$; Fig. 9b) provides a different picture. Computational analysis revealed that these functional groups also have preference for co-planarity with their aryl neighbors. An analysis similar to that done for the tetrazole shows that both compounds pay little energy in order for $\theta_3$ to adopt a good binding geometry. When examining $\theta_2$ however, the acyl sulfonamide fragment pays a much higher penalty to orient itself toward the guest than the aryl sulfonamide. These lessons inform on sulfonamide recognition in general, as they can also explain the failure of the Crabtree bis-sulfonamide **7**. Clearly, the chosen scaffold was not ideal for maximum host–guest binding and within the right structural context, the rigid acyl sulfonamides have potential to bind anions with more potency.

Tetrazoles are the most highly acidic of these isosteres, and should be the best hydrogen bonders of anions. Their utility as anion recognition elements has been demonstrated in a variety of contexts that are informed on by considerations of host shape and electronics. A tripodal receptor bearing three tetrazoles (**15**) was compared to a carboxylic acid analog (**14**) [52]. Binding studies showed that the tetrazole-functionalized host bound halides up to six orders of magnitude more strongly than did its carboxyl functionalized counterpart despite their nearly identical acidities. Again, the apparent similarity covers up major differences in shape and stereoelectronics that drive molecular recognition. While the acidic OH and acidic NH groups in **14** and **15** are arrayed in nearly identical positions in three-dimensional space, the OH groups are directed outward and away from the guest because of a carboxylic acid's strong preference for a *syn* OH conformation. On the contrary, the tetrazole's NH (which prefers strongly to exist as the 1H-tetrazole tautomer in polar solutions) is oriented such that all three host NHs can bind the guest simultaneously. The impact of these stereoelectronic effects all cooperating in

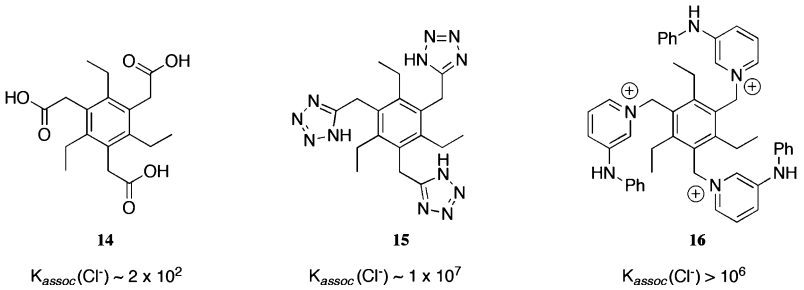


**Fig. 10** Tripodal anion receptors with association constants measured in acetonitrile

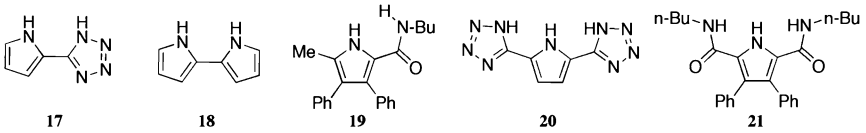


**Fig. 11** Previously studied amidopyrroles and new tetrazole containing anion binders

the binding of the guest is dramatic. Comparison with other recognition elements appended to the same scaffold reveals that this host is among the strongest neutral binders of anions ever reported—even comparable to triply cationic hosts such as **16** [53] (Fig. 10).

Further inspiration from medicinal chemistry was found in the natural product prodigiosin and its synthetic analogs (prodigiosenes), which are being heavily investigated as novel therapeutics in a variety of contexts and *whose highly potent binding of anions are linked to their biological activities* [54]. The core anion-binding motif of prodigiosenes is a 2,2′-bipyrrole; prior lessons, the shape-preferences, and anion-binding potencies of tetrazoles suggested the utility of a scaffold in which tetrazoles are installed next to a neighboring, conjugated pyrrole. 5-(Pyrrolyl)-tetrazole (**17**) is a small fragment that satisfies these design criteria (Fig. 11) [55]. The recognition properties of the tetrazole are revealed by comparisons to two key relatives, 2,2′-bipyrrole (**18**), which reproduces the geometry of **17** with a nearly isosteric tetrazole-for-pyrrole swap and a simple amidopyrrole (**19**), which is a representative member of a larger family of amidopyrroles that have been shown to bind and transport anions. Compound **16** binds almost any anion with tenfold greater affinity than the structurally related bipyrrole (**18**) (Table 2). The similar selectivity trends observed for **17** and **18** suggest that their structures and directional hydrogen bonding are similar and that the acidity of the tetrazole is the key to higher affinities. Compound **17** also binds anions at least 47-fold stronger than amidopyrroles. The tridentate 2,5-(bis-tetrazolyl)-pyrrole (**20**) was also produced and compared to a forerunner that used bis(amido)pyrroles for anion binding (**21**) [56]. Affinities improved again—this time by two orders of magnitude. These amide-to-tetrazole comparisons revealed

**Table 2** Association constants for selected anions with compounds **17–21**

| $K_{assoc}(M^{-1})$ in $CD_3CN$ | | | | |
|---|---|---|---|---|
| Host | $Bu_4N^+Cl^-$ | $Bu_4N^+NBr^-$ | $Bu_4N^+OTs^-$ | $Bu_4N^+NO_3^-$ |
| **17** | 3,300(1,200) | 450(50) | 900(50) | 160(20) |
| **18** | 310(10) | 50(3) | 40(4) | 20(1) |
| **19** | 28 | <10 | nd | nd |
| **20** | 26,000(2,300) | 1,500(430) | 34,000(3,500) | 1,600(300) |
| **21** | 138 | <10 | nd | nd |

Errors in brackets are standard deviations of 2–3 replicate titrations. For compounds **19** and **21**, errors are estimated to be <15% [56]

that the tetrazole-functionalized hosts encoded a non-Hofmeister preference for sulfonate/sulfate type anions relative to chloride, showing that even a subtle structural change (in this case, the angle of the N–H donors relative to each other) can have large effects on binding selectivities.

## 3 Receptors That Mimic Natural Aromatic Cage Motifs

As with receptors inspired by common drug motifs, inspection of natural binding motifs in proteins provides a multitude of lessons on molecular recognition. One that has found particular resonance and utility in supramolecular chemistry is the aromatic cage motif that is used throughout nature to bind tertiary and quaternary ammonium cations. This motif is typically defined as a rigid cluster of 2–4 aromatic amino acid side chains (Trp, Phe, Tyr) describing a central binding site [57]. Notable examples include the choline-binding proteins, which include acetylcholinesterase and the nicotinic acetylcholine receptor, and several gene regulation proteins that recognize and bind to the "histone code" marks of dimethyllysine and trimethyllysine side chains (see Fig. 2) [58]. The latter are especially interesting, because they have evolved in a competitive environment where they must reject binding of proteins that are identical to their targets but are unmethylated at the critical lysine side chain. Thus, their aromatic cages are the sole and unique hot spots that are responsible for this biologically driven selectivity.

It is notable that, for tertiary and especially quaternary ammonium ions, nature rarely employs a negatively charged molecular recognition element such as carboxylate or phosphate. This leads one to the idea that the cation–pi interaction and possibly the hydrophobic effect are the key operators. A large set of synthetic receptors have been used both to demonstrate that the cation–pi interaction is adept at encoding the strong and selective binding of quaternary ammonium ions in water. The variety of such structures is large, and this area has been extensively reviewed [59]. Exemplary evidence is provided by data from a single family of macrocyclic hosts invented by Dougherty (Fig. 12). Host **22**, a synthetic aromatic cage functionalized with polar solubilizing groups, demonstrated the ability to bind ammonium ion guests such as **23** and **24** with high affinities in pure water.

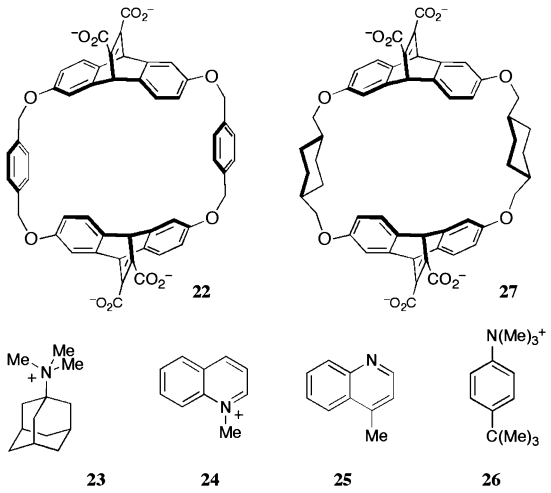


**Fig. 12** Macrocyclic hosts that mimic aromatic cage motifs, and their guests that have been used to understand the nature of their interactions

It has always been a confounding element of these motifs (natural and synthetic) that a hydrophobic component (and not the cation–pi interaction) might be sufficient to drive the observed binding of quaternary ammonium ions (which are much more hydrophobic than primary ammonium ions). But comparisons in water of nearly isostructural guests such as **24** and **25** that bear different charges revealed a significant difference in affinities that can only be explained by the multiple cation–pi interactions that exist between the guest **24** and the cyclophane. Further, the guest **26**, with isosteric aliphatic $–C(Me)_3$ and ammonium $–N(Me)_3^+$ ends, shows a strong preference for binding with the charged (and slightly less hydrophobic) $–N(Me)_3^+$ end inside the aromatic cage of **22**; if hydrophobicity were the prime driver for guest binding, then the more hydrophobic $–C(Me)_3$ end would win out. Finally, the host analog **27**, with two of the benzene rings of **22** replaced with cyclohexane walls that are more hydrophobic and more polarizable than the benzenes in the parent host, also shows decreased affinities for ammonium ion guests that support the key role of cation–pi interactions in the parent host **22**.

Other informative comparisons have been conducted using natural, protein-based aromatic cage motifs as "receptors" themselves for physical organic studies. In one kind of study, strong-binding quaternary ammonium ion ligands of the type $–N(Me)_3^+$ (**31**) are compared to ligand analogs that are isostructural except for $–C(Me)_3$ substitutions (**32**, Fig. 13) [60, 61]. The affinities in both studies are higher for the charged species that can form multiple cation–pi interactions with their protein binding partners. In a different kind of study, the protein itself is altered to reduce the strength of the proposed cation–pi interaction by installation of unnatural aromatic amino acids with decreased electron density relative to the native Trp residues (e.g., F-Trp, $F_2$-Trp, $F_3$-Trp) [62]. Again, the dominant role of certain cation–pi contacts in these biologically important recognition events is supported

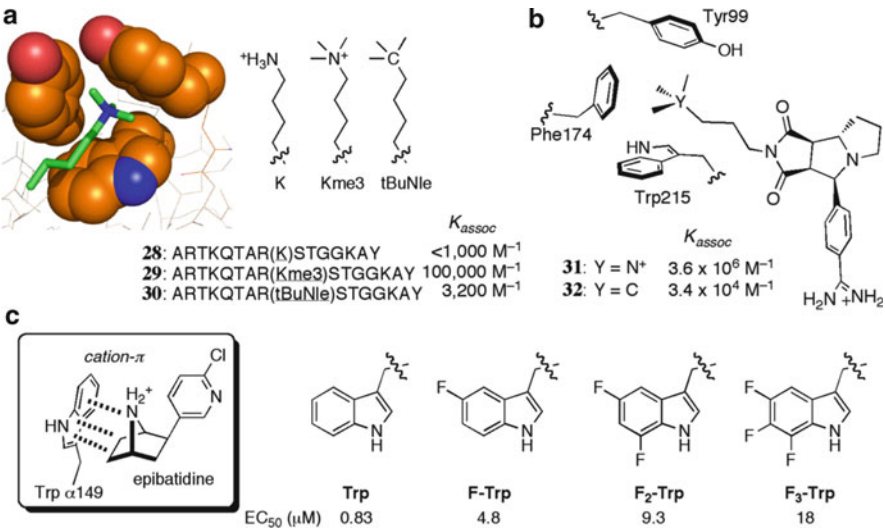


**Fig. 13** Perturbation of protein–ligand interactions by modulating the cation–pi interaction. (**a**) Peptides bearing trimethyllysine $–N(Me_3)^+$ and neutral $–C(Me_3)^+$ side chains, as well as the parent peptide with an unmodified lysine ($–NH_3^+$) side chain, and their affinities for the aromatic cage protein HP1. The HP1 binding pocket containing bound trimethyllysine is shown at *left* (1KNE). (**b**) Two inhibitors of the enzyme Factor Xa that show 100-fold increased potency for the $–N(Me_3)^+$ type inhibitor. (**c**) Electron-deficient Trp analogs are introduced into nicotinic acetylcholine receptor (nAChR) at position α149 by mutagenesis. The resulting receptors show reduced activation by the agonist epibatidine with increasing fluorination, demonstrating the importance of the Trp149-epibatidine cation–pi interaction

by the dramatic weakening of protein–ligand interactions for the proteins with more highly fluorinated, and therefore more electron-poor aromatics.

Trp is the most electron-rich of the aromatic amino acids, and as such is the aromatic residue most frequently identified as participating in strong cation–pi interactions in nature [63]. In recent work, our group has been creating new aromatic cage mimics that involve the incorporation of Trp into receptor frameworks of different types. One variation on this theme involves the construction of small Trp-rich peptides, with augmentation of their aromatic character by incorporation of *N*-benzyltryptophan as a building block [64]. This artificial amino acid presents both the electron-rich indole and an appended benzyl substituent to cationic binding partners. The peptide-derived receptor Trp(Bn)–Trp(Bn) (**33**) shows the ability to bind to quaternary ammonium ions in water, with selectivity over unmethylated primary quaternary ammonium ions in this highly competitive medium. More telling, the construction and comparison of receptors based on Trp–Trp peptides (**33**–**35**) bearing zero, one, and two appended benzyl substituents showed an increasing ability to bind to acetylcholine in water with increasing benzylation (Fig. 14).

In another approach, we have created a variety of hosts based on indole carboxylic acids, including receptor **37** made from three copies of the de-aminated Trp building block indole-3-propionic acid (**36**, Fig. 15) [65]. Despite its extreme

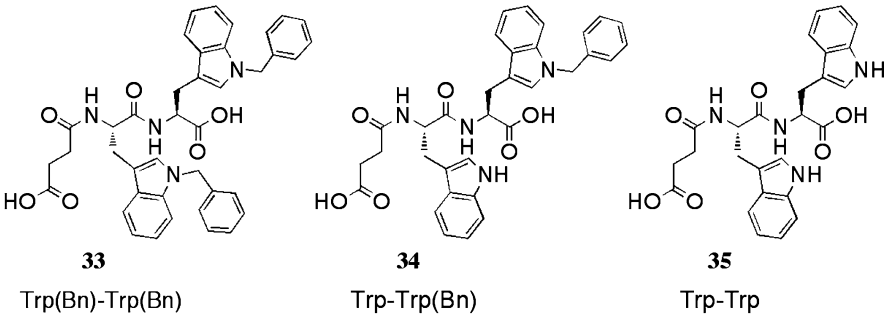


**Fig. 14** Hosts for quaternary ammonium ions based on *N*-benzylated Trp residues

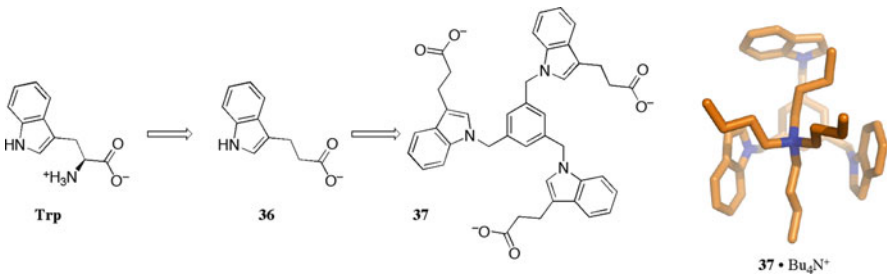


**Fig. 15** A host based on Trp derivative indole-3-propionic acid binds quaternary ammonium ions in buffered water

simplicity and lack of pre-organized binding pocket, host **37** binds quaternary ammonium ions like $Me_4N^+$, AcCh, and Kme3 in buffered water, with two- to fivefold selectivity over some analogous primary and secondary ammonium ions. In this case, however, studies that extended to examination of more hydrophobic cations showed more dramatic increases in affinity, including a >800-fold higher affinity for $Bu_4N^+$ than $Me_4N^+$. This difference is completely absent in $CDCl_3$ (where in fact, $Bu_4N^+$ is not bound at all), demonstrating that it must be largely driven by the hydrophobic effect and not by specific weak interactions between host **37** and $Bu_4N^+$. This lesson—that hydrophobicity can cooperate with cation–pi interactions in aqueous medium—likely extends to naturally occurring aromatic cages that can recognize peptides bearing quaternary trimethyllysine side chains over their unmethylated congeners. As previously stated, the observation that a substrate bearing neutral $–C(Me)_3$ side chains (**32**) binds more weakly to an aromatic cage protein than does the natural partner (**31**) bearing $–N(Me)_3^+$ side chains demonstrates the importance of the cation–pi interaction. But the comparison of the methylated peptide **29** to its unmethylated analog **28** raises a question: all other things being equal, the strength of the cation–pi between the aromatic cage protein and unmethylated **28** should be stronger, because the cation is more compact (and charge dense) and can form shorter cation–pi contacts. But of course,

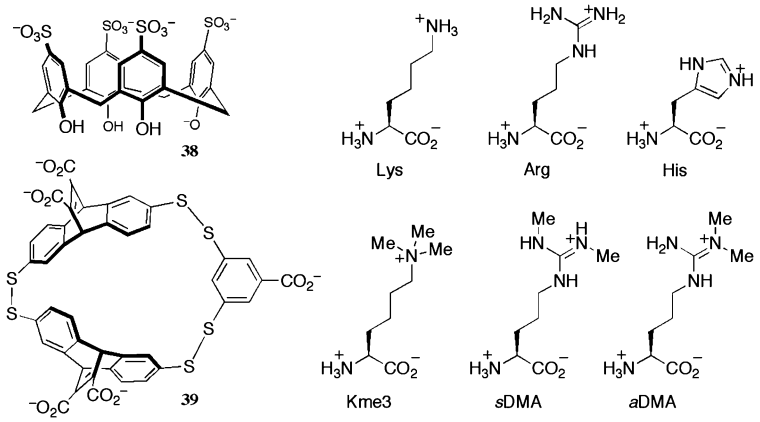


**Fig. 16** *p*-Sulfonatocalix[4]arene (**38**) and macrocycle **39** mimic the aromatic-rich pockets of methyllysine-binding proteins in different ways. Both bind peptides bearing trimethyllysine side chains in buffered water with protein-like affinities and selectivities

the unmethylated peptide **28** has a handicap, in that the primary $-NH_3^+$ cation of its side chain is strongly engaged with a shell of hydrating water molecules. The affinities for these peptides with their partner protein HP1 are provided in Fig. 13 [61]. The hydrophobic effect that is induced upon methylation is an important player that works alongside cation–pi interactions and dispersive interactions in generating the natural selectivities of aromatic cage proteins for their trimethyllysine targets over unmethylated counterparts.

The calixarenes, a family of hosts that present multiple phenol rings in a macrocyclic arrangement, have been widely explored as biomimetic hosts. A particularly large body of work exists on sulfonated calixarenes, which are soluble in water and capable of binding a wide variety of natural and unnatural guests presenting ammonium ions. Hosts like *p*-sulfonatocalix[4]arene (**38**) have long been known to bind biologically important quaternary ammonium ions like acetylcholine [66, 67]. More recent work has broadly explored their abilities to bind cationic amino acids, peptides, and proteins [68, 69]. Host **38** binds lysine, arginine, and histidine in buffered water with affinities of 520, 330, and 20 $M^{-1}$, respectively (Fig. 16) [70–72]. Affinities rise upon methylation, with asymmetric dimethylarginine (*a*DMA) and symmetric dimethylarginine (*s*DMA) binding threefold stronger than unmethylated arginine [72]. The lysine series displays even stronger dependence on methylation state, with affinities ranging up to 37,000 $M^{-1}$ for the 1:1 complexation of trimethyllysine (Kme3) and 96,600 $M^{-1}$ for the short trimethyllysine-containing peptide R-(Kme3)-S-T [72]. Such high selectivities for the quaternary ammonium ions over primary counterparts are dramatic and consistently displayed for this host. The difference probably arises from a change in binding mode: unmethylated lysine side chains bury their most hydrophobic elements, the aliphatic $CH_2$ groups, inside of the aromatic cavity of host **38** and leave the polar $-NH_3^+$ group exposed to the polar sulfonate functional groups and external solvent [73]. Trimethyllysine, instead, buries its $-NMe_3^+$ group

snugly inside the cavity of the host, and in doing so forms multiple strong cation–pi, CH–pi, and dispersive interactions assisted by the hydrophobic effect [72].

Other synthetic hosts have recently been explored as receptors for trimethyllysine, including macrocycle **39**, which contains disulfide bridges that enable its participation in and selection from a dynamically equilibrating host library. This aromatic-rich host also binds trimethyllysine-containing peptides with excellent selectivities, displaying an association constant of 40,000 $M^{-1}$ for a trimethylated peptide and >50-fold weaker binding to the unmethylated analog [74]. The affinities of both calixarene **38** and cyclophane **39** for their trimethyllysine targets are of the same order as the naturally evolved aromatic cage proteins, which typically range from $K_{assoc}$ = 50,000–200,000 $M^{-1}$ ($K_d$ = 5–20 μM). In both of these cases, protein-like affinities and selectivities, which are rarely displayed by supramolecular hosts, are achieved. Both receptors profit from the approach of "teaching old dogs new tricks," i.e., identifying existing host scaffolds that mimic naturally evolved protein binding partners, and using them to engage biological targets that had previously been unconsidered by supramolecular chemistry. The creation of biomimetic receptors for post-translationally methylated protein residues is a promising area for future developments in biotechnology and biomedical research.

## 4 Conclusions

The technological promise of biomimetic receptor-type compounds as both sensors and disruptors of biological pathways is only now beginning to be realized [67, 75] but there remain many challenges to converting this type of biologically inspired receptor into advances that are biomedically important. The most fundamental is that strong and specific molecular recognition in the medium of life—pure, warm, salty water—remains difficult to achieve using the simple scaffolds that are familiar in the world of supramolecular chemistry. Examples of success of the types described here are relatively rare. As we continue to seek simple molecules that can achieve complex recognition tasks, we find an almost inexhaustible source of inspiration for these studies in the huge diversity of proteins and drugs that are known encode strong and selective binding in water.

**Acknowledgements** We thank NSERC for funding. FH is a Career Scholar of the Michael Smith Foundation for Health Research and Canada Research Chair.

## References

1. Perrin CL, Nielson JB (1997) "Strong" hydrogen bonds in chemistry and biology. Annu Rev Phys Chem 48(1):511–544
2. Xu D, Tsai CJ, Nussinov R (1997) Hydrogen bonds and salt bridges across protein–protein interfaces. Protein Eng 10(9):999–1012

3. Wahl MC, Sundaralingam M (1997) C–H. . .O hydrogen bonding in biology. Trends Biochem Sci 22(3):97–102
4. Resnati G (2002) Book review: the weak hydrogen bond—in structural chemistry and biology. Gautam Desiraju and Thomas steiner. ChemPhysChem 3(2):225–226
5. Gao J, Bosco DA, Powers ET, Kelly JW (2009) Localized thermodynamic coupling between hydrogen bonding and microenvironment polarity substantially stabilizes proteins. Nat Struct Mol Biol 16(7):684–690
6. Alkorta I, Elguero J (1998) Non-conventional hydrogen bonds. Chem Soc Rev 27(2):163–170
7. Dutzler R, Campbell EB, Cadene M, Chait BT, MacKinnon R (2002) X-ray structure of a ClC chloride channel at 3.0 Å reveals the molecular basis of anion selectivity. Nature 415 (6869):287–294
8. Thiyagarajan N, Smith BD, Raines RT, Acharya KR (2011) Functional and structural analyses of *N*-acylsulfonamide-linked dinucleoside inhibitors of rnase a. FEBS J 278(3):541–549
9. Serrano L, Bycroft M, Fersht AR (1991) Aromatic-aromatic interactions and protein stability: investigation by double-mutant cycles. J Mol Biol 218(2):465–475
10. Bhattacharyya R, Samanta U, Chakrabarti P (2002) Aromatic-aromatic interactions in and around α-helices. Protein Eng 15(2):91–100
11. Hunter CA, Lawson KR, Perkins J, Urch CJ (2001) Aromatic interactions. J Chem Soc Perkin Trans 2 (5):651–669
12. Ma M, Kuang Y, Gao Y, Zhang Y, Gao P, Xu B (2010) Aromatic-aromatic interactions induce the self-assembly of pentapeptidic derivatives in water to form nanofibers and supramolecular hydrogels. J Am Chem Soc 132(8):2719–2728
13. Spolar RS, Ha JH, Record MT (1989) Hydrophobic effect in protein folding and other noncovalent processes involving proteins. Proc Natl Acad Sci USA 86(21):8382–8385
14. Lins L, Brasseur R (1995) The hydrophobic effect in protein folding. FASEB J 9(7):535–540
15. Southall NT, Dill KA, Haymet ADJ (2001) A view of the hydrophobic effect. J Phys Chem B 106(3):521–533
16. Tsai C-J, Lin SL, Wolfson HJ, Nussinov R (1997) Studies of protein–protein interfaces: a statistical analysis of the hydrophobic effect. Protein Sci 6(1):53–64
17. Bachmann IM, Halvorsen OJ, Collett K, Stefansson IM, Or S, Haukaas SA, Salvesen HB, Otte AP, Akslen LA (2006) EZH2 expression is associated with high proliferation rate and aggressive tumor subgroups in cutaneous melanoma and cancers of the endometrium, prostate, and breast. J Clin Oncol 24(2):268–273
18. Eckerle S, Brune V, Doring C, Tiacci E, Bohle V, Sundstrom C, Kodet R, Paulli M, Falini B, Klapper W, Chaubert AB, Willenbrock K, Metzler D, Brauninger A, Kuppers R, Hansmann ML (2009) Gene expression profiling of isolated tumour cells from anaplastic large cell lymphomas: insights into its cellular origin, pathogenesis and relation to Hodgkin lymphoma. Leukemia 23(11):2129–2138
19. Fiskus W, Wang Y, Sreekumar A, Buckley KM, Shi H, Jillella A, Ustun C, Rao R, Fernandez P, Chen J, Balusu R, Koul S, Atadja P, Marquez VE, Bhalla KN (2009) Combined epigenetic therapy with the histone methyltransferase EZH2 inhibitor 3-deazaneplanocin A and the histone deacetylase inhibitor panobinostat against human AML cells. Blood 114(13): 2733–2743
20. Basak SK, Veena MS, Oh S, Huang G, Srivatsan E, Huang M, Sharma S, Batra RK (2009) The malignant pleural effusion as a model to investigate intratumoral heterogeneity in lung cancer. PLoS One 4(6):e5884
21. Kaustov L, Ouyang H, Amaya M, Lemak A, Nady N, Duan S, Wasney GA, Li Z, Vedadi M, Schapira M, Min J, Arrowsmith CH (2011) Recognition and specificity determinants of the human CBX chromodomains. J Biol Chem 286(1):521–529
22. Gribkoff VK, Champigny G, Barbry P, Dworetzky SI, Meanwell NA, Lazdunski M (1994) The substituted benzimidazolone NS004 is an opener of the cystic fibrosis chloride channel. J Biol Chem 269(15):10983–10986

23. Cox JPD, Yamamoto K, Christie PT, Wooding C, Feest T, Flinter FA, Goodyer PR, Leumann E, Neuhaus T, Reid C, Williams PF, Wrong O, Thakker RV (1999) Renal chloride channel, CLCN5, mutations in Dent's disease. J Bone Miner Res 14(9):1536–1542
24. Schlingmann KP, Konrad M, Jeck N, Waldegger P, Reinalter SC, Holder M, Seyberth HrW, Waldegger S (2004) Salt wasting and deafness resulting from mutations in two chloride channels. N Engl J Med 350(13):1314–1319
25. Beer PD, Gale PA (2001) Anion recognition and sensing: the state of the art and future perspectives. Angew Chem Int Ed 40(3):486–516
26. Gale PA, Quesada R (2006) Anion coordination and anion-templated assembly: highlights from 2002 to 2004. Coord Chem Rev 250(23–24):3219–3244
27. Lavigne JJ, Anslyn EV (2001) Sensing a paradigm shift in the field of molecular recognition: from selective to differential receptors. Angew Chem Int Ed 40(17):3118–3130
28. Herr RJ (2002) 5-Substituted-1H-tetrazoles as carboxylic acid isosteres: medicinal chemistry and synthetic methods. Bioorg Med Chem 10(11):3379–3393
29. Patani GA, LaVoie EJ (1996) Bioisosterism: a rational approach in drug design. Chem Rev 96(8):3147–3176
30. Olesen PH (2001) The use of bioisosteric groups in lead optimization. Curr Opin Drug Discov Devel 4:471–478
31. Myznikov LV, Hrabalek A, Koldobskii GI (2007) Drugs in the tetrazole series. Chem Heterocycl Compd 43(1):1–9
32. Nattel S (2005) Can losartan prevent new-onset atrial fibrillation in hypertensive patients more effectively than atenolol? Nat Clin Pract Cardiovasc Med 2(7):332–333
33. Skulnick HI, Johnson PD, Aristoff PA, Morris JK, Lovasz KD, Howe WJ, Watenpaugh KD, Janakiraman MN, Anderson DJ, Reischer RJ, Schwartz TM, Banitt LS, Tomich PK, Lynn JC, Horng M-M, Chong K-T, Hinshaw RR, Dolak LA, Seest EP, Schwende FJ, Rush BD, Howard GM, Toth LN, Wilkinson KR, Kakuk TJ, Johnson CW, Cole SL, Zaya RM, Zipp GL, Possert PL, Dalga RJ, Zhong W-Z, Williams MG, Romines KR (1997) Structure-based design of nonpeptidic HIV protease inhibitors: the sulfonamide-substituted cyclooctylpyranones. J Med Chem 40(7):1149–1164
34. Reddy NS, Mallireddigari MR, Cosenza S, Gumireddy K, Bell SC, Reddy EP, Reddy MVR (2004) Synthesis of new coumarin 3-(*N*-aryl) sulfonamides and their anticancer activity. Bioorg Med Chem Lett 14(15):4093–4097
35. Lobb KL, Hipskind PA, Aikins JA, Alvarez E, Cheung Y-Y, Considine EL, De Dios A, Durst GL, Ferritto R, Grossman CS, Giera DD, Hollister BA, Huang Z, Iversen PW, Law KL, Li T, Lin H-S, Lopez B, Lopez JE, Cabrejas LMM, McCann DJ, Molero V, Reilly JE, Richett ME, Shih C, Teicher B, Wikel JH, White WT, Mader MM (2004) Acyl sulfonamide anti-proliferatives: benzene substituent structure, activity relationships for a novel class of antitumor agents. J Med Chem 47(22):5367–5380
36. Stansfield I, Pompei M, Conte I, Ercolani C, Migliaccio G, Jairaj M, Giuliano C, Rowley M, Narjes F (2007) Development of carboxylic acid replacements in indole-*N*-acetamide inhibitors of hepatitis C virus NS5B polymerase. Bioorg Med Chem Lett 17(18):5143–5149
37. Hu X, Sun J, Wang H-G, Manetsch R (2008) Bcl-xl-templated assembly of its own protein-protein interaction modulator from fragments decorated with thio acids and sulfonyl azides. J Am Chem Soc 130(42):13820–13821
38. Mahnke DJ, McDonald R, Hof F (2007) A shape-dependent hydrophobic effect for tetrazoles. Chem Commun (36):3738–3740
39. Tominey A, Andrew D, Oliphant L, Rosair GM, Dupre J, Kraft A (2006) Supramolecular binding of protonated amines to a receptor microgel in aqueous medium. Chem Commun (23):2492–2494
40. Peters L, Froehlich R, Boyd ASF, Kraft A (2001) Noncovalent interactions between tetrazole and an N, N-diethyl-substituted benzamidine. J Org Chem 66(10):3291–3298
41. Khazaei A, Manesh AA (2006) Selective oxidation of oximes to carbonyl compounds using n-bromo-n-benzoyl-4-toluenesulfonamide. Mendeleev Commun 16(2):109–111

42. Chen Y-J, Chung W-S (2009) Tetrazoles and para-substituted phenylazo-coupled calix[4] arenes as highly sensitive chromogenic sensors for ca2+. Eur J Org Chem 2009 (28):4770–4776
43. Boyko V, Rodik R, Danylyuk O, Tsymbal L, Lampeka Y, Suwinska K, Lipkowski J, Kalchenko V (2005) Tetrazolecalix[4]arenes as new ligands for palladium(ii). Tetrahedron 61(52):12282–12287
44. Starnes SD, Arungundram S, Saunders CH (2002) Anion sensors based on β, β′-disubstituted porphyrin derivatives. Tetrahedron Lett 43(43):7785–7788
45. Chung YM, Raman B, Kim D-S, Ahn KH (2006) Fluorescence modulation in anion sensing by introducing intramolecular H-bonding interactions in host-guest adducts. Chem Commun (2):186–188
46. Carroll CN, Naleway JJ, Haley MM, Johnson DW (2010) Arylethynyl receptors for neutral molecules and anions: emerging applications in cellular imaging. Chem Soc Rev 39 (10):3875–3888
47. Berryman OB, Johnson CA, Zakharov LN, Haley MM, Johnson DW (2008) Water and hydrogen halides serve the same structural role in a series of 2+2 hydrogen-bonded dimers based on 2,6-bis(2-anilinoethynyl)pyridine sulfonamide receptors. Angew Chem Int Ed 47(1):117–120
48. Kavallieratos K, Bertao CM, Crabtree RH (1999) Hydrogen bonding in anion recognition: a family of versatile, nonpreorganized neutral and acyclic receptors. J Org Chem 64(5): 1675–1683
49. Shchipanov VP, Krashina KI, Skachilova AA (1973) Tetrazole derivatives. Chem Heterocycl Compd 9(11):1423–1426
50. Seydel JK (1968) Sulfonamides, structure-activity relationship, and mode of action. Structural problems of the antibacterial action of 4-aminobenzoic acid (paba) antagonists. J Pharm Sci 57(9):1455–1478
51. Hof F, Schütz A, Fäh C, Meyer S, Bur D, Liu J, Goldberg DE, Diederich F (2006) Starving the malaria parasite: inhibitors active against the aspartic proteases plasmepsins I, II, and IV. Angew Chem Int Ed 45(13):2138–2141
52. McKie AH, Friedland S, Hof F (2008) Tetrazoles are potent anion recognition elements that emulate the disfavored anti conformations of carboxylic acids. Org Lett 10(20):4653–4655
53. Wallace KJ, Belcher WJ, Turner DR, Syed KF, Steed JW (2003) Slow anion exchange, conformational equilibria, and fluorescent sensing in venus flytrap aminopyridinium-based anion hosts. J Am Chem Soc 125(32):9699–9715
54. Davis J (2010) Anion binding and transport by prodigiosin and its analogs anion recognition in supramolecular chemistry. In: Gale PA, Dehaen W (eds) Topics in heterocyclic chemistry, vol 24. Springer, Berlin, pp 145–176
55. Courtemanche RJM, Pinter T, Hof F (2011) Just add tetrazole: 5-(2-Pyrrolo)tetrazoles are simple, highly potent anion recognition elements. Chem Commun 47 (47):12688–12690
56. Gale PA, Camiolo S, Tizzard GJ, Chapman CP, Light ME, Coles SJ, Hursthouse MB (2001) 2-Amidopyrroles and 2,5-diamidopyrroles as simple anion binding agents. J Org Chem 66(23):7849–7853
57. Campagna-Slater V, Schapira M (2010) Finding inspiration in the Protein Data Bank to chemically antagonize readers of the histone code. Mol Inform 29(4):322–331
58. Taverna SD, Li H, Ruthenburg AJ, Allis CD, Patel DJ (2007) How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. Nat Struct Mol Biol 14(11):1025–1040
59. Ma JC, Dougherty DA (1997) The cation-π interaction. Chem Rev 97(5):1303–1324
60. Schärer K, Morgenthaler M, Paulini R, Obst-Sander U, Banner DW, Schlatter D, Benz J, Stihle M, Diederich F (2005) Quantification of cation–π interactions in protein–ligand complexes: crystal-structure analysis of Factor Xa bound to a quaternary ammonium ion ligand. Angew Chem Int Ed 44(28):4400–4404

61. Hughes RM, Wiggins KR, Khorasanizadeh S, Waters ML (2007) Recognition of trimethyllysine by a chromodomain is not driven by the hydrophobic effect. Proc Natl Acad Sci USA 104(27):11184–11188
62. Cashin AL, Petersson EJ, Lester HA, Dougherty DA (2004) Using physical chemistry to differentiate nicotinic from cholinergic agonists at the nicotinic acetylcholine receptor. J Am Chem Soc 127(1):350–356
63. Gallivan JP, Dougherty DA (1999) Cation-π interactions in structural biology. Proc Natl Acad Sci USA 96(17):9459–9464
64. Beshara CS, Hof F (2010) Modular incorporation of 1-benzyltryptophan into dipeptide hosts that bind acetylcholine in pure water. Can J Chem 88(10):1009–1016
65. Whiting AL, Neufeld NM, Hof F (2009) A tryptophan-analog host whose interactions with ammonium ions in water are dominated by the hydrophobic effect. Tetrahedron Lett 50(50):7035–7037
66. Koh KN, Araki K, Ikeda A, Otsuka H, Shinkai S (1996) Reinvestigation of calixarene-based artificial-signaling acetylcholine receptors useful in neutral aqueous (water/methanol) solution. J Am Chem Soc 118(4):755–758
67. Guo D-S, Uzunova VD, Su X, Liu Y, Nau WM (2011) Operational calixarene-based fluorescent sensing systems for choline and acetylcholine and their application to enzymatic reactions. Chem Sci 2(9):1722–1734
68. Perret F, Lazar AN, Coleman AW (2006) Biochemistry of the *para*-sulfonato-calix[n]arenes. Chem Commun (23):2425–2438
69. Coleman A, Perret F, Moussa A, Dupin M, Guo Y, Perron H (2007) Calix[n]arenes as protein sensors. Creative chemical sensor systems. In: Schrader T (ed) Topics in current chemistry, vol 277. Springer, Berlin, pp 31–88
70. Arena G, Casnati A, Contino A, Magri A, Sansone F, Sciotto D, Ungaro R (2006) Inclusion of naturally occurring amino acids in water soluble calix[4]arenes: a microcalorimetric and $^{1}$H NMR investigation supported by molecular modeling. Org Biomol Chem 4(2):243–249
71. Douteau-Guevel N, Coleman AW, Morel J-P, Morel-Desrosiers N (1999) 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 (3):629–634
72. Beshara CS, Jones CE, Daze KD, Lilgert BJ, Hof F (2010) A simple calixarene recognizes post-translationally methylated lysine. Chembiochem 11(1):63–66
73. Selkti M, Coleman AW, Nicolis I, Douteau-Guevel N, Villain F, Tomas A, de Rango C (2000) The first example of a substrate spanning the calix[4]arene bilayer: the solid state complex of *p*-sulfonatocalix[4]arene with -lysine. Chem Commun (2):161–162
74. Ingerman LA, Cuellar ME, Waters ML (2010) A small molecule receptor that selectively recognizes trimethyl lysine in a histone peptide with native protein-like affinity. Chem Commun 46(11):1839–1841
75. Wang K, Guo D-S, Zhang H-Q, Li D, Zheng X-L, Liu Y (2009) Highly effective binding of viologens by *p*-sulfonatocalixarenes for the treatment of viologen poisoning. J Med Chem 52 (20):6402–6412

http://www.springer.com/978-3-642-32328-7

Designing Receptors for the Next Generation of Biosensors

Piletsky, S.A.; Whitcombe, M.J. (Eds.)
2013, XII, 264 p., Hardcover
ISBN: 978-3-642-32328-7