ACCOUNTS
of chemical research

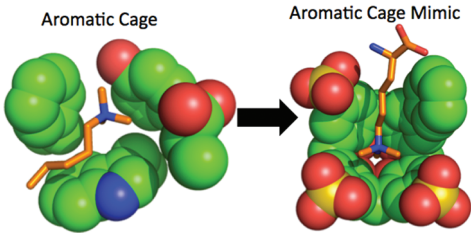
The Cation–π Interaction at Protein–Protein Interaction Interfaces: Developing and Learning from Synthetic Mimics of Proteins That Bind Methylated Lysines

KEVIN D. DAZE AND FRASER HOF*

Department of Chemistry, University of Victoria, 3800 Finnerty Road, Victoria, British Columbia V8W 3V6, Canada

RECEIVED ON MARCH 6, 2012

CONSPECTUS



First discovered over 60 years ago, post-translational methylation was considered an irreversible modification until the initial discoveries of demethylase enzymes in 2004. Now researchers understand that this process serves as a dynamic and complex control mechanism that is misregulated in numerous diseases. Lysine methylation is most often found on histone proteins and can effect gene regulation, epigenetic inheritance, and cancer. Because of this connection to disease, many enzymes responsible for methylation are considered targets for new cancer therapies. Although our understanding of the biology of post-translational methylation has advanced at an astonishing rate within the last 5 years, chemical approaches for studying and disrupting these pathways are only now gaining momentum.

In general, enzymes methylate lysine and arginine residues with very high specificity for both the location and methylation state. Each methylated target serves as the focused hot spot for an inducible protein–protein interaction (PPI). Conceptually, lysine or arginine methylation is a subtle modification that leads to no change in charge and small changes in size, but it significantly alters the hydration energies and hydrogen bonding potential of these side chains. Nature has evolved a special motif for recognizing the methylation states of lysine, called the "aromatic cage", a collection of aromatic protein residues, often accompanied by one or more neighboring anionic residues. The combination of favorable cation–π, electrostatic, and van der Waals interactions, as well as size matching, gives these proteins a high degree of specificity for the methylation state.

This Account summarizes the development of various supramolecular host system scaffolds developed to recognize and bind to ammonium cations, such as trimethyllysine, on the basis of their methylation state. Early systems bound to their targets in pure, buffered water but failed to achieve biochemically relevant affinities and selectivities. Surprisingly, the use of the simple and very well-known *p*-sulfonatocalix[4]arene provides protein-like affinities and selectivities for trimethyllysine in water. New analogs, created by synthetic modification of the same scaffold, allow for further tuning of affinities and selectivities for trimethyllysine. Our studies of each family of hosts paint a consistent picture: cation–π interactions and electrostatics are important, and solvation effects are complex. Rigidity is especially important for host–guest systems that function in pure water. Despite their simplicity, synthetic systems that take these lessons into account can achieve affinities that rival or surpass those of their naturally evolved counterparts. The stage is now set for the next act: the use of such compounds as tunable and adaptable tools for modern chemical biology.

Introduction

Cationic amino acids play a critical role in protein structure, engaging with acidic and aromatic amino acids (aspartic acid and glutamic acid and tryptophan, tyrosine, and phenylalanine, respectively) through favorable electrostatic[1] and cation–π interactions.[2] Surveys have shown that cation–π pairs are particularly important at protein–protein interaction (PPI) interfaces, appearing in nearly half of all examples that have been studied.[2] While arginine is relatively over-represented at protein interaction interfaces, its cationic counterpart lysine is notably absent from the list of residues that are important at protein interaction hot spots.[2,3] But nature has evolved a separate path that brings lysine into a role as an important player in protein–protein interactions.

Published on the Web 06/22/2012 www.pubs.acs.org/accounts
10.1021/ar300072g

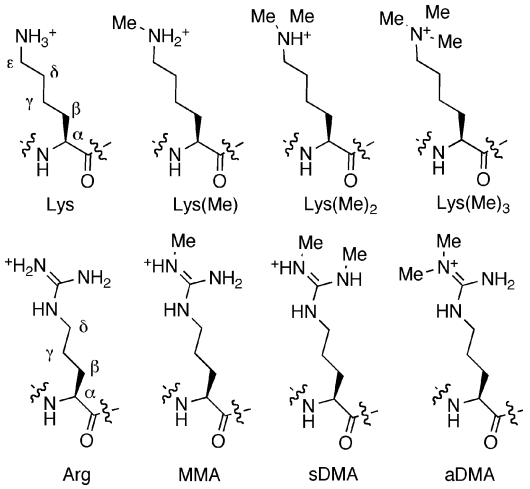


FIGURE 1. Known methylation states of lysine and arginine (MMA = monomethylarginine; sDMA = symmetric dimethylarginine; aDMA = asymmetric dimethylarginine).

A diverse set of lysine post-translational modifications expand the toolkit of chemical properties beyond those of the ribosomally translated amino acid, and almost all of them exist exclusively to trigger PPIs. Well known examples include lysine acetylation, which turns on binding by bromodomains,[4] and ubiquitination, which can recruit various partners that are involved in gene regulation.[5] Of special interest in a discussion of aromatic interactions are the post-translational methylations of lysine.[6,7] Lysine can exist in three distinct methylation states (mono-, di-, and trimethyllysine), which are under the control of highly specific methyl transferases or demethylases. Arginine is also methylated by a complementary set of enzymes (Figure 1). Post-translational methylation of lysine and arginine can be considered a minimalist set of covalent changes for a variety of reasons: the addition of CH_3 is the smallest possible change to a protein's structure; each side chain remains positively charged in all modified states; and most methylation states retain some hydrogen bond donating ability (with the exception of trimethyllysine). Almost all known examples of post-translational methylation serve to turn on new protein–protein interactions with partners that specifically recognize the methylated side chain.[6,8] These interactions have diverse downstream cellular effects, but are most closely associated with histone proteins and gene regulation pathways.[8,9] How do these seemingly small changes become the dominant hot-spots that trigger protein–protein interactions?

Structural Aspects and Interactions

There are several protein domains identified that recognize post-translationally methylated lysine or arginine (or both). The methylated lysine recognition domains tudor, MBT, chromo, and PWWP are collectively referred to as the "royal family".[10,11] All proteins containing domains of the "royal family" are implicated in chromatin and genetic control.[12] Tudor domains, in particular, can be categorized base on their binding of either methylated arginine or methylated lysine. Tudor domains that bind methylated arginine are typically associated with RNA metabolism and those that bind methylated lysine residues are associated with chromatin compaction and post-translational modifications on histone proteins.[12,13] Plant homeobox domains (PHD domains)[14] and chromodomains[15] are small α/β mixed domains that bind to chromatin-related methylated lysine residues. PWWP domains bind methylated lysine residues within a central β-barrel.[16] Proteins consisting of multiple WD40 repeats form a β-propeller structure that binds methylated lysines through its central pore.[17] MBT domains recognize lower methylation states of lysine through two distinct mechanisms: using a deep, buried aromatic pocket with specific size constraints and positioning a favorable anionic complement near the pocket to make hydrogen bond contact with the methylammonium group.[18] Despite the evolutionary and structural diversity among these proteins, they all engage methylated amino acid side chains using a common motif: the aromatic cage.

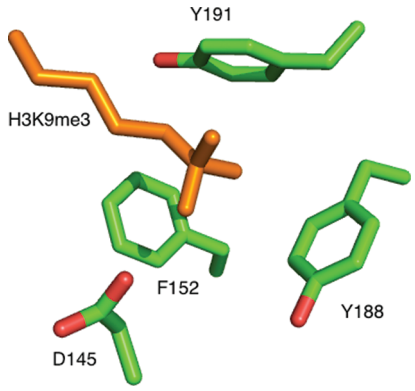


FIGURE 2. Aromatic cage recognition of a trimethyllysine-containing peptide (PHD domain, PDB id 3ASK).

The "aromatic cage" is a preorganized collection of aromatic amino acids that coordinate to produce a desolvated, highly π-electron-rich "cage" occasionally containing an adjacent anionic residue (Figure 2).[13,19] Recognition and affinity is mainly derived from cation–π interactions between the positively charged cationic side chain and the electron-rich π-surfaces of one or more nearby aromatic rings. Recent work has shown, computationally, that replacement of trimethyllysine with a neutral, isosteric analogue (*tert*-butyl norleucine) that cannot form cation–π

interactions results in a 3.1 kcal mol^{-1} penalty in binding.[20] Prior work using experimental systems based on both the trimethyllysine-binding HP1 domain and an aromatic cage model system constructed using a β-hairpin peptide also showed that replacement of trimethyllysine with the same neutral analogue severely weakened affinity.[21] The hydrophobic effect also plays a role, because the cationic (and somewhat hydrophobic) methylated ammonium groups are desolvated as they bind into the aromatic cage. Studies of ING4 (a PHD-containing protein that binds methyllysine partners) has shown that binding to a peptide containing a methylated lysine is driven overall by large favorable enthalpic contributions and opposed by smaller unfavorable entropic costs.[22] When the methylation state of the key lysine residue was considered, it was found that the selectivity for trimethyllysine was entropically driven, suggesting a unique role for the hydrophobic effect in driving selectivities among similar partners. Further evidence for the contributions of these interactions in methyllysine recognition is reported in a growing list of papers.[18,23,24]

Chemical Mimics of Methyllysine-Binding Proteins

Efforts to create chemicals that recognize PPI elements can teach fundamental lessons about how these surface motifs balance the competing needs of encoding molecular recognition and preserving protein solubility. These chemicals can also serve as disruptors of their targeted PPIs, which permit their use in both basic studies of chemical biology and as novel therapeutics. The first small molecules that target and bind to aromatic cage proteins have recently been reported.[25,26] Agents that bind to the methyllysine-containing partners in this family of PPIs would provide complementary information. Unfortunately, trimethyllysine residues are almost always present on flexible, unstructured protein tails and as such do not present any concave binding pocket of the type typically exploited for small-molecule intervention. We, and others, have proposed a supramolecular approach using host scaffolds with defined binding pockets. A suitable concave host could simulate the aromatic cage motif, while negatively charged groups could provide favorable electrostatic interactions known to also play a role in the natural complexes.[18]

This idea arises naturally from a large body of work, too vast to summarize here, on ammonium-ion-binding supramolecular hosts. Where supramolecular approaches have been used to bind ammonium ions of biological importance, the targets have most often been cationic neurotransmitters such as catecholamines, nicotine, and especially cholines. The structural similarity of cholines to trimethyllysine (they both contain $RNMe_3^+$ functional groups) provides an obvious starting point for this new work, where the biological targets are post-translationally methylated amino acids, peptides, and proteins (and the interactions they encode), rather than small molecules. The constant challenge in the currently reported work and in all biologically motivated supramolecular chemistry is to arrive at hosts that operate in biologically relevant solutions (sometimes referred to as "pure, warm, salty water") with affinities and selectivities that are relevant to the biochemistry under study. What follows is a brief account of our own group's recent work in the area. We report on a variety of motifs that have succeeded to greater or lesser extents against these stringent criteria and on lessons learned along the way.

Tetrazolates and Indole Carboxylates

To engage the hydrophobic cation of trimethyllysine a supramolecular system is required that has rigidity, water solubility and some combination of aromatic and anionic character to encourage cation–π interactions and electrostatic attraction. To have some selectivity for a more highly methylated ammonium ion over a less methylated analog, some degree of hydrophobicity is required. We explored two heterocyclic systems, tetrazolates and indoles functionalized with carboxylates, that satisfy these requirements in different ways. Tetrazoles are acidic, aromatic heterocycles finding increased use in medicinal chemistry as metabolically stable replacements for carboxylic acids.[27] In terms of molecular recognition, the deprotonated, anionic tetrazolates are more hydrophobic than their carboxylate counterparts, and we anticipated that these characteristics would make them good complements to hydrophobic quaternary ammonium cations like the side chain of trimethyllysine.[28,29] Indoles are electron-rich aromatic rings that, as the side chains of tryptophan residues,[30,31] are key participants in the vast majority of natural protein–trimethyllysine complexes. To use them as recognition elements in synthetic hosts, we appended them with carboxylates or additional aromatic functionality in a variety of ways in an effort to provide for attractive electrostatic and cation–π interactions and water solubility.

We first used 1,3,5-trisubstituted benzene scaffolds in order to allow for the installation of multiple recognition elements and to provide them with some degree of organization.[32,33] Tetrazolates (and carboxylates, as a point of comparison) were appended onto the well-known 2,4,6-triethyl benzene scaffold[32,34] to give compounds

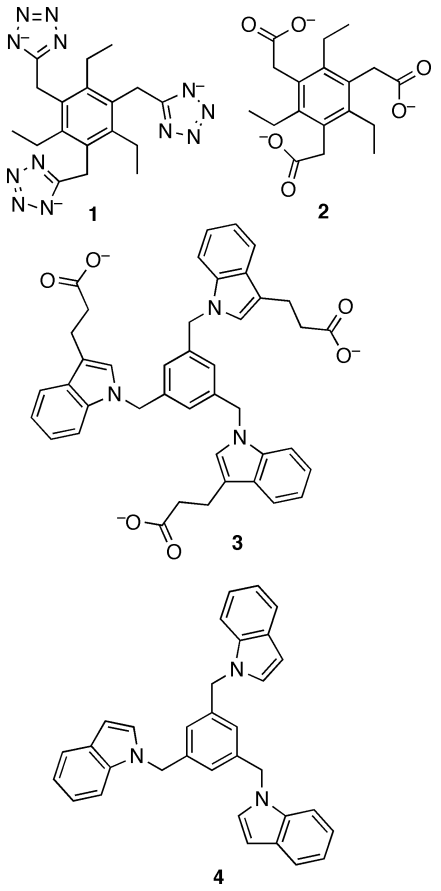


FIGURE 3. Hosts explored based on 1,3,5-trisubstituted benzene scaffold.

1 and **2** (Figure 3).[35] Tetrazolates have previously been explored as recognition elements that can form salt bridges with amidinium-type cations in organic solvents.[36] We found that in mixed organic/aqueous solvent, the tetrazoles of host **1** also form energetically stable salt bridges with ammonium ions and therefore bind ammonium ions of type RNH_3^+ more strongly than quaternary ammonium ions. But in pure water, the selectivities change. Host **1** binds Me_4N^+ with higher affinity than it does RNH_3^+. Its carboxylate counterpart **2** does not measurably bind the same quaternary ammonium ion. Further, NMR shifts upon formation of the **1**·Me_4N^+ complex in pure water revealed that the interaction between the tetrazolates and quaternary ammonium ion occurs via a face-on orientation of the tetrazolate. This is consistent with our understanding of these binding elements, because this orientation maximizes dispersive and hydrophobic interactions between the tetrazolates' π-electrons and the polarizable, hydrophobic quaternary ammonium ion.

To explore the use of indole carboxylates (Figure 3) and associated solvent effects on similar scaffolding, the cation

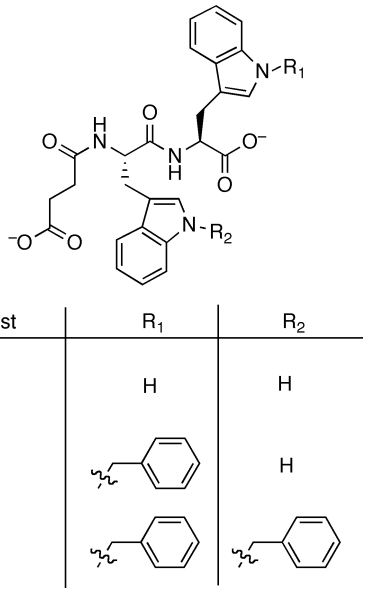


FIGURE 4. Trp and Trp(Bn) hosts explored as dipeptides.

binding properties of host **3** and its neutral counterpart **4** were studied in different solvent systems.[37] NMR titration studies of **3** and **4** with a variety of methylated ammonium cations, as well as acetylcholine and trimethyllysine, provided insight into the nature of interactions occurring in this host–guest system. Host **3**, in pure, buffered water, bound acetylcholine and trimethyllysine (K_d = 23.8 and 15.9 mM) with modest affinities. Again, increasing methylation tracked with increasing affinities in pure water for host **3**. When affinities were compared between **3** and **4**, the trend of increased affinity with increasing methylation state was not observed. In this instance, we also investigated the affinities of non-natural Et_4N^+, Pr_4N^+, and Bu_4N^+ and found dramatic increases in affinity, reaching a strong K_d of 41.7 μM for the complex of **3** and Bu_4N^+ in pure water. The structure/function and solvent effects for this system are consistent with a dominant role for the hydrophobic effect and only a small impact of cation–π interactions.

We sought a different scaffold for further studies of the indole carboxylate motifs. Previous work using dipeptides as synthetic hosts has shown that they can bind ammonium cations in chloroform.[38] Moderate affinity between a tyrosine-based dipeptide host and acetylcholine were measured in chloroform (K_d = 7.69 mM). We created new hosts based on this synthetic scaffold with two key features: the addition of carboxylates for improved electrostatic attraction and water solubility and the introduction of indole-*N*-benzyl substituents as additional aromatic surfaces for cation binding (Figure 4).[39] This resulted in the hosts **5** [Trp-Trp], **6** [Trp(Bn)-Trp], and **7** [Trp(Bn)-Trp(Bn)]. NMR binding studies

in buffered water showed the Trp(Bn)-Trp(Bn) host **7** was able to bind acetylcholine in water ($K_d = 71.4$ mM) and that binding was dependent on the presence of the introduced benzyl groups. Again, the binding measured was weak.

Each of these systems displayed the general trend of binding more highly methylated ammonium ions more strongly. They fail to reproduce the low-micromolar affinities displayed by naturally evolved binders of trimethyllysine. More importantly, the studies of solvent effects in each synthetic system highlighted that they rely more on the hydrophobic effect than the cation−π interaction for binding and selectivity. NMR studies in each case suggest that each of these hosts possess a large degree of conformational freedom and that indole-based hosts **3**−**7** tend to undergo hydrophobic collapse or self-aggregation in pure water. In general we observe that flexibility is very detrimental to affinities and that these effects are greater in water than in organic solvents. Can we solve this problem of flexibility to achieve biochemically relevant affinities of trimethylammonium groups in water?

Macrocycles: Disulfide Linked

The systems described above failed to supply strong binding due to inherent rotational freedom of binding elements and collapsibility of the binding pocket. We sought a scaffold that is inherently rigid. There are a significant number of macrocyclic systems that have proven to be effective supramolecular hosts in pure water. The existence of the cation−π interaction has been demonstrated using various aromatic macrocycles,[40,41] and a family of aromatic sulfur-linked dynamic combinatorial macrocycles provide an example of hosts capable of high-affinity interactions with acetylcholine and various ammonium cations.[42,43] These systems produced hosts that engage their ammonium guest via cation−π interactions much more so than electrostatics, because the charged elements are presented outside the aromatic cage cavity (Figure 5A).[43] Recently one of these dynamic combinatorial systems was used to screen for the creation of trimethyllysine binders.[44] Three to four building blocks capable of forming several different macrocycles under equilibrating conditions were treated with a trimethyllysine-containing dipeptide as a template. This protocol amplified and produced from the mixture two trimethyllysine-selective hosts (represented by **8**). Not surprisingly, these hosts were some that had been previously identified as strong binders of acetylcholine in a prior study.[43] These hosts displayed protein-like (low-micromolar) affinities for trimethyllysine-containing peptides, 2-fold selectivity over dimethylated lysine, and even higher selectivities over unmethylated lysine. These results demonstrate that rigid, organized aromatic systems are critical to provide strong affinities and methylation state selectivity.

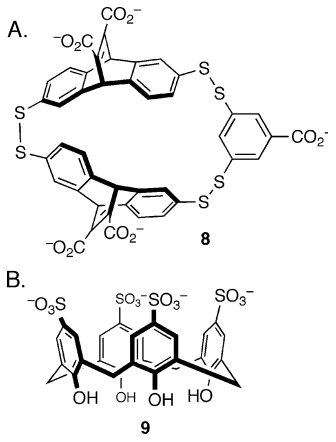


FIGURE 5. (A) Previously explored macrocyclic acetylcholine[43] and trimethyllysine[44] host, **8**. (B) *p*-Sulfonatocalix[4]arene host **9**.

Macrocycles: Calixarenes

Calixarenes are multiaromatic macrocycles whose use as supramolecular hosts is well studied.[45,46] This sustained interest is owed to their ease of synthesis, control of size (calix[*x*]arene, $x = 4$, 5, 6, and 8), well-defined cavity, and the amenability of the upper and lower rim positions to synthetic modifications that tune binding properties. Sulfonated calixarenes have been used as hosts for ammonium ions of all types; a survey of the extensive literature is impossible here. The portion of this body of work most relevant to the discussion here involves their use as hosts for cationic neurotransmitters, amino acids, and peptides.[47−50,60] Previous work has found **9** to complex cationic amino acids with moderate affinity ($K_d = 0.66$ and 1.36 mM, for arginine and lysine).[51] These complexation events arose from both favorable enthalpies and favorable entropies, and the relative strength of the enthalpic contributions suggested a large role for charge−charge interactions in the formation of the complexes. Other investigators have shown that **9** binds to cholines, which bear an $RNMe_3^+$ group analogous to the side chain of trimethyllysine, with strong K_d values in the 2.5 μM range.[47,52−54] These results naturally lead us to explore whether these simple sulfonated macrocycles could bind post-translationally modified amino acids such as methylated lysines and arginines.[50] We found that **9** is very well suited to bind trimethyllysine, displaying a K_d of 27.0 μM, with 70-fold selectivity over lysine, >100-fold selectivity over arginine, and even higher selectivity over all other

TABLE 1. Selectivities and Salt Effects Observed for the Binding of **9** to Various Methylation States of Lysine As the Free Amino Acid and in the Context of a Short Peptide

guest	K_d (μM)[a]	medium
Lys(Me)	333 ± 189	PB
Lys(Me)	769 ± 15	PBS
$Lys(Me)_2$	95.2 ± 18.1	PB
$Lys(Me)_2$	323 ± 11	PBS
$Lys(Me)_3$	28.0 ± 2.0	PB
$Lys(Me)_3$	66.7 ± 2.1	PBS
RKST	182 ± 33	PB
RKST	909 ± 4	PBS
$RK(Me)_3ST$	10.4 ± 1.1	PB
$RK(Me)_3ST$	36.1 ± 0.7	PBS

[a]Determined by ITC at 30 °C. PB = 40 mM Na_2HPO_4/NaH_2PO_4, pH 7.4. PBS = 40 mM Na_2HPO_4/NaH_2PO_4, pH 7.4, 137 mM NaCl, 2.7 mM KCl. Results are averages of 2–3 replicate titrations.

known unmodified amino acids.[51,55] Methylation of arginine improves affinity to a lesser degree, meaning that **9** is >30-fold selective for trimethyllysine over the most highly methylated states of arginine. To gain insight into the system, we further analyzed our NMR titration data and observed that methylated arginine analogs did not bind tightly into the calixarene cavity (based upon $\Delta\delta$ of shifts during titration) but instead adopted a side-on binding mode (with side-chain CH_2 groups deepest within the cavity) like that observed for lysine.[49] Modeling of trimethyllysine bound to **9** produced a clear difference, with *N*-methyl groups buried in the pocket in an overall end-on structure, consistent with our NMR data. Our observed K_d of 10 μM for a short peptide containing trimethyllysine is similar to the affinities of the naturally evolved trimethyllysine binding partners.[21,56] The thermodynamics of these binding events, as measured by isothermal titration calorimetry (ITC), are consistent with previously reported thermodynamics of sulfonatocalixarene based host–guest systems.[51,57]

Our prior work focused on using solvent changes to understand the forces that drive binding. The addition of salt to aqueous recognition systems is a related experiment that should screen and weaken electrostatic interactions and increase the power of hydrophobic interactions. Subjecting host **9** to physiological salt concentrations (137 mM NaCl, 2.7 mM KCl) was detrimental to affinities for a variety of methylated lysine guests (Table 1), in each case causing ~3-fold loss in affinity. We also studied the effect of varying temperature on the complexation of **9** and trimethyllysine and saw small decreases in affinity with increasing temperature (Table 2). These studies show that affinity of the simple host **9** for trimethyllysine under the most physiologically relevant conditions (K_d = 66.0 μM at 37 °C in isotonic salt) remains within the range of values reported for binding of

TABLE 2. Thermodynamic Parameters for the Binding of **9** to $Lys(Me)_3$ at Various Temperatures

temp (°C)	K_d (μM)[a]
37	66.0 ± 2.9
30	66.7 ± 2.1
21	46.9 ± 1.2
16	39.5 ± 1.9
9	28.0 ± 2.2

[a]Determined by ITC in 40 mM Na_2HPO_4/NaH_2PO_4, pH 7.4, 137 mM NaCl, 2.7 mM KCl. [$Lys(Me_3)$] in cell = 0.2 mM; [**9**] in syringe = 5 mM. Results are averages of 2–3 replicate titrations.

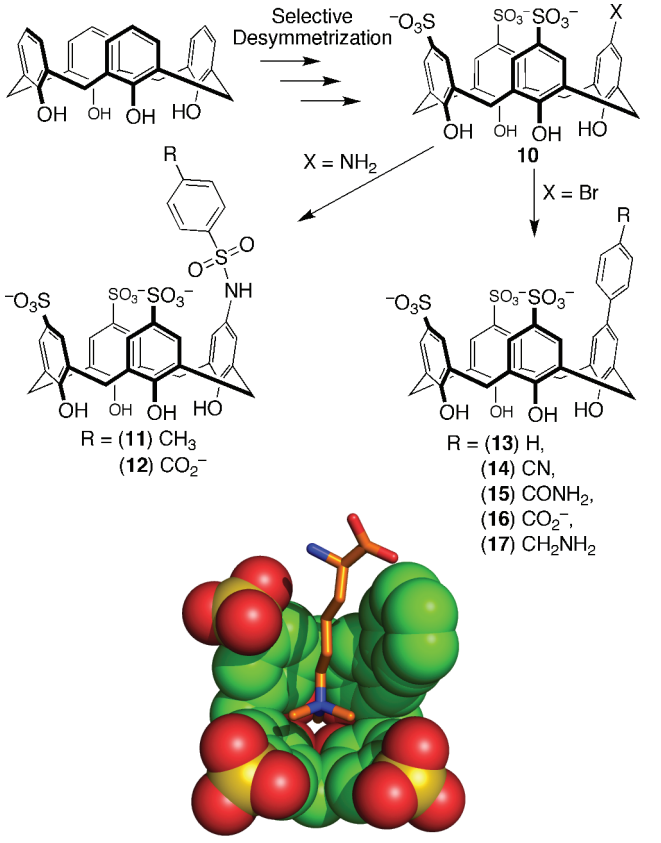


FIGURE 6. Synthesis of selectively modified hosts **10**–**17** and a model of **13**·$Lys(Me)_3$ that shows extended contacts between the introduced aryl ring and the amino acid guest.

trimethyllysine-containing peptides by their evolved recognition proteins.[58]

Thus, the parent sulfonated calixarene **9** possesses impressive affinity and selectivity for trimethyllysine despite its relative simplicity. One of the most attractive features of calixarenes is the ease with which they can be synthetically modified to tune binding properties. We recently devised routes by which a single sulfonate on **9** could be selectively replaced by a distinct, newly introduced substituent.[59] One major goal was to add functionality that could engage a greater portion of the trimethyllysine guests, as the naturally

evolved proteins manage to bind both the NMe_3^+ functionality and the neighboring functional groups in the targeted peptide partners. The outline of such a synthetic approach to two such families of functionalized hosts is shown in Figure 6. NMR titration of sulfonamide-derived hosts showed a diminished affinity for trimethyllysine ($K_d = 1.75$ and 0.34 mM for **11** and **12**, respectively) relative to parent host **9**. No NMR chemical shift occurred for distal protons on the trimethyllysine, showing that the goal of engaging the distal protons of the guest was not achieved. More rigid biaryl hosts prepared via Suzuki couplings proved more promising. NMR titration studies showed that host **13** has the strongest affinity for trimethyllysine (K_d of 15.6 μM) and best selectivity over unmethylated lysine (~150-fold) that we have observed to date. NMR titration data show that all Suzuki-coupled hosts **13**–**17** engage guest protons all the way down to the most distal CHα proton of the trimethyllysine guest, validating our structural hypothesis and showing again the benefits of rigid linkages in these synthetic hosts. Our most recent work in this area has generated calixarenes (the best scaffolds for binding trimethyllysine) appended with tetrazolates (found in our hands to be one of the most complementary binding groups for trimethyllysine), and preliminary results suggest them to be potent and selective agents that can disrupt methylation-dependent biochemical signaling events in vitro.

Outlook

The fundamental nature of the aromatic and electrostatic interactions that drive trimethyllysine recognition events is well understood. But the actual implementation of this understanding remains complicated, largely due to our inability to predict the complex relationships between structure, solvation, and binding. One lesson that arises repeatedly from the studies of supramolecular systems is that host rigidity is far more important in water, where hydrophobic collapse is possible and must be avoided, than in conventional organic solvents. Interestingly, the known structures of methyllysine-binding proteins also offers a similar insight, because their aromatic cage binding pockets are very rigid and tend to adopt the same structures whether in free or bound states.[18]

Of course, programming extreme rigidity in a host system is a recipe for success only if the *correct* shape, size, and chemical functionality are installed. It seems that in the well-known calix[4]arenes, we have identified a host system that is inherently a very good match for the side chain of trimethyllysine. Challenges remain: macrocycles that bind other important signaling motifs, including mono- and dimethyllysine, each form of methylated arginine (including isomeric forms of dimethylarginine), and acetylated lysine, all await discovery. Given the nature of the analytes, it is almost certain that these developments will rely on existing knowledge about aromatic interactions. Because these systems must operate in physiologically relevant solutions of pure (salty) water, they will also almost certainly teach us new lessons on how the complex interplay of structure, weak interactions, and solvation effects provide affinity and selectivity in water.

In addition to teaching us basic lessons, these approaches promise to provide new tools for chemical investigations of these biologically important pathways.[61] Post-translational modifications of the types discussed here constitute a specific subdivision of the bigger, growing field of protein–protein interactions. The reliance of these protein interactions on post-translational modifications, which are highly localized and structurally very well understood on/off triggers, offers continued encouragement for those who would use supramolecular approaches to target, probe, sense, disrupt, and learn from them.

We thank the numerous co-workers who contributed to the projects described herein. We gratefully acknowledge support of this work by NSERC, Prostate Cancer Canada, the Prostate Cancer Foundation of British Columbia, and the West Coast Ride to Live.

BIOGRAPHICAL INFORMATION

Kevin D. Daze attended Simon Fraser University and obtained a B.Sc. in Cell and Molecular Biology (2008). Afterward, he worked on the synthesis of novel penam-based antibiotics under the supervision of Dr. Saul Wolfe. In 2009, Kevin joined the laboratory of Dr. Fraser Hof for his graduate studies at the University of Victoria, where his thesis research is focused on creating novel therapies for prostate cancer that target epigenetic pathways. In 2012, Kevin received joint funding for his Ph.D. research from the Prostate Cancer Foundation of British Columbia (PCFBC, Vancouver, BC) and the WestCoast Ride to Live (Victoria, BC).

Fraser Hof was born in Medicine Hat, Canada, in 1976. He obtained a B.Sc. in Chemistry (1998) from the University of Alberta and a Ph.D. in Organic Chemistry (2003) from the Scripps Research Institute and was a Novartis and HFSP postdoctoral fellow in medicinal chemistry at ETH Zurich (2003–2005). His research interests include biological applications of supramolecular chemistry, protein–protein interactions, and epigenetic approaches to novel cancer therapeutics. He is an Associate Professor of Chemistry at the University of Victoria, Canada, where he is the Canada

Research Chair of Supramolecular and Medicinal Chemistry and a Career Scholar of the Michael Smith Foundation for Health Research.

FOOTNOTES

The authors declare no competing financial interest.

REFERENCES

1 Strop, P.; Mayo, S. L. Contribution of Surface Salt Bridges to Protein Stability. *Biochemistry* **2000**, *39*, 1251–1255.

2 Crowley, P. B.; Golovin, A. Cation–π Interactions in Protein–Protein Interfaces. *Proteins: Struct., Funct., Bioinf.* **2005**, *59*, 231–239.

3 Dougherty, D. A. Cation-π Interactions Involving Aromatic Amino Acids. *J. Nutr.* **2007**, *137*, 1504S–1508S.

4 Zeng, L.; Zhou, M.-M. Bromodomain: An Acetyl-Lysine Binding Domain. *FEBS Lett.* **2002**, *513*, 124–128.

5 Weake, V. M.; Workman, J. L. Histone Ubiquitination: Triggering Gene Activity. *Mol. Cell* **2008**, *29*, 653–663.

6 Bedford, M. T.; Richard, S. Arginine Methylation: An Emerging Regulator of Protein Function. *Mol. Cell* **2005**, *18*, 263–272.

7 Kubicek, S.; Jenuwein, T. A Crack in Histone Lysine Methylation. *Cell* **2004**, *119*, 903–906.

8 Dambacher, S.; Hahn, M.; Schotta, G. Epigenetic Regulation of Development by Histone Lysine Methylation. *Heredity* **2010**, *105*, 24–37.

9 Alfred, J. Histone Methylation Takes the Lead. *Nat. Rev. Genet.* **2001**, *2*, 916–916.

10 Maurer-Stroh, S.; Dickens, N. J.; Hughes-Davies, L.; Kouzarides, T.; Eisenhaber, F.; Ponting, C. P. The Tudor Domain 'Royal Family': Tudor, Plant Agenet, Chromo, PWWP and MBT Domains. *Trends Biochem. Sci.* **2003**, *28*, 69–74.

11 Kim, J.; Daniel, J.; Espejo, A.; Lake, A.; Krishna, M.; Xia, L.; Zhang, Y.; Bedford, M. T. Tudor, MBT and chromo domains gauge the degree of lysine methylation. *EMBO Rep.* **2006**, *7*, 397–403.

12 Chen, C.; Nott, T. J.; Jin, J.; Pawson, T. Deciphering Arginine Methylation: Tudor Tells the Tale. *Nat. Rev. Mol. Cell Biol.* **2011**, *12*, 629–642.

13 Bedford, M. T.; Clarke, S. G. Protein Arginine Methylation in Mammals: Who, What, and Why. *Mol. Cell* **2009**, *33*, 1–13.

14 Sepideh, K. Recognition of Methylated Histones: New Twists and Variations. *Curr. Opin. Struct. Biol.* **2011**, *21*, 744–749.

15 Blus, B. J.; Wiggins, K.; Khorasanizadeh, S. Epigenetic Virtues of Chromodomains. *Crit. Rev. Biochem. Mol. Biol.* **2011**, *46*, 507–526.

16 Wu, H.; Zeng, H.; Lam, R.; Tempel, W.; Amaya, M. F.; Xu, C.; Dombrovski, L.; Qiu, W.; Wang, Y.; Min, J. Structural and Histone Binding Ability Characterizations of Human PWWP Domains. *PLoS ONE* **2011**, *6*, No. e18919.

17 Xu, C.; Bian, C.; Yang, W.; Galka, M.; Ouyang, H.; Chen, C.; Qiu, W.; Liu, H.; Jones, A. E.; MacKenzie, F.; Pan, P.; Li, S. S.-C.; Wang, H.; Min, J. Binding of Different Histone Marks Differentially Regulates the Activity and Specificity of Polycomb Repressive Complex 2 (PRC2). *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 19266–19271.

18 Taverna, S. D.; Li, H.; Ruthenburg, A. J.; Allis, C. D.; Patel, D. J. How Chromatin-Binding Modules Interpret Histone Modifications: Lessons from Professional Pocket Pickers. *Nat. Struct. Mol. Biol.* **2007**, *14*, 1025–1040.

19 Jane, M. It Takes a PHD to Read the Histone Code. *Cell* **2006**, *126*, 22–24.

20 Lu, Z.; Lai, J.; Zhang, Y. Importance of Charge Independent Effects in Readout of the Trimethyllysine Mark by HP1 Chromodomain. *J. Am. Chem. Soc.* **2009**, *131*, 14928–14931.

21 Hughes, R. M.; Wiggins, K. R.; Khorasanizadeh, S.; Waters, M. L. Recognition of Trimethyllysine by a Chromodomain Is Not Driven by the Hydrophobic Effect. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 11184–11188.

22 Palacios, A.; Munoz, I. G.; Pantoja-Uceda, D.; Marcaida, M. J.; Torres, D.; Martin-Garcia, J. M.; Luque, I.; Montoya, G.; Blanco, F. J. Molecular Basis of Histone H3K4me3 Recognition by ING4. *J. Biol. Chem.* **2008**, *283*, 15956–15964.

23 Ruthenburg, A. J.; Allis, C. D.; Wysocka, J. Methylation of Lysine 4 on Histone H3: Intricacy of Writing and Reading a Single Epigenetic Mark. *Mol. Cell* **2007**, *25*, 15–30.

24 Li, H.; Ilin, S.; Wang, W.; Duncan, E.; Wysocka, J.; Allis, C.; Patel, D. Molecular Basis for Site-Specific Read-out of Histone H3K4me3 by the BPTF PHD Finger of NURF. *Nature* **2006**, *442*, 91–95.

25 Herold, J. M.; Wigle, T. J.; Norris, J. L.; Lam, R.; Korboukh, V. K.; Gao, C.; Ingerman, L. A.; Kireev, D. B.; Senisterra, G.; Vedadi, M.; Tripathy, A.; Brown, P. J.; Arrowsmith, C. H.; Jin, J.; Janzen, W. P.; Frye, S. V. Small-Molecule Ligands of Methyl-Lysine Binding Proteins. *J. Med. Chem.* **2011**, *54*, 2504–2511.

26 Gao, C.; Herold, J. M.; Kireev, D.; Wigle, T.; Norris, J. L.; Frye, S. Biophysical Probes Reveal a Compromise Nature of the Methyl-lysine Binding Pocket in L3MBTL1. *J. Am. Chem. Soc.* **2011**, *133*, 5357–5362.

27 Herr, R. J. 5-Substituted-1H-Tetrazoles As Carboxylic Acid Isosteres: Medicinal Chemistry and Synthetic Methods. *Bioorg. Med. Chem.* **2002**, *10*, 3379–3393.

28 Noda, K.; Saad, Y.; Kinoshita, A.; Boyle, T. P.; Graham, R. M.; Husain, A.; Karnik, S. S. Tetrazole and Carboxylate Groups of Angiotensin Receptor Antagonists Bind to the Same Subsite by Different Mechanisms. *J. Biol. Chem.* **1995**, *270*, 2284–2289.

29 McKie, A. H.; Friedland, S.; Hof, F. Tetrazoles Are Potent Anion Recognition Elements That Emulate the Disfavored Anti Conformations of Carboxylic Acids. *Org. Lett.* **2008**, *10*, 4653–4655.

30 Hoppmann, V.; Thorstensen, T.; Kristiansen, P. E.; Veiseth, S. V.; Rahman, M. A.; Finne, K.; Aalen, R. B.; Aasland, R. The CW Domain, a New Histone Recognition Module in Chromatin Proteins. *EMBO J.* **2011**, *30*, 1939–1952.

31 Li, H.; Ilin, S.; Wang, W.; Duncan, E. M.; Wysocka, J.; Allis, C. D.; Patel, D. J. Molecular Basis for Site-Specific Read-out of Histone H3K4me3 by the BPTF PHD Finger of NURF. *Nature* **2006**, *442*, 91–95.

32 Hennrich, G.; Anslyn, E. V. 1,3,5-2,4,6-Functionalized, Facially Segregated Benzenes—Exploitation of Sterically Predisposed Systems in Supramolecular Chemistry. *Chem.—Eur. J.* **2002**, *8*, 2218–2224.

33 Wang, X.; Hof, F. (How) Does 1,3,5-Triethylbenzene Scaffolding Work? Analyzing the Abilities of 1,3,5-Triethylbenzene- And 1,3,5-Trimethylbenzene-Based Scaffolds To Pre-organize the Binding Elements of Supramolecular Hosts and To Improve Binding of Targets. *Beilstein J. Org. Chem.* **2012**, *8*, 1–10.

34 Wiskur, S. L.; Ait-Haddou, H.; Lavigne, J. J.; Anslyn, E. V. Teaching Old Indicators New Tricks. *Acc. Chem. Res.* **2001**, *34*, 963–972.

35 Mahnke, D. J.; McDonald, R.; Hof, F. A Shape-Dependent Hydrophobic Effect for Tetrazoles. *Chem. Commun.* **2007**, 3738–3740.

36 Tominey, A.; Andrew, D.; Oliphant, L.; Rosair, G. M.; Dupre, J.; Kraft, A. Supramolecular Binding of Protonated Amines to a Receptor Microgel in Aqueous Medium. *Chem. Commun.* **2006**, 2492–2494.

37 Whiting, A. L.; Neufeld, N. M.; Hof, F. A Tryptophan-Analog Host Whose Interactions with Ammonium Ions in Water Are Dominated by the Hydrophobic Effect. *Tetrahedron Lett.* **2009**, *50*, 7035–7037.

38 Ito, K.; Nagase, K.; Morohashi, N.; Ohba, Y. Interaction between Quaternary Ammonium Ions and Dipeptides: Positive Anion Allosteric Effect. *Chem. Pharm. Bull.* **2005**, *53*, 90–94.

39 Beshara, C. S.; Hof, F. Modular Incorporation of 1-Benzyltryptophan into Dipeptide Hosts That Bind Acetylcholine in Pure Water. *Can. J. Chem.* **2010**, *88*, 1009–1016.

40 Kearney, P. C.; Mizoue, L. S.; Kumpf, R. A.; Forman, J. E.; McCurdy, A.; Dougherty, D. A. Molecular Recognition in Aqueous Media. New Binding Studies Provide Further Insights into the Cation-.Pi. Interaction and Related Phenomena. *J. Am. Chem. Soc.* **1993**, *115*, 9907–9919.

41 Ma, J. C.; Dougherty, D. A. The Cation–π Interaction. *Chem. Rev.* **1997**, *97*, 1303–1324.

42 Otto, S.; Furlan, R. L. E.; Sanders, J. K. M. Selection and Amplification of Hosts From Dynamic Combinatorial Libraries of Macrocyclic Disulfides. *Science* **2002**, *297*, 590–593.

43 Corbett, P. T.; Sanders, J. K. M.; Otto, S. Exploring the Relation between Amplification and Binding in Dynamic Combinatorial Libraries of Macrocyclic Synthetic Receptors in Water. *Chem.—Eur. J.* **2008**, *14*, 2153–2166.

44 Ingerman, L. A.; Cuellar, M. E.; Waters, M. L. A Small Molecule Receptor That Selectively Recognizes Trimethyl Lysine in a Histone Peptide with Native Protein-Like Affinity. *Chem. Commun.* **2011**, *46*, 1839–1841.

45 Mokhtari, B.; Pourabdollah, K.; Dallali, N. A Review of Calixarene Applications in Nuclear Industries. *J. Radioanal. Nucl. Chem.* **2010**, *287*, 921–934.

46 Pinter, T.; Jana, S.; Courtemanche, R. J. M.; Hof, F. Recognition Properties of Carboxylic Acid Bioisosteres: Anion Binding by Tetrazoles, Aryl Sulfonamides, and Acyl Sulfonamides on a Calix[4]arene Scaffold. *J. Org. Chem.* **2011**, *76*, 3733–3741.

47 Perret, F.; Lazar, A. N.; Coleman, A. W. Biochemistry of the para-Sulfonato-calix[n]arenes. *Chem. Commun.* **2006**, 2425–2438.

48 Perret, F.; Coleman, A. W. Biochemistry of Anionic Calix[n]arenes. *Chem. Commun.* **2011**, *47*, 7303–7319.

49 Selkti, M.; Coleman, A. W.; Nicolis, I.; Douteau-Guevel, N.; Villain, F.; Tomas, A.; de Rango, C. The First Example of a Substrate Spanning the Calix[4]arene Bilayer: The Solid State Complex of p-Sulfonatocalix[4]arene with L-Lysine. *Chem. Commun.* **2000**, 161–162.

50 Beshara, C. S.; Jones, C. E.; Daze, K. D.; Lilgert, B. J.; Hof, F. A Simple Calixarene Recognizes Post-translationally Methylated Lysine. *ChemBioChem* **2010**, *11*, 63–66.

51 Douteau-Guevel, N.; W. Coleman, A.; Morel, J.-P.; Morel-Desrosiers, N. Complexation of the Basic Amino Acids Lysine and Arginine by Three Sulfonatocalix[n]arenes (n = 4, 6 and 8) in Water: Microcalorimetric Determination of the Gibbs Energies, Enthalpies and Entropies of Complexation. *J. Chem. Soc., Perkin Trans. 2* **1999**, 629–634.

52 Lehn, J.-M.; Meric, R.; Vigneron, J.-P.; Cesario, M.; Guilhem, J.; Pascard, C.; Asfari, Z.; Vicens, J. Binding of Acetylcholine and Other Quaternary Ammonium Cations by Sulfonated

Calixarenes. Crystal Structure of a [Choline-Tetrasulfonated Calix[4]arene] Complex. *Supramol. Chem.* **1995**, *5*, 97–103.

53 Guo, D.-S.; Uzunova, V. D.; Su, X.; Liu, Y.; Nau, W. M. Operational Calixarene-Based Fluorescent Sensing Systems for Choline and Acetylcholine and Their Application to Enzymatic Reactions. *Chem. Sci.* **2011**, *2*, 1722–1734.

54 Koh, K. N.; Araki, K.; Ikeda, A.; Otsuka, H.; Shinkai, S. Reinvestigation of Calixarene-Based Artificial-Signaling Acetylcholine Receptors Useful in Neutral Aqueous (Water/Methanol) Solution. *J. Am. Chem. Soc.* **1996**, *118*, 755–758.

55 Buschmann, H. J.; Mutihac, L.; Schollmeyer, E. Complexation of Some Amino Acids and Peptides by p-Sulfonatocalix[4]arene and Hexasodium p-Sulfonatocalix[6]-arene in Aqueous Solution. *J. Inclusion Phenom. Macrocycl. Chem.* **2003**, *46*, 133–137.

56 Fischle, W.; Wang, Y.; Jacobs, S. A.; Kim, Y.; Allis, C. D.; Khorasanizadeh, S. Molecular Basis for the Discrimination of Repressive Methyl-lysine Marks in Histone H3 by Polycomb and HP1 Chromodomains. *Genes Dev.* **2003**, *17*, 1870–1881.

57 Douteau-Guevel, N.; Perret, F.; Coleman, A. W.; Morel, J.-P.; Morel-Desrosiers, N. Binding of Dipeptides and Tripeptides Containing Lysine or Arginine by p-Sulfonatocalixarenes in Water: NMR and Microcalorimetric Studies. *J. Chem. Soc., Perkin Trans. 2* **2002**, 524–532.

58 Kaustov, L.; Ouyang, H.; Amaya, M.; Lemak, A.; Nady, N.; Duan, S.; Wasney, G. A.; Li, Z.; Vedadi, M.; Schapira, M.; Min, J.; Arrowsmith, C. H. Recognition and Specificity Determinants of the Human Cbx Chromodomains. *J. Biol. Chem.* **2011**, *286*, 521–529.

59 Daze, K. D.; Ma, M. C. F.; Pineux, F.; Hof, F. Synthesis of New Trisulfonated Calix[4]arenes Functionalized at the Upper Rim, and Their Complexation with the Trimethyllysine Epigenetic Mark. *Org. Lett.* **2012**, *14*, 1512–1515.

60 Arena, G.; Casnati, A.; Contino, A.; Magri, A.; Sansone, F.; Sciotto, D.; Ungaro, R. Inclusion of Naturally Occurring Amino Acids in Water Soluble Calix[4]arenes: a Microcalorimetric and 1H NMR Investigation Supported by Molecular Modeling. *Org. & Biomol. Chem.* **2006**, *4*, 243–249.

61 Minaker, S. A.; Daze, K. D.; Ma, M. C. F.; Hof, F. Antibody-free Reading of the Histone Code Using a Simple Chemical Sensor Array. *J. Am. Chem. Soc.* **2012**, DOI: http://dx.doi.org/10.1021/ja303465x.