DOI: 10.1002/cbic.200900633

A Simple Calixarene Recognizes Post-translationally Methylated Lysine

Cory S. Beshara, Catherine E. Jones, Kevin D. Daze, Brandin J. Lilgert, and Fraser Hof*[a]

The methylation of amino acid side chains is a post-translational modification (PTM) important for several gene regulation and developmental signaling pathways.[1–4] Lysine residues that are specifically mono-, di-, and trimethylated by methyltransferase enzymes carry out their signaling function by acting as recruitment points for new protein–protein interactions.[2] Histone methylation, in particular, is controlled by histone lysine methyl transferases (HKMTs) and lysine-specific demethylases (LSDs).[5–7] Growing interest in methylation pathways has recently driven efforts at therapeutic intervention through the discovery and synthesis of HKMT and LSD inhibitors. The first members of these enzyme inhibitor families are being pursued as novel cancer therapies[8–10] and as chemical agents for the generation of pluripotent cells from differentiated cell lines.[11–13] The various effector domains that recognize and bind to di- and trimethylated lysine residues share a common recognition motif referred to as an "aromatic cage".[2] These motifs consist of a highly preorganized cluster of aromatic side chains that bind the cationic Lys($Me_{2/3}$) side chains through cation–π interactions.[14] Aromatic cage residues are often accompanied by carboxylate side chains that provide an additional electrostatic driving force for binding (Figure 1 A).[2,15]

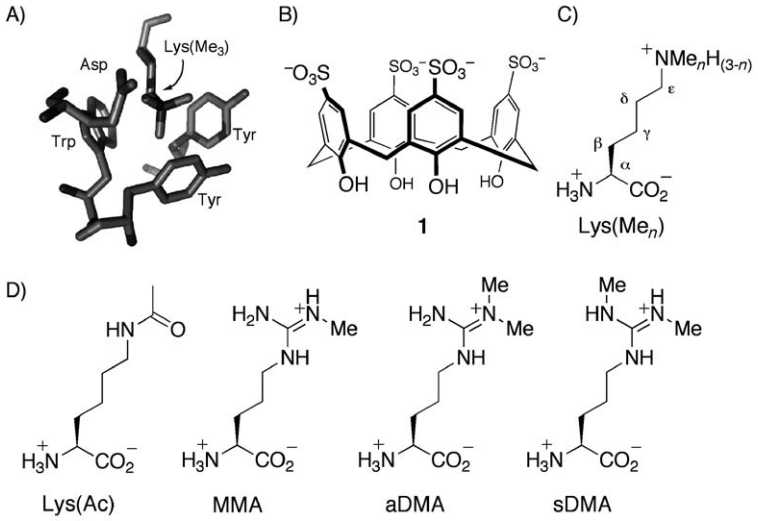

Figure 1. A) The aromatic cage binding motif of heterochromatin protein 1 (HP1) chromodomain bound to a trimethylated lysine side chain from histone 3.[35] B) Host *p*-sulfonatocalix[4]arene (**1**). C) Methylated lysine derivatives used in this study. D) Other amino acids used in this study; MMA: monomethylarginine; aDMA: asymmetric dimethylarginine; sDMA: symmetric dimethylarginine.

There exist no synthetic molecules that can distinguish between post-translational methylation states. We hypothesized that Lys(Me_3) sites might be selectively recognized on the basis of their methylation state by an appropriate concave host molecule. The well-known host *p*-sulfonatocalix[4]arene (**1**, Figure 1 B) imitates the rigid multiaromatic cavity and charge complementarity of natural aromatic cages, and has been reported to bind ammonium ions in pure water.[16–22] We first examined its affinity for the free amino acid lysine with all possible degrees of side-chain methylation (Lys, Lys(Me), Lys(Me_2), and Lys(Me_3)) using NMR titrations in D_2O (40 mM Na_2HPO_4/NaH_2PO_4, pD 7.0 = pH 7.4[23]). Fitting the chemical shift data to 1:1 binding isotherms provided K_{assoc} values for each host–guest pair, and the 1:1 stoichiometry was confirmed by Job plot (Figure S1 in the Supporting Information). The affinity of **1** for lysine derivatives increases with increasing methylation, with an overall 70-fold selectivity for Lys(Me_3) versus Lys (Table 1).

Isothermal titration calorimetry (ITC) was also carried out in order to confirm K_{assoc} values and to determine thermodynamic parameters for the binding events. The K_{assoc} values and stoichiometries ($N = 0.67–1.33$ across all guests studied) determined by ITC agree well with the NMR spectroscopy results (Table 1). The complexation of each lysine derivative by **1**, as measured by ITC at 303 K, has a large favorable enthalpic component and a smaller favorable entropic component. Increasing methylation of lysine is accompanied by significant increases in enthalpic driving force and smaller favorable changes in entropy.

The binding of host **1** to other free amino acids in water has been previously studied.[19,22,24,25] The maximum reported affinity of **1** for any free amino acid under similar buffered conditions is for Arg ($K_{assoc} = 1520\ M^{-1}$, pH 8, 10 mM phosphate buffer[24]). Whether comparing to this literature value or to the value we observe under our slightly more competitive experimental conditions ($K_{assoc} = 330\ M^{-1}$, pH 7.4, 40 mM phosphate buffer), the affinity of **1** for Lys(Me_3) is far higher than for any other amino acid. We further explored this selectivity by comparison to other post-translationally modified lysines and arginines. We found, unsurprisingly, that the biologically important product of post-translational lysine acetylation (Lys(Ac))[2] displays only very weak binding to **1** (Table 1). The weakness of this interaction is almost certainly due to the fact that this modification renders the side chain neutral—prior studies of **1** have generally demonstrated weak binding of neutral amino acids in phosphate buffer.[22]

[a] *Dr. C. S. Beshara, C. E. Jones, K. D. Daze, B. J. Lilgert, Prof. F. Hof*
Department of Chemistry, University of Victoria
P.O. Box 3065, STN CSC, Victoria, BC V8W 3V6 (Canada)
Fax: (+1) 250-721-7147
E-mail: fhof@uvic.ca

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.200900633.

Table 1. Thermodynamic parameters for the binding of **1** to amino acids and peptides.

	Guest	K_{assoc} [M^{-1}] in D_2O (NMR)[a]	K_{assoc} [M^{-1}] in H_2O (ITC)[b]	ΔH [kJ mol^{-1}][b]	$T\Delta S$ [kJ mol^{-1}][b]
1	Lys	520 ± 300	n.d.[c]	n.d.[c]	n.d.[c]
2	Lys(Me)	4000 ± 3000	3000 ± 1700	−16.4 ± 8.3	3.5 ± 9.8
3	Lys(Me_2)	16200 ± 4300	10500 ± 2000	−19.9 ± 3.0	3.4 ± 3.5
4	Lys(Me_3)	37000 ± 18000	35700 ± 2500	−22.1 ± 0.6	4.3 ± 0.8
5	Lys(Ac)	12 ± 34	n.d.[c]	n.d.[c]	n.d.[c]
6	Arg	330 ± 260	n.d.[c]	n.d.[c]	n.d.[c]
7	MMA	760 ± 330	n.d.[c]	n.d.[c]	n.d.[c]c
8	aDMA	1100 ± 460	1300 ± 180	−55 ± 13	−37 ± 10
9	sDMA[d]	n.d.[d]	1100 ± 100[d]	−13.4 ± 0.4[d]	4.4 ± 8.2[d]
10	Ac-RKST-NH_2	15200 ± 4000	5500 ± 1000	−16.4 ± 1.5	5.3 ± 0.5
11	Ac-R(KMe_3)ST-NH_2	$> 10^5$[e]	96600 ± 10000	−23.2 ± 0.7	5.6 ± 0.5

[a] Determined by ^{1}H NMR spectroscopy (500 MHz) at 298 K in D_2O (40 mM Na_2HPO_4/NaH_2PO_4, pD 7.0 = pH 7.4[23]) by titration of a solution of **1** (20–300 mM) into a solution of amino acid or peptide (1–2 mM). The K_{assoc} values reported are the averages of values arising from all 2–5 trackable NMR signals from each of 2–3 replicate titrations per guest. Errors reported are standard deviations. See Figure S2 in the Supporting Information for exemplary curve fits. [b] Determined by ITC at 303 K in H_2O (40 mM Na_2HPO_4/NaH_2PO_4, pH 7.4) by titration of a solution of **1** (1–10 mM) into a solution of amino acid or peptide (50–600 μM). Values reported are the averages 2–3 replicate titrations. See Figure S3 in the Supporting Information for exemplary curve fits. Errors are calculated by using fitting errors propagated by standard means; [c] n.d. = K_{assoc} value below the limit measurable by using ITC; [d] sDMA gave a complex NMR curve that could not be fit to any simple 1:1, 2:1, or 1:2 binding isotherm (Figure S5 in the Supporting Information). Satisfactory curve fitting for sDMA ITC data was obtained by using an *N* value of 1.00, but in light of the problems with the NMR spectroscopy data the resulting thermodynamic data are suspect and are reported only for the information of the reader; [e] K_{assoc} greater than the limit measurable by NMR spectroscopy under these conditions.

dimethylated lysine counterpart Lys(Me_2), and > 30-fold weaker than observed for Lys(Me_3).

Why is Lys(Me_3) so much better a guest for **1** than the other cationic amino acid side chains? We used both NMR chemical shift trends and molecular modeling to understand the structural details of each complex. Like others before us, we observe for unmodified Lys significant up-field shifts for β, γ, δ, and ε methylene groups upon complexation with **1** (Figure 2 A).[24,29] These shifts arise due to a "side-on" binding mode that has previously been observed in an X-ray co-crystal structure of Lys and **1**.[30] We observe for Lys(Me_3) significant up-field shifts of Me and CH_2ε protons and, unlike unmethylated lysine, no up-field shifts for CH_2β and CH_2γ protons (Figure 2 B). This is suggestive of a different binding mode in which the $-NMe_3^+$ functionality at the end of the Lys(Me_3) side chain is bound deep within the cavity of **1**. Computational energy minimizations[31] of Lys and Lys(Me_3) complexed to **1** provide models that are consistent with our NMR spectroscopy data (Figure S4 in the Supporting Information).

More interesting is the comparison to arginine, because it is also cationic and is also subject to post-translational methylations that render it more hydrophobic than its parent unmethylated form.[26,27] Arginine can exist in nature as three distinct methylated forms: monomethylarginine (MMA), symmetric dimethylarginine (sDMA), and asymmetric dimethylarginine (aDMA; Figure 1).[2,28] For Arg, as for Lys, our studies of binding to **1** show that affinity increases with increasing methylation (Table 1). But the affinities of **1** for both dimethylated arginine isomers remain tenfold weaker than those observed for their

Each of the methylated Arg derivatives displays significant up-field shifts for its *N*-methyl group(s) upon binding to **1**. Arg, MMA and sDMA also display up-field shifts for the side-chain β, γ methylene groups. Molecular models for Arg, MMA, and

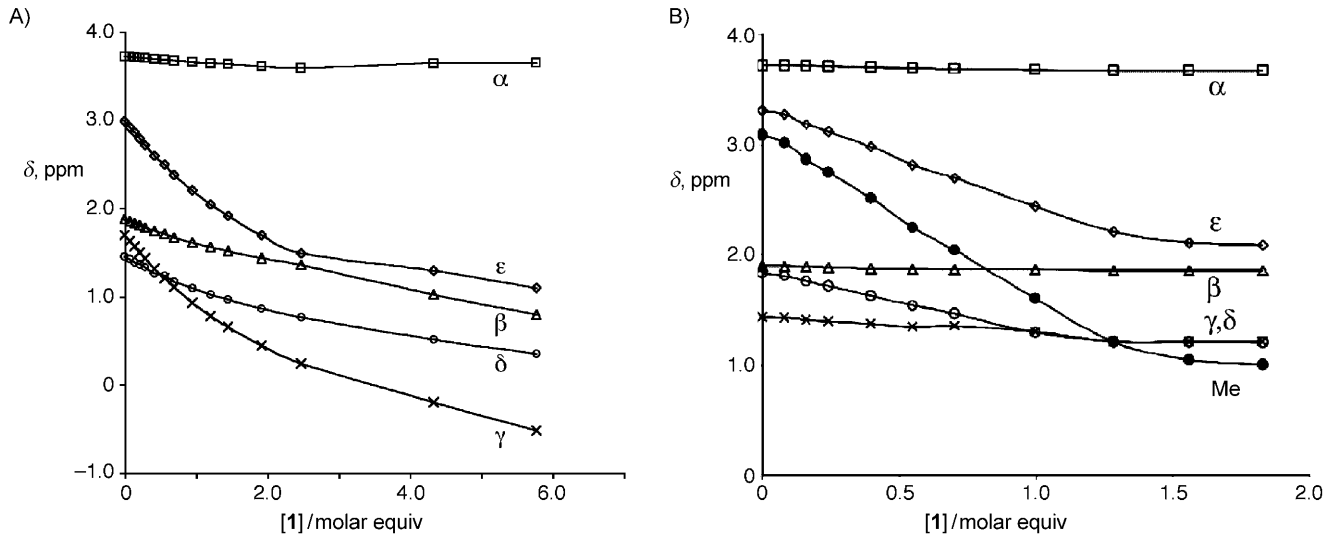

Figure 2. ^{1}H NMR titration data (500 MHz) from the addition of **1** to: A) Lys, and B) Lys(Me_3). Conditions are as described in Table 1. See Figure 1 for lettering of resonances. Lines are added to guide the reader's eye.

sDMA are wholly consistent with the NMR spectroscopy data, and show that the side-chain methylenes are engaged within the cavity, and that their methyls are not bound as deeply within the cavity of the calixarene as are the methyls of Lys(Me_3); see Figure S4 in the Supporting Information. Only in the case of aDMA is the modeling, which suggests a side-on structure with methylenes engaged within the cavity (Figure S4 in the Supporting Information), at odds with the structure suggested by the NMR spectroscopy data, in which the β and γ methylenes experience no up-field shift at all (Figure S5 in the Supporting Information). It is also interesting that while the K_{assoc} value for aDMA is in line with the trends observed for other cationic amino acids and peptides, it is the only cationic amino acid among the examples studied here to display a large, negative ΔH and compensating large, negative $T\Delta S$ of binding to **1**. The disagreement of our gas-phase calculations and solution-phase NMR spectroscopy data, taken together with the large deviations from the expected enthalpy and entropy of binding, suggest an anomalous solvation effect that operates only in the case of aDMA. We and others have looked, and failed to find any significant solvation differences between aDMA and sDMA in prior studies,[26,27] and the origins of the differences that we observe for the binding of these two isomers to **1** remain unclear.

To see if these lessons in affinity, selectivity, and structure extended to the binding of **1** to Lys(Me_3) in the context of a peptide, we used the histone tail peptides Arg-Lys-Ser-Thr and Arg-Lys(Me_3)-Ser-Thr (each synthesized as N-terminal acetamides and C-terminal primary amides), taken from the sequence of histone 3A around Lys9. This region near the N terminus of histone 3 is a known methylation site of importance in gene silencing.[2,32] The K_{assoc} value measured by ITC for **1**·Arg-Lys-Ser-Thr is 5500 M^{-1}, somewhat less than the value determined by averaging K_{assoc} values arising from the fit of different NMR resonances. This type of discrepancy can arise when multiple binding modes are present; the ITC measures the heat arising from all binding modes, while the limitation that only clearly resolved NMR resonances can be used in curve fitting can lead to an unavoidable bias in the NMR-derived K_{assoc} value. Indeed, the addition of 1 equiv of **1** to Arg-Lys-Ser-Thr produces significant up-field shifts of NMR resonances for both the lysine- and arginine-derived ^{1}H resonances; this is suggestive of competing inclusion of both cationic side chains within the cavity of **1** (Figure 3 A). In contrast, the addition of the first equivalent of **1** to the methylated peptide

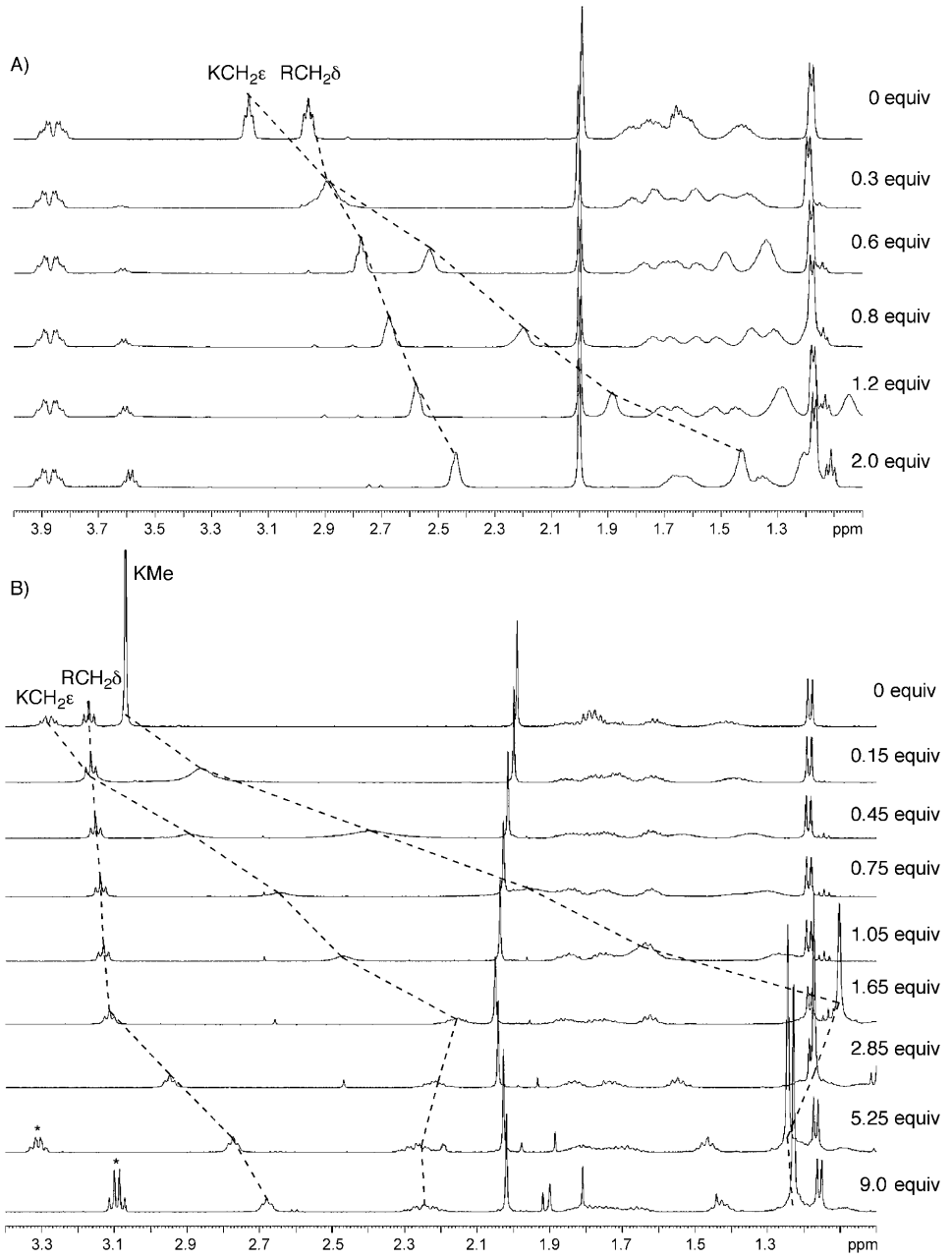

Figure 3. A) ^{1}H NMR titration data for addition of **1** to unmethylated Arg-Lys-Ser-Thr. Up-field shifts of both Arg- and Lys-derived resonances suggest competing inclusion of both side chains. B) Addition of **1** to methylated peptide Arg-Lys(Me_3)-Ser-Thr shows exclusive up-field shifts for Lys(Me_3) side chain resonances up to saturation at ~1.6 equiv, followed by the up-field shift of the Arg side chain resonances after ~1.6 equiv; trace EtOH is indicated (*).

causes striking up-field shifts for the lysine's Me and $CH_2\varepsilon$ resonances and no detectable shift for arginine resonances or any other peptide resonances (Figure 3B). The arginine signals begin to shift only after the response of the Lys(Me_3) side chain is saturated. We conclude from the chemical shift data that the first equivalent of **1** binds to this peptide by engaging the Lys(Me_3) side chain with high site selectivity. The K_{assoc} value measured by ITC for **1**·Arg-Lys(Me_3)-Ser-Thr is 96600 M^{-1}, a ~threefold increase over the affinity for free Lys(Me_3); this suggests that the other nearby amino acids, while not included within the cavity, might assist binding by making contacts with the upper rim sulfonates.

As a whole, the NMR chemical shift data demonstrate that methylated lysines bind to **1** by an "end-on" binding mode that operates both on the free amino acid and on a simple histone-tail peptide. The high affinity of **1** for Lys(Me_3) suggests that this side chain has the combination of size, shape, and charge that complements the anionic, aromatic binding pocket of **1** better than all other cationic amino acid side chains. While it is clear that tetraanionic host **1** can also bind strongly to other highly cationic protein surface sites,[19–21,25,33] we find that **1** has an affinity for Lys(Me_3) that is at least 30-fold higher than any other unmodified or methylated individual amino acid. Despite its simplicity, the affinity of **1** for the histone-tail peptide Arg-Lys(Me_3)-Ser-Thr is on the same order of magnitude as the affinities of the highly evolved effector proteins that bind to such Lys(Me_3) sites on histone tails.[34] We are working to improve the affinity and specificity of calixarenes for these biochemically important targets.

Acknowledgements

We thank Christine Greenwood for assistance with the NMR spectroscopy. F.H. is a Career Scholar of the Michael Smith Foundation for Health Research (MSFHR). This work was supported by NSERC and MSFHR.

Keywords: amino acids · calixarenes · methylated lysine · molecular recognition · post-translational modifications

[1] J. Huang, R. Sengupta, A. B. Espejo, M. G. Lee, J. A. Dorsey, M. Richter, S. Opravil, R. Shiekhattar, M. T. Bedford, T. Jenuwein, S. L. Berger, *Nature* **2007**, *449*, 105.
[2] S. D. Taverna, H. Li, A. J. Ruthenburg, C. D. Allis, D. J. Patel, *Nat. Struct. Mol. Biol.* **2007**, *14*, 1025.
[3] B. E. Bernstein, T. S. Mikkelsen, X. Xie, M. Kamal, D. J. Huebert, J. Cuff, B. Fry, A. Meissner, M. Wernig, K. Plath, R. Jaenisch, A. Wagschal, R. Feil, S. L. Schreiber, E. S. Lander, *Cell* **2006**, *125*, 315.
[4] P. Völkel, P.-O. Angrand, *Biochimie* **2007**, *89*, 1.
[5] A. J. Bannister, T. Kouzarides, *Nature* **2005**, *436*, 1103.
[6] J. Schneider, A. Shilatifard, *ACS Chem. Biol.* **2006**, *1*, 75.
[7] T. Swigut, J. Wysocka, *Cell* **2007**, *131*, 29.
[8] S. B. Hake, A. Xiao, C. D. Allis, *Br. J. Cancer* **2004**, *90*, 761.
[9] S. V. Antonello Mai, D. Cheng, A. Perrone, R. Ragno, S. Simeoni, G. Sbardella, G. Brosch, A. Nebbioso, M. Conte, L. Altucci, M. T. Bedford, *ChemMedChem* **2007**, *2*, 987.
[10] D. Greiner, T. Bonaldi, R. Eskeland, E. Roemer, A. Imhof, *Nat. Chem. Biol.* **2005**, *1*, 143.
[11] Y. Shi, J. Tae Do, C. Desponts, H. S. Hahm, H. R. Schöler, S. Ding, *Cell Stem Cell* **2008**, *2*, 525.
[12] Y. Chang, X. Zhang, J. R. Horton, A. K. Upadhyay, A. Spannhoff, J. Liu, J. P. Snyder, M. T. Bedford, X. Cheng, *Nat. Struct. Mol. Biol.* **2009**, *16*, 312.
[13] S. Kubicek, R. J. O'Sullivan, E. M. August, E. R. Hickey, Q. Zhang, M. L. Teodoro, S. Rea, K. Mechtler, J. A. Kowalski, C. A. Homon, T. A. Kelly, T. Jenuwein, *Mol. Cell* **2007**, *25*, 473.
[14] J. C. Ma, D. A. Dougherty, *Chem. Rev.* **1997**, *97*, 1303.
[15] >Our own survey of protein data bank structures with proteins bound to partners containing Lys(Me_3) shows that 25/31 such structures use an aromatic cage motif in which the Lys(Me_3) side chain is in contact with two or more aromatic residues. Of those 25, a total of 12 also included a neighboring Asp or Glu side chain to complement the cationic Lys(Me_3) side chain.
[16] J. L. Atwood, L. J. Barbour, P. C. Junk, G. W. Orr, *Supramol. Chem.* **1995**, *5*, 105.
[17] H.-J. Buschmann, L. Mutihac, E. Schollmeyer, *J. Inclusion Phenom. Macrocyclic Chem.* **2003**, *46*, 133.
[18] N. Douteau-Guével, A. W. Coleman, J.-P. Morel, N. Morel-Desrosiers, *J. Chem. Soc. Perkin Trans. 2* **1999**, 629.
[19] F. Perret, A. N. Lazar, A. W. Coleman, *Chem. Commun.* **2006**, 2425.
[20] L. Memmi, A. N. Lazar, A. Brioude, V. Ball, A. W. Coleman, *Chem. Commun.* **2001**, 2474.
[21] N. Douteau-Guével, F. Perret, A. W. Coleman, J.-P. Moreland, N. Morel-Desrosiers, *J. Chem. Soc. Perkin Trans. 2* **2002**, 524.
[22] G. Arena, A. Casnatic, A. Continoa, A. Magrì, F. Sansone, D. Sciottoa, R. Ungaro, *Org. Biomol. Chem.* **2006**, *4*, 243.
[23] P. K. Glasoe, F. A. Long, *J. Phys. Chem.* **1960**, *64*, 188.
[24] N. Douteau-Guével, A. W. Coleman, J.-P. Morel, N. Morel-Desrosier, *J. Chem. Soc. Perkin Trans. 2* **1999**, 629.
[25] H.-J. Buschmann, L. Mutihac, E. Schollmeyer, *J. Inclusion Phenom. Macrocyclic Chem.* **2003**, *46*, 133.
[26] R. M. Hughes, M. L. Waters, *J. Am. Chem. Soc.* **2006**, *128*, 12735.
[27] K. McQuinn, J. S. McIndoe, F. Hof, *Chem. Eur. J.* **2008**, *14*, 6483.
[28] A. E. McBride, P. A. Silver, *Cell* **2001**, *106*, 5.
[29] N. Douteau-Guével, A. W. Coleman, J.-P. Morel, N. Morel-Desrosier, *J. Phys. Org. Chem.* **1998**, *11*, 693.
[30] M. Selkti, A. W. Coleman, I. Nicolis, N. Douteau-Guével, F. Villain, A. Tomas, C. de Rango, *Chem. Commun.* **2000**, 161.
[31] G. B. Rocha, R. O. Freire, A. M. Simas, J. J. P. Stewart, *J. Comput. Chem.* **2006**, *27*, 1101.
[32] H. Ryu, J. Lee, S. W. Hagerty, B. Y. Soh, S. E. McAlpin, K. A. Cormier, K. M. Smith, R. J. Ferrante, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 19176.
[33] A. Coleman, F. Perret, A. Moussa, M. Dupin, Y. Guo, H. Perron, *Topics Curr. Chem.* **2007**, *277*, 31.
[34] Reported K_d values for Lys(Me_3)-recognizing effector proteins binding their methylated peptide partners range from 0.7–5 μM. See, for example, W. Fischle, Y. Wang, S. A. Jacobs, Y. Kim, C. D. Allis, S. Khorasanizadeh, *Genes Dev.* **2003**, *17*, 1870–1881; P. R. Nielsen, D. Nietlispach, H. R. Mott, J. Callaghan, A. Bannister, T. Kouzarides, A. G. Murzin, N. V. Murzina, E. D. Laue, *Nature* **2002**, *416*, 103–107.
[35] S. A. Jacobs, S. Khorasanizadeh, *Science* **2002**, *295*, 2080.

Received: October 15, 2009
Published online on November 20, 2009

