A EUROPEAN JOURNAL OF CHEMICAL BIOLOGY

CHEMBIOCHEM

SYNTHETIC BIOLOGY & BIO-NANOTECHNOLOGY

Accepted Article

Title: Synthetic receptors for the recognition and discrimination of post-translationally methylated lysines

Authors: Tobias Gruber

To be cited as: *ChemBioChem* 10.1002/cbic.201800398

Link to VoR: http://dx.doi.org/10.1002/cbic.201800398

A Journal of ChemPubSoc Europe

WILEY-VCH www.chembiochem.org

10.1002/cbic.201800398

Synthetic receptors for the recognition and discrimination of post-translationally methylated lysines

Dr. Tobias Gruber*

School of Pharmacy, University of Lincoln, Joseph Banks Laboratories, Green Lane, Lincoln LN6 7DL, United Kingdom.
**E-mail: tgruber@lincoln.ac.uk*

10.1002/cbic.201800398

Synthetic receptors for the recognition and discrimination of post-translationally methylated lysines

Dr. Tobias Gruber, University of Lincoln

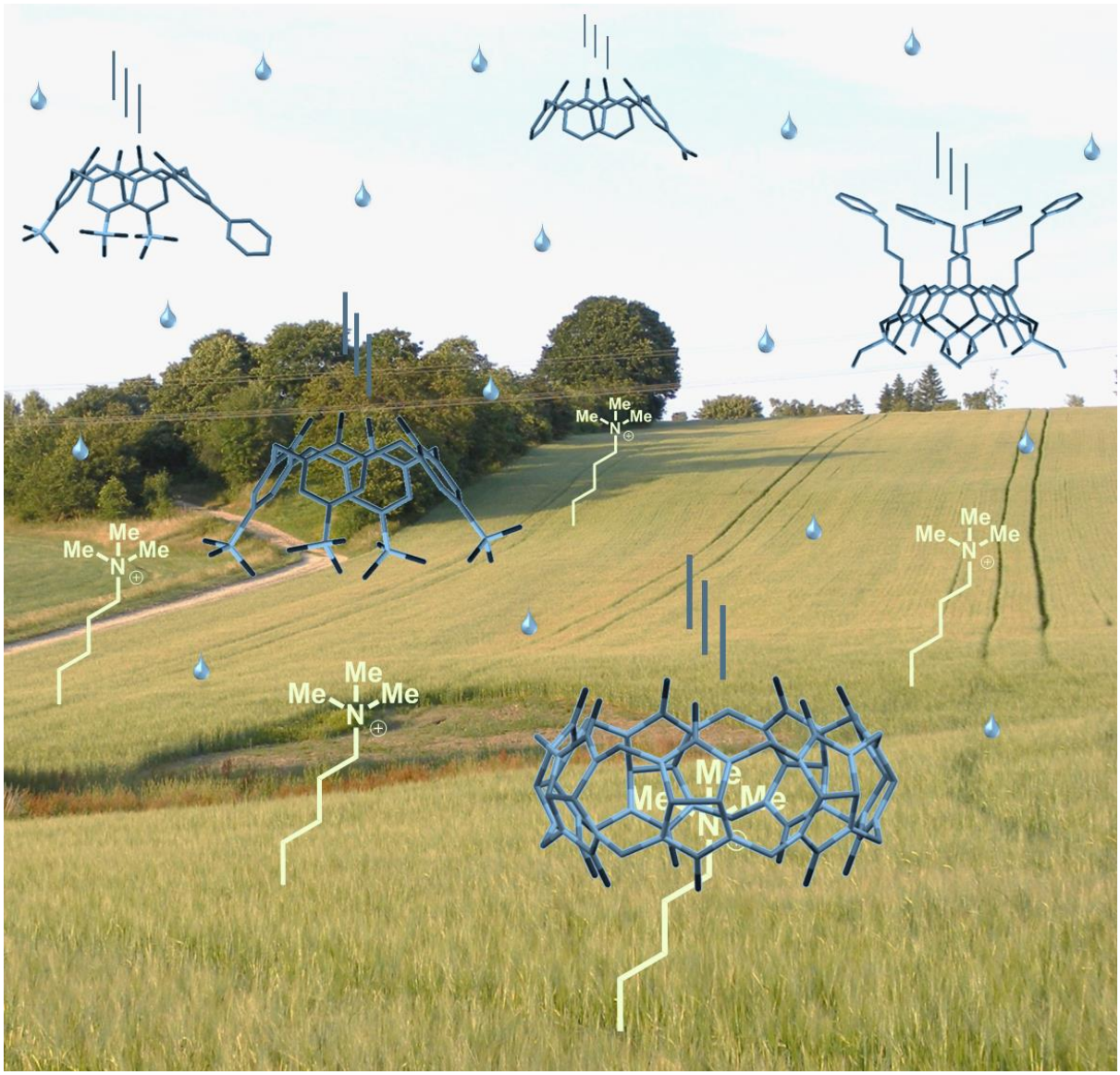


Photo: Tobias Gruber; 50.495 N, 11.467 O

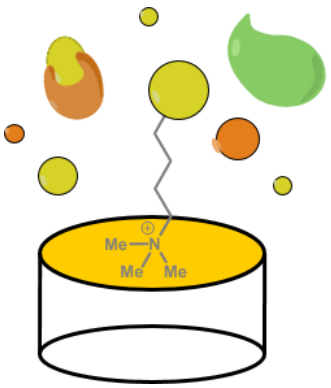
Abstract

Post-translational modifications (PTMs) describe the chemical alteration of proteins after their biosynthesis in ribosomes. PTMs play important roles in cell biology including the regulation of gene expression, cell-cell interactions and the development of different diseases. A prominent class of PTMs is the side chain methylation of lysine. For the analysis and discrimination of differently methylated lysines antibodies are widely used, though, methylated peptide and protein targets are known to be particularly difficult to be differentiated by antibody-based affinity reagents; an additional challenge can be batch-to-batch reproducibility. The application of mass spectrometry techniques for methyllysine discrimination requires a complex sample preparation and is not suited for working in cells. The desire to overcome above-named challenges promoted the development of synthetic receptor molecules that recognize and bind methyllysines. Such 'artificial antibodies' are of interest for a number of applications, e.g. as reagents in biochemical assays, for the isolation and purification of post-translationally methylated proteins and for the tracking of signalling pathways. Moreover, they offer new approaches in diagnostics and therapy. This review delivers an overview of the broad field of methyllysine binding and covers a wide range of synthetic receptors used for the recognition of methylated lysines including calixarenes, resorcinarenes, pillararenes, disulfide cyclophanes, cucurbituriles and acyclic receptors.

Keywords:

Post-translational modification
Lysine methylation
Artificial antibodies
Synthetic receptors
Epigenetics

Graphical Abstract



Very distinguished: The side chain methylation of lysine is a prominent post-translational modification. The desire to understand the function of this modification promoted the development of synthetic molecules that recognize and bind methylated lysines. This review delivers an overview of established methyllysine hosts and provides a summary of their affinities and applications.

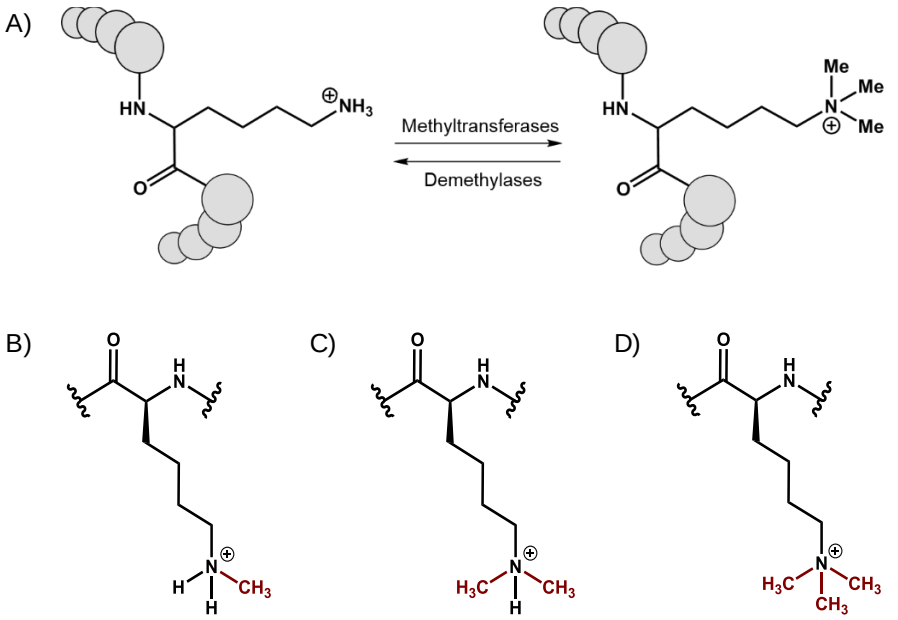
10.1002/cbic.201800398

Tobias Gruber studied chemistry at the University of Freiberg/Saxony and earned his "Diplom" (= M.Sc.) in 2004. Later he carried out research on calixarenes and related receptors in the group of Edwin Weber and received his PhD in 2008. After one year with Manfred Jung in Freiburg/Breisgau were he has been working on histone demethylase inhibitors, Tobias moved to Oxford joining the group of Christopher Schofield where he studied the (bio)syntheses of β- and γ-lactams. In 2011, he started his independent research career in Braunschweig and moved later back to Freiberg. Since 2015, Tobias is a Senior Lecturer in the School of Pharmacy at the University of Lincoln. His research interests include: synthetic receptors as tools for chemical biology and drug delivery, crystal engineering and pre-formulation of APIs, new antimicrobial agents.

1. Introduction

Post-translational modifications (PTMs) describe the chemical alteration of proteins after their biosynthesis in ribosomes. They occur in almost all known proteins and only after post-translational modification is the protein able to fulfil its biological duty in or out of the cell. One can differentiate four different types of PTMs: 1. addition of other proteins or peptides (*e.g.* ubiquitination), 2. change of the chemical nature of an amino acid (e.g. arginine → citrulline), 3. structural alterations (e.g. formation of disulfide bridges) and 4. addition of other functionalities (e.g. methylation, phosphorylation, glycosylation). Together, PTMs play important roles in cell-biology including the regulation of gene expression, cell-cell interactions and development of diseases.[1]

A well-studied class of proteins subjected to PTMs are histones. They are basic, i.e. cationic, proteins found in the eukaryotic cell nuclei and form complexes with the anionic DNA to form chromatin. To date five core histone proteins have been described, viz. H1, H2A, H2B, H3 and H4. Histones are rich in lysine and the accessibility of the DNA for transcription is regulated by PTMs in the lysine side-chains, generally by an interplay of acetylation and methylation.[2] Thereby, methylation and demethylation is catalyzed by respective enzymes. Methyltransferases introduce one, two or three methyl groups to the N_ε positions of different lysine residues resulting in mono-, di- or trimethylation; demethylases promote the loss of methyl groups (Scheme 1).[3] Lysine methyltransferases and lysine demethylases as well as methyllysine binding proteins, e.g. chromodomain- and plant homeodomain-containing proteins, are essential drug targets.[4] One of the largest families of histone binding domains (histone readers) are plant homeodomain (PHD) fingers. They stabilize the chromatin by binding to post-translationally modified and unmodified H3 proteins. PHD fingers control the activity of nuclear enzymes maintaining the physiological PTM equilibrium and, therefore, are crucial for cell homeostasis.

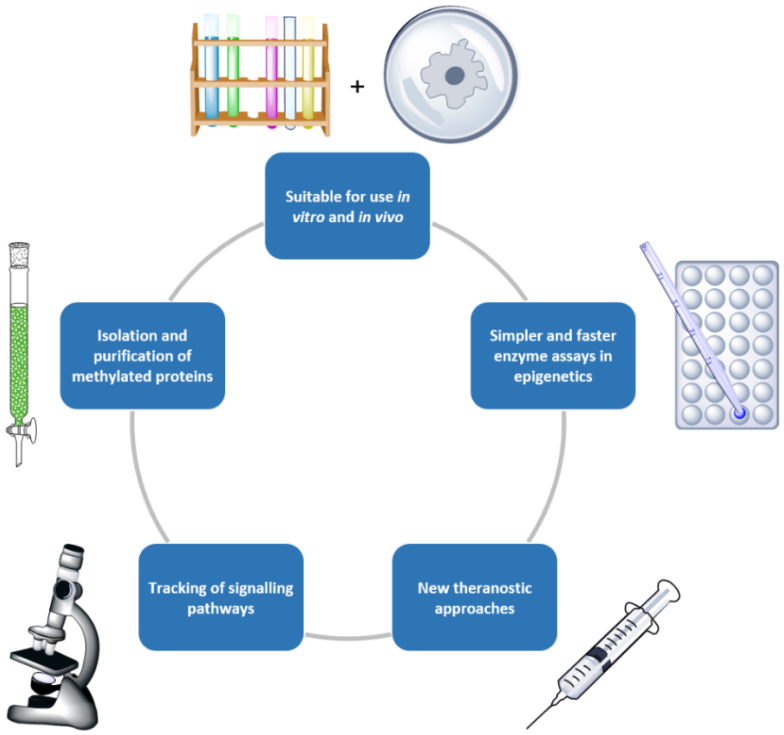


Scheme 1. A) Lysine (de-)methylation is catalyzed by methyltransferases and demethylases leading to B) monomethyl-, C) dimethyl- or D) trimethyllysine. Under physiological conditions all three species appear as ammonium ions.

So far, the best characterized histone methylations are H3K4me$_x$ (histone H3 x times methylated at Lys in position 4), H3K9me$_x$, H3K27me$_x$ and H4K20me$_x$. Methylation at H3K4

is connected with active euchromatin, i.e. actively transcripted genes,[5] and methylation at H3K9 is typically found in gene silencing.[6] H3K27me$_3$ is associated with inactive gene promoters,[7] whereas H4K20 methylation is a key player in genomic integrity.[8] Histone methylation patterns can predict the clinical outcome of a variety of cancers and especially the loss of H4K20 trimethylation is regarded as a hallmark of human cancer.[9] Lysine methylation is also playing a fundamental role in neuronal disorders such as depression. In the absence of stress the brain-derived neutrotropic factor (Bdnf) shows modest levels of histone H3 acetylation and no H3K27 dimethylation. Chronic defeat stress – an animal model of human depression – induces demethylation of histone H3K27. Moreover, related patterns of histone acetylation and methylation play a crucial role in addiction.[10]

The desire to analyze, understand and quantify lysine methylation promoted the development of synthetic molecules that recognize and bind methyllysines.[11] Antibodies against differently methylated lysines are widely used,[12] although methylated peptide and protein targets are known to be particularly difficult to differentiate by antibody-based affinity reagents.[13] Furthermore, most antibodies cannot penetrate cells[14] and they are expensive to produce requiring, in many cases, the use of animals. An additional challenge can be batch-to-batch reproducibility.[15] This leads to massive extra-costs for validating an antibody and to replace ineffective batches.[16] The application of other techniques for methyllysine discrimination such as mass spectrometry generated important tools for PTM analysis,[17] though, often require a complex sample preparation.[18] Mass spectroscopic analyses can be time-consuming and sometimes complicated due to complex mixtures containing low PTM concentrations and large ionized proteins.[19] Additionally, instrument-based techniques are not suited for working in cells.



Scheme 2. Possible applications of synthetic receptors ('artificial antibodies') for post-translational modifications.

Above drawbacks in analyzing and manipulating methyllysines and respective proteins led to the development of synthetic receptors ('artificial antibodies') for the discrimination and

detection of methyllysines in water and biological media as well as in cells.[20] They are well-defined small molecules, stable under ambient conditions and can be produced – more or less cost-efficient – without the need for animals. Actual and potential applications of synthetic receptors for methyllysines are shown in Scheme 2.

In the last years a broad range of synthetic ammonium ion receptors have been reported such as cyclic peptides, crown ethers, calixarenes and cyclodextrins.[21] Thereby, for the recognition of ammonium ions four types of interactions between host and guest are characteristic:

1. **Steric and molecular complementarity**
2. **Ion pairs and salt bridges**
3. **Hydrogen bonds**
4. **Cation···π interactions**

Organic solvents allow strong hydrogen bonding between host and guest resulting in higher associations constants than observed for polar solvents, such as methanol and water.[22] In these more challenging media, hydrogen bonds between host and guest will provide less impetus for binding.[23] As a consequence, other non-covalent interactions are amplified. In case of methyllysines their recognition is based on the complexation of methylammonium ions. Additional methyl groups at the NH_2 moiety in ε-position of lysine does not change the overall charge, however, result in less hydrophilic side chains. Hence, C-H···π interactions and the hydrophobic effect have a larger influence on the complexation.
Obviously, the ideal synthetic receptor for methyllysines will be able to discriminate them from other amino acids and methylamines. However, it will also be able to distinguish between four lysines featuring the most subtle differences in size, shape, basicity and lipophilicity resulting from the presence of 0-3 methyl groups at the N_ε position. Crystallographic studies of methylated lysines complexed by peptides and proteins revealed that di- and trimethylated lysine residues prefer so-called "aromatic cages".[24, 25] The latter were the inspiration for quite a number of synthetic methyllysine receptors. However, some successful candidates not even contain phenyl moieties as discussed later.

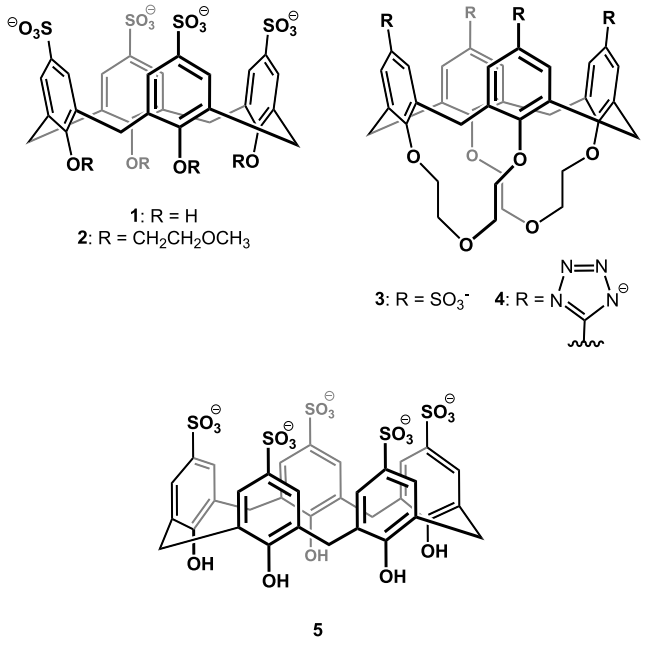
The dissociation constant (K_d) is a measure for the thermodynamic stability of a host/guest complex. It can be used to describe the affinity between receptor (host) and ligand (guest); a high affinity of both binding partners is expressed by a low dissociation constant of the complex. The most commonly used techniques for determining K_d values are: NMR spectroscopy and isothermal titration calorimetry (ITC),[26] fluorescence displacement (FD),[27] fluorescence anisotropy (FA),[28] and surface plasmon resonance SPR.[29] For comparison reasons this review uses only K_d values in µM. For some methyllysine receptors these were calculated from the respective association constant K_a (K_d=1/K_a). A detailed compilation of the dissociation constants for all compounds discussed here can be found in Tables 1-11 and the ESI (Tables S1-S8). As a matter of lucidity the lowest K_d values and highest selectivity are shaded in each table.
Due their rather easy experimental set-up complexation studies in solution only deliver restricted information on the geometry of host/guest complexes. Complementary X-ray structures allow to study interactions between hosts and guests in more detail and give insight into respective conformational parameter such as distances and torsions angles. So far, five complexes of synthetic receptors with methylated lysines have been described in the Cambridge Structural Database (CSD) and the Protein Data Bank (PDB), respectively. They will be discussed in the corresponding sections of this review.

10.1002/cbic.201800398

2. Calixarenes

2.1. Sulfonatocalixarenes

One way to mimic "aromatic cages" found in proteins are aromatic macrocycles with cyclophanes[30] as typical examples. Cyclophanes (=cyclic phenylalkanes) were one of the first synthetic receptors used for the recognition and complexation of alkylammonium ions. A prominent subfamily are so-called calix[n]arenes, *i.e. m*-cyclophanes featuring phenolic groups and a defined cavity; the digit in brackets defines the number of aromatic units.[31] The phenolic region of a calixarene is called lower rim, the opposite section upper rim. The best researched calixarenes are the cyclic tetramers, i.e. calix[4]arenes. They can exist in four different main conformations (cone, partial cone, 1,3-alternate and 1,2-alternate), which are determined by the respective lower rim substituents, temperature, solvent and possible complexing ions (e.g. Na^+ will force a flexible calix[4]arene into the cone form).[32]
Calixarenes and their derivatives are rather easy to synthesize and some show remarkable bioactivity.[33] This is also true for sulfonatocalixarenes,[34] which have already been studied extensively with respect to their inclusion properties in aqueous solution (ammonium ions,[35] amino acids,[36] biological buffer components,[37] neutral species[38]) as well as in solid state.[39] Studies carried out by Hof and co-workers[40] found that *p*-tetrasulfonatocalix[4]arene (**1**) (Scheme 3) is able to discriminate between tri-, di-, mono- and unmethylated lysine with dissociation constants of 27, 62, 250 and 1,923 µM, respectively, and a Kme_0/Kme_3 selectivity of 71 (NMR, sodium phosphate buffer) (Table 1). Using fluorescence displacement assays and glycine buffer even lower dissociation constants (7.7 µM for trimethyllysine) and higher selectivity (>130) have been observed.[41] Primary drivers for the recognition events are favorable enthalpies of binding (electrostatic interactions, non-classical hydrophobic effects).[42]



Scheme 3. Structures of calixarenes **1**-**5** featuring varying lower and upper rim substituents as well as ring sizes.

Further research revealed that biologically more relevant concentrations of sodium and potassium ions in the buffer medium decrease the affinity of methyllysines to **1** by a factor of

2-3; lowering the temperature decreases dissociations constants (ca. 2 μM/K) (Table S1).[43] Complexation studies with methylated and unmethylated peptides showed that both have higher affinities to **1**, though lower selectivity, than observed for the simple amino acids. As one reason repulsive forces between the carboxylate anions of the amino acid guests and the sulfonato group of the host have been discussed. Another reason could be attractive secondary interactions between **1** and the amide backbone and neighboring side chains in the histone peptides.[44] By way of interest, **1** binds H3K4me$_3$, H3K9me$_3$ and H3K27me$_3$ equally well (K_d =5.0-9.1 μM) as determined by ITC and, hence, possess a rather low specificity towards different trimethylated lysines sites.

The interactions of tetrasulfonatocalix[4]arene (**1**) with trimethyllysines are in some cases strong enough to compete with those of natural protein receptors. It has been shown that **1** is able to interrupt the binding of histone H3K4me$_3$ with its native protein binder ING2 PHD. By way of interest the calixarene host disrupts the interaction of ING2 PHD and the trimethylated histone (IC_{50} = 108 μM), though not between ING2 PHD and the dimethylated histone.[45]

Table 1. Dissociation constants for the binding of *p*-tetrasulfonatocalix[4]arene (**1**) to differently methylated lysines and respective histone peptides (40 mM sodium phosphate buffer, pH 7.4).

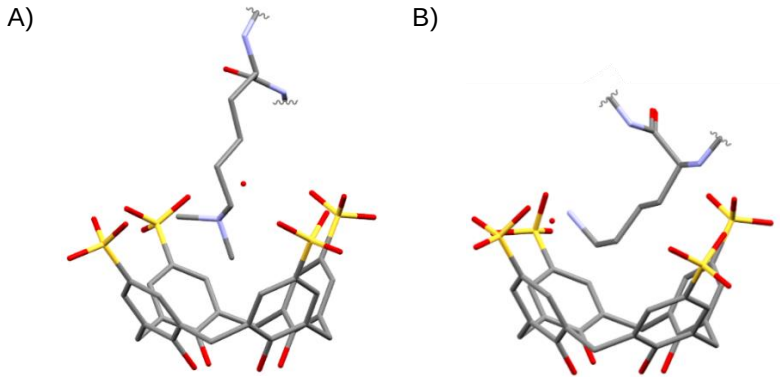
Ref.	Guest	K_d [μM]	S[a]	T [K]	method
[40]	Kme$_0$	1923	-	298	NMR
[40]	Kme$_1$	250	8	298	NMR
[40]	Kme$_1$	333	-	303	ITC
[40]	Kme$_2$	62	31	298	NMR
[40]	Kme$_2$	95	-	303	ITC
[40]	Kme$_3$	27	71	298	NMR
[40]	Kme$_3$	28	-	303	ITC
[42]	H3K4[b]	46		303	ITC
[42]	H3K4me$_3$	5.0	9	303	ITC
[42]	H3K9	101		303	ITC
[42]	H3K9me$_3$	7.2	14	303	ITC
[42]	H3K27	220		303	ITC
[42]	H3K27me$_3$	5.4	41	303	ITC
[42]	H3K36	128		303	ITC
[42]	H3K36me$_3$	9.1	14	303	ITC

[a] S: Kme$_0$/Kme$_x$ selectivity

[b] H3K4 = $^+$H$_3$N-ART<u>K</u>QTAY-CONH$_2$; H3K9 = Ac-TAR<u>K</u>STGY-CONH$_2$; H3K27 = Ac-AAR<u>K</u>SAPY-CONH$_2$; H3K36 = Ac-GGV<u>K</u>KPHY-CONH$_2$

The complexes of tetrasulfonatocalixarene (**1**) with dimethyllysine hen egg-white lysozyme[46] and cytochrome c containing nonmethylated lysines[47] give interesting insights into the driving forces for methyllysine recognition. In the asymmetric unit of the dimethyllysine hen egg-white lysozyme calixarene complex two host molecules bind to dimethyllysine residues and two bind to arginine residues. The calixarenes that complex the dimethyllysine are in a symmetrical cone conformation with interplanary angles of 61.7/62.7 ° and 65.1/66.6 °, respectively (Scheme 4a). One of the N_ε methyl groups of each guest points directly into the calixarene cavity. The distance between the carbon atom and the centroids of the calixarene phenyl rings range from

3.51 to 3.91 Å (four C-H···π interactions). The second methyl group points towards two of the sulfonato residues [d(C···O)=3.57, 3.46 Å] and develops only one C-H···π interaction (3.89 Å) with the aromatic cavity.
The asymmetric unit of the complex of **1** with cytochrome c contains three guest molecules, which all cap the protein at lysine residues. The calixarene host is again in a cone conformation, though its geometry is somewhat distorted (interplanary angles 51.4/86.7 °, 58.4/80.4 °, 59.0/88.9 °). The lysine guests are interacting with the receptor cavity via their alkyl backbone. The found C-H···π interactions involve carbon atoms C_ε (2-3 contacts per host/guest unit, 3.27-3.99 Å) and C_δ (1-2 contacts per host/guest unit, 3.73-3.97 Å) (Scheme 4b). The ε-ammonium unit is bend towards two the sulfonate groups preventing cation···π interactions; a similar feature is observed in the complex of lysine with **1** (CSD code: WIXSOL)[48]. Taking under consideration the rather unspecific inclusion of lysine, dimethyllysine and arginine residues, the binding behaviour of **1** seems to be rather 'promiscuous' – at least in the solid state.



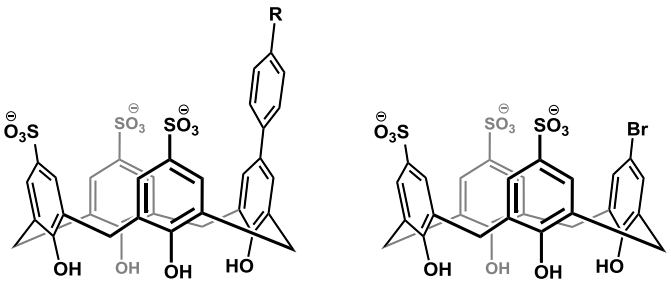
Scheme 4. Details of the X-ray structures of tetrasulfonatocalixarene **1** with A) dimethyllysine hen egg-white lysozyme (PDB code: 4N0J/4PRU)[47] and B) cytochrome c (PDB codes: 3TYI;[46] unpublished alternative binding mode: 4YE1). (For both complexes only one host/guest unit is shown. Hydrogen atoms are omitted for clarity.)

The conformation of the calixarene host is of crucial importance for its inclusion behaviour. The etherified calixarene **2** (Scheme 3) is in a so-called pinched cone conformation. Its closed cavity prevents efficient trimethyllysine uptake. For calixarenes **3** and **4** the ethylene glycol handles facilitate an open cavity leading to H3K27me_3 affinities of 85 and 20 μM, respectively.[42] Also the ring size of the calixarene has been varied. The resulting hexasulfonatocalix[6]arene (**5**) still possesses a moderate selectivity for trimethylated lysine over lysine (Kme_0/Kme_3=5), though presents only low affinities towards both (5,000 μM for Kme_0 and 1,074 μM for Kme_3) (Table S2).[43]

In order to further increase the affinity of *p*-tetrasulfonatocalixarene **1** towards higher methylated lysines, Hof and co-workers suggested to exchange one SO_3^- group against a phenyl moiety or a bromine atom.[49] The resulting trisulfonato receptors (**6-12**) (Scheme 5) have the potential to develop additional C-H···π interactions towards the backbone of the lysine guests. Within the series of receptors **6-12** only the calixarene featuring the underivatized phenyl substituent (**6**) showed higher affinity (16 μM) and selectivity (150) towards trimethylysine in comparison to **1** (27 μM, 71); no information has been given for the affinities towards the recognition mono- and dimethyllysines. In later studies hosts **6-8**, **11** and **12** have been employed in fluorescence displacement assays to study their affinity to histone peptide H3K27 in its trimethylated and unmethylated form.[50] As observed for mother compound **1**, the affinities for the peptides are much higher than for the isolated amino acids (0.34-1.86 μM),

10.1002/cbic.201800398

though the Kme_0/Kme_x selectivity decreases in most cases (e.g. from 150 to 25 for **6**) (Table 3).



6: R=H **9**: R=COOH
7: R=CN **10**: R=CH_2NH_2
8: R=$CONH_2$ **11**: R=Me

12

Scheme 5. For an improved binding of higher methylated lysines one SO_3^- group in **1** has been exchanged against a phenyl moiety or a bromine atom leading to trisulfonato calixarenes **6-12**

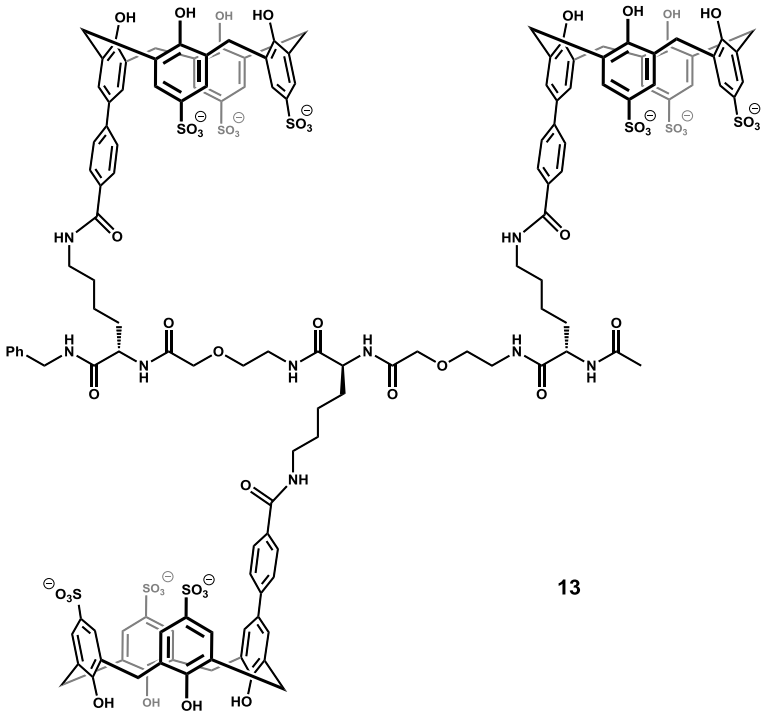
Table 3. Dissociation constants for the complexes of trisulfonatocalixarenes **6-12** with lysine and trimethyllysine and the respective histone peptide H3K27me$_x$ (*pH* 7.4).

Host	Guest	K_d [μM]	S[a]	solv.	method
6[49]	Kme_0	2380	-	A[b]	NMR
6[49]	Kme_3	16	150	A	NMR
6[49]	Kme_3	13	-	A	ITC
6[50]	H3K27	19	-	B[c]	FD
6[50]	H3K27me$_3$	0.75	25	B	FD
7[49]	Kme_0	4762	-	A	NMR
7[49]	Kme_3	476	10	A	NMR
7[50]	H3K27	11.3	-	B	FD
7[50]	H3K27me$_3$	0.88	13	B	FD
8[49]	Kme_0	7143	-	A	NMR
8[49]	Kme_3	169	42	A	NMR
8[50]	H3K27	23	-	B	FD
8[50]	H3K27me$_3$	1.86	12	B	FD
9[49]	Kme_0	9091	-	A	NMR
9[49]	Kme_3	588	16	A	NMR
10[49]	Kme_0	5000	-	A	NMR
10[49]	Kme_3	192	26	A	NMR
11[50]	H3K27	2.7	-	B	FD
11[50]	H3K27me$_3$	0.34	8	B	FD
12[49]	Kme_0	2273	-	A	NMR
12[49]	Kme_3	256	9	A	NMR
12[50]	H3K27	11.3	-	B	FD
12[50]	H3K27me$_3$	0.88	13	B	FD

[a] S: Kme_0/Kme_x selectivity
[b] A: 40 mM sodium phosphate buffer
[c] B: 10 mM sodium phosphate buffer

Only recently, the interactions of calixarenes **6** and **12** with cytochrome c have been studied in solution and the crystalline state (PDB codes: 5KPF for **6** and 5LFT for **12**).[51] Though only unmethylated lysines are available in cytochrome c, the dissociation constant for the complex with **12** is 20 μM. This is slightly lower than for the complex of cytochrome c with parent calixarene **1** (K_d=28 μM). In the crystalline complexes of **6** and **12** the NH_3^+ moieties of the lysine residues avoid the cavity of the calixarenes similar to the respective X-ray structure of **1**.

Kimura *et al.* employed trisulfonated calixarenes for the design of multivalent ligands.[52] The conjugation of **6** via short amide linkers delivered mono-, di- and trivalent (**13**) receptors (Scheme 6). The di- and trivalent ligands have higher binding affinities for methylated and nonmethylated histones (0.39-0.86 μM) than the monovalent receptor (8.9-20.4 μM) (Table 4). Noteworthy, for the trivalent system the K_d differences of respective complexes with histones H3 and H4, acetylated histone H4KAc and H3K27me$_3$ are only rather small. H3 and H4 are rich in arginine and lysine, which suggests that nonspecific electrostatic interactions between the charged guest and the calixarene host overrule the more discriminating cation···π and C-H···π interactions between the methylated lysines the host cavity.[53] It should also be considered that in this case the peptides were immobilized on a chip and the binding has been studied using surface plasmon resonance (SPR).



Scheme 6. For trivalent receptor **13** three molecules of calixarene **6** have been conjugated via amide linkers. The affinity of **13** to trimethyllysine is almost 23 times higher than observed for the respective monovalent host.

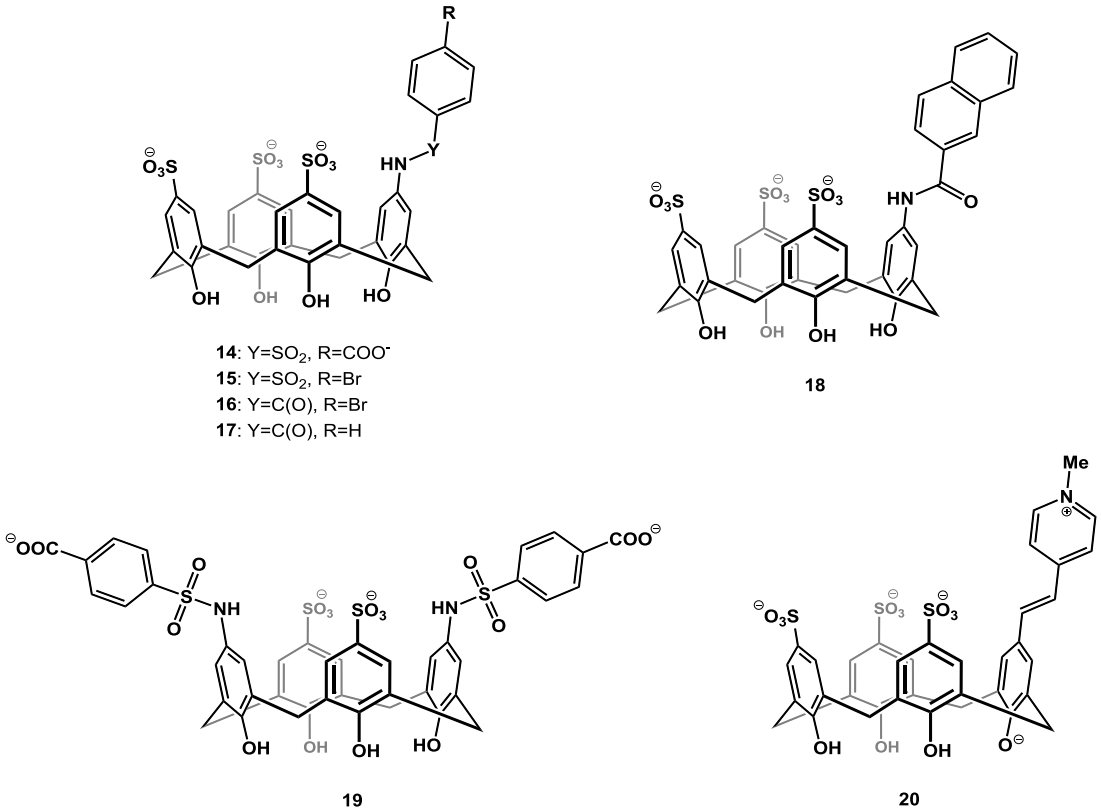
Further research focussed on discriminating lysine trimethylation at specific lysine residues, *viz.* H3K9me$_3$ and H3K4me$_3$. For that purpose a series of calixarene amides and sulfonamides (**14-19**) has been synthesized (Scheme 7).[44],[45] They feature a somewhat higher flexibility in comparison to the biphenyl calixarenes discussed above. Most (sulfon)amide receptors bind to H3K9me$_3$ and H3K4me$_3$ with dissociation constants below 1 μM as determined by fluorescence displacement assays (**1**: K_d=0.12 μM [H3K9me$_3$], 0.02 μM [H3K4me$_3$]; **14**, **16**-

18: K_d=0.14-0.49 µM [H3K9me$_3$], 0.02-0.09 µM [H3K4me$_3$]). Only for hosts **15** and **19** lower affinities have been found (**15**: K_d=4.8 µM [H3K9me$_3$], 7.8 µM [H3K4me$_3$]; **19**: K_d=0.51 µM [H3K9me$_3$], 1.6 µM [H3K4me$_3$]). The H3K9me$_3$/H3K4me$_3$ selectivity is rather low for **15**, **18** and **19** (0.3-1.6) and somewhat higher for **14** (9.5) and **17** (9.8), which is similar to the one of calixarene **1** (7.8) (Table S3). (N.B.: Due to the different analytical methods applied a direct comparison with data from Tables 1 and 2 is difficult.)

Table 4. Dissociation constants K_d [µM] for receptor **6** as mono-, di - and trivalent amide conjugate to H4 and H3 tails (SPR, chip-immobilized peptides).

	Peptide			
Host	H4[a]	H4KAc[b]	H3[c]	H3K27me$_3$[d]
Monovalent	17.8	10.0	20.4	8.9
Divalent	1.8	3.3	1.8	0.39
Trivalent (**13**)	0.73	0.86	0.51	0.39

[a] H-SGRGKGGKGLGKGGAKRHRKGGK(biotin)-NH$_2$
[b] H-SGRG-K(Ac)-GG-K(Ac)-GLG-K(Ac)-GGA-K(Ac)-RHR-K(Ac)-GGK(biotin)-NH$_2$
[c] Ac-RKSTGGKAPRKQLATKAAR-Kme$_0$-GGK(biotin)-NH$_2$
[d] Ac-RKSTGGKAPRKQLATKAAR-Kme$_3$-GGK(biotin)-NH$_2$

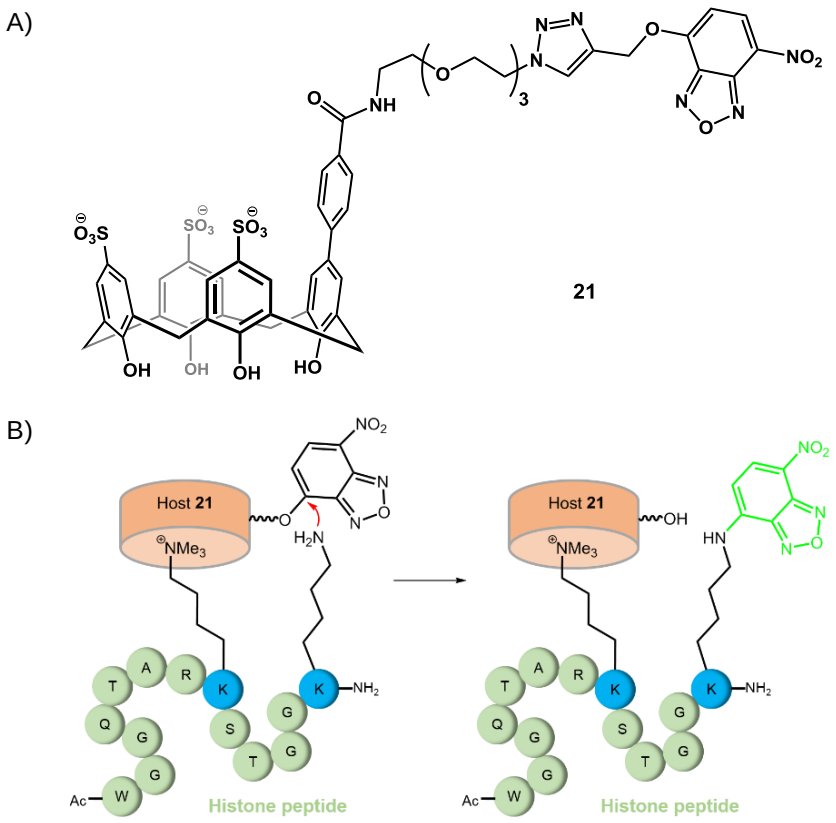


Scheme 7. Structures of sulfonatocalixarenes **14**-**20** featuring one or two aromatic substituents at the upper rim.

By way of interest, calixarenes **1** and **14**-**16** are able to disrupt the binding between the natural histone binder CHD4 PHD2[54] and histone H3K9me$_3$ (K_d=0.9 µM) *in vitro* without disturbing the interaction between CHD4 and the unmethylated histone H3K9me$_0$ (K_d=19 µM);

compound **14** showed the highest activity. Furthermore, all four compounds are active in disrupting heterochromatin markers in cells.[45]
Due to its high affinities for trimethylated peptides, calixarene **14** has been employed in supramolecular affinity chromatography for methylation-targeted proteomics. Linking **14** to agarose beads allowed the resolution of histone peptides (H3K4me$_x$, H3K27me$_x$ and H3K27me$_x$) on the basis of their methylation.[55, 56]

The recognition and purification of methyllysines with calixarenes is a crucial achievement, though also their quantification has been in the focus. In 2013, a macrocyclic sensor array and its method of use have been patented.[57] Thereby, a macrocycle is connected to a dye for identifying histone-code-related analytes. Several macrocycles (e.g. calix[n]arenes, cyclodextrines, cucurbit[n]urils), analytes (e.g. methyllysines, methylarginines, phosphotyrosines) and fluorophores (e.g. fluorescein, dansyl, pyrene) have been disclosed. Only recently, Beatty *et al.* presented trisulfonatocalixarene **20** (Scheme 7) as a tool for photochemical sensing of trimethyllysines in biological media.[58] Macrocycle **20** features intrinsic fluorescence[59] and the recognition and sensing processes tolerate various salts, metal ions and enzymatic cofactors.



Scheme 8. A) Structure of NBD-labelled host **21**. B) After the selective binding of **21** to a trimethylated site of the peptide, a free lysine residue nearby reacts with the NBD moiety in an S$_N$Ar reaction.

Gober and Waters used trisulfonatocalixarene **21** for the affinity labelling of Kme$_3$-containing histone peptides (Scheme 8). Thereby, the receptor unit of the molecule complexes trimethyllysine in the peptide guest and the nitrobenzoxadiazole (NBD) group acts as the reagent covalently labelling a nonmethylated lysine in the same peptide (S$_N$Ar mechanism). The selectivity and rate of the labelling reaction proved to be significantly dependent on salt

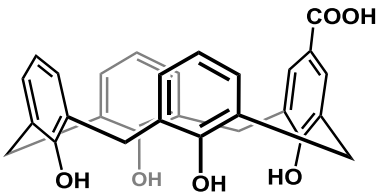
10.1002/cbic.201800398

and reagent concentration as well as pH. The utility of this new tool has been demonstrated in a turn-on fluorescence HDAC assay.[60] NBD is a particularly attractive fluorophore as its photophysical properties allow the use of a fluorescein isothiocyanate (FITC) filter.[61]

2.2. Carboxycalixarenes

Leung, Gruber and co-workers described a simple, readily synthesized monocarboxycalixarene (**22**) that selectively binds to di/trimethylammonium groups (Table 6).[62] In comparison to respective tri- and tetrasulfonatocalixarenes a monocarboxycalixarene is bearing only a single charge with possible advantages with respect to cell permeability. Furthermore, the complex of a monocarboxycalixarene and lysines will be neutral, which could be of interest for the extraction or the delivery of (methylated) lysines. Carboxycalixarene **22** binds di- and trimethyllysine with dissociations constants of 70 µM and 60 µM, respectively, though the affinities towards non- and monomethylated lysines are very low (K_d>500 µM) (Table 5). Receptor **22** is able to recognize methyllysines even in complex mixtures such as *E. coli* cell lysate as demonstrated by NMR spectroscopy.

Table 5. Dissociation constants of carboxycalixarene **22** with differently methylated lysine guests (300 K, NMR, 50 mM sodium phosphate buffer, *pH* 7.5).

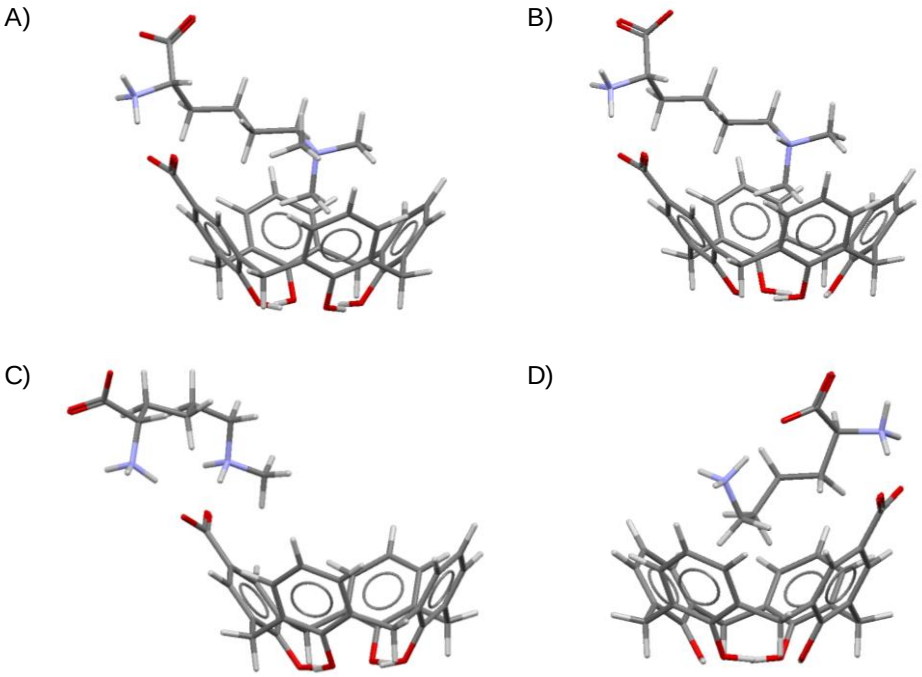


22

Guest	K_d [µM]
Kme_0	>500
Kme_1	>500
Kme_2	70
Kme_3	60
PATGGV-Kme_1-KPHRY	>500
PATGGV-Kme_2-KPHRY	60
PATGGV-Kme_3-KPHRY	65
AR-Kme_1-STGGK	>500
AR-Kme_2-STGGK	50
AR-Kme_3-STGGK	50
choline	65
carnitine	60
meldonium	50
4-dimethylaminopyridine (DMAP)	95

The energy-minimized complexes of **22** with methyllysines reveal, that only for higher methylated lysines cation···π-interactions contribute to the recognition event (Scheme 9). Furthermore, C-H···π-interactions seem to play a vital role for the formation and the stability of the host/guest complexes. Di- und trimethyllysine develop four and five of these contacts, respectively. Thereby, all of the *N*-methyl groups are involved in the interaction. In the case of monomethyllysine two N-H···O-hydrogen bond prevents the occurrence of C-H···π- and cation···π-contacts.

In the energy-minimized structure of **22** with lysine three C-H···π-interactions and an intramolecular N-H···O-hydrogen bond involving the guest carboxylate and the ε-ammonium unit has been observed. The N-H···O prevent a close contact of the ammonium cation with the π-electron rich cavity. Such a behavior has already been found in the X-ray structure of lysine in the complex with tetrasulfonatocalixarene (**1**),[48] cytochrome c[47] and hexaphosphonatocalix[6]arene (PDB code: 5LYC),[63] respectively.



Scheme 9. The energy-minimized host/guest complexes of **22** with trimethyllysine (A), dimethyllysine (B), monomethyllysine (C) and lysine (D) demonstrate the preference for higher methylated lysines. Only in A) and B) stabilizing C-H···π and cation···π interactions are observed. (MacroModel V.9.8; OPLS_2001 force field; MCMM; solvent: water; 20,000 steps.)

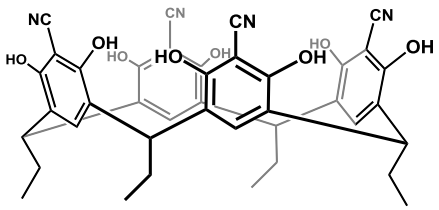
In contrast to tetra- and trisulfonatocalixarenes the binding of **22** to the di- and trimethyllysine motif does not improve significantly when changing from free amino acids to peptides (K_d=50-65 μM) (Table 5). The peptide PATGGV-Kme$_3$-KPHRY is actually bound even worse than the free amino acid – despite the fact that the trimethyllysine has a direct lysine neighbor in the peptide. Host **22** only contains one anionic charge at the upper rim, though recognizes di- and trimethyllysines in the same order of magnitude as tetrasulfonatocalixarene **1**. Hence, the repulsive forces between host and the amino acid carboxylate are much lower in **22** compared to **1**. As a consequence, the more or less unchanged affinities between **22** and the methyllysine peptides must be primarily attributed to the lack secondary interactions between host and guest.

2.3. Structure-activity relationship of calixarenes

Due to the diverse analytical methods employed for the determination of the dissociation constants, structure-activity relationships of the different calixarene hosts needs to be discussed cautiously. Here, additional work seems necessary in order to gain more reliable data. In general, it can be stated that higher charged receptors lead to higher affinities and selectivity (**1** vs. **22**). Calix[4]arenes are favoured over calix[6]arenes (**1** vs. **5**) and alternative anionic groups at the upper rim lead to similar affinities for trimethyllysines.
In tetrasulfonatocalixarenes, the substitution of one SO_3^- group against a non-functionalized phenyl substituent improves the binding of trimethyllysine and the Kme_0/Kme_x selectivity (**1** vs. **6**). However, functionalized phenyl substituents as in **7**-**11** result in drastically decreased binding affinities and selectivity. In contrast, for histone peptide H3K27me$_3$ the electronically-activated *p*-methylphenyl group at the upper rim (**11**) delivered the highest affinity. The phenyl moiety can also be attached to the calixarene via amide and sulphonamide bonds (**14**-**19**). The additional flexibility resulted in highly active receptors for histone peptides H3K4me$_3$ and H3K9me$_3$. It is somewhat surprising that also the plain tetrasulfonatocalixarene **1** is able to complex both peptides with nanomolar affinity.

3. Resorcinarenes

Similar to calixarenes, resorcinarenes are *m*-cyclophanes, though bear two phenol groups at each aromatic moiety instead of one. Resulting from their synthesis most resorcinarenes feature additional alkyl groups at the methylene bridges. Their cavity size is comparable to those of calixarenes. Four of the eight phenolic protons of resorcinarenes can be dissociated at pH >12. The introduction of electron-withdrawing groups (such as –CN) increases the OH acidity of the resorcinarene resulting in a p*K*a in the physiological pH region.[64] Chen and co-workers deployed this phenomenon in tetracyano receptor **23** (Scheme 10), which proved to be a powerful receptor for tetraalkylammonium ions with dissociation constants of 0.9-1.7 µM.[65] Hamilton and co-workers further studied this receptor as methyllysine host revealing a dissociations constant of 21 µM for Kme$_3$, 68 µM for Kme$_2$, 476 µM for Kme$_1$ and over 1,000 µM for nonmethylated lysine.[66] These are comparable to the values found for tetrasulfonatocalixarene (**1**). Both receptors, **1** and **23**, have been successfully screened for their potential to inhibit the KDM4A(=JMJD2A)-catalysed demethylation of a histone peptide (H3K9me$_3$) *in vitro*.

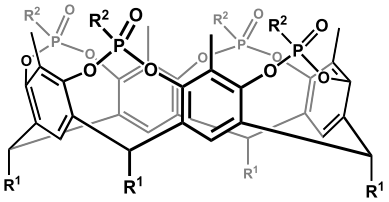


23

Scheme 10. Structure of tetracyanoresorcinarene **23**. The electron-withdrawing cyano groups result in a lower p*Ka* compared to the mother resorcinarene and, hence, improve receptor solubility and the binding of alkylammonium ions.

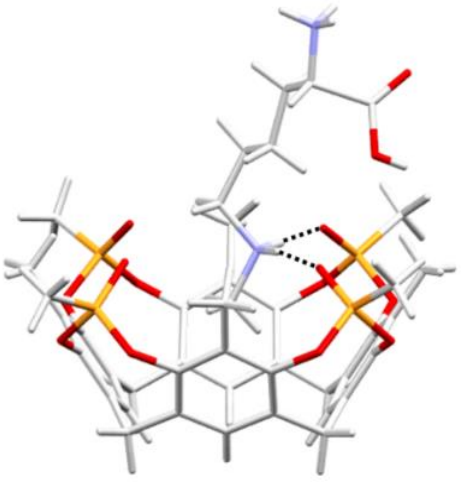
10.1002/cbic.201800398

Despite their strong O-H···O hydrogen bonds between adjacent OH functions resorcinarenes are quite flexible molecules; bridging the phenol moieties leads to more rigid cavitands. An example are tetraphosphonate hosts in which two OH groups of neighboring arene units are each connected via a P=O bridge. This family of cavitands proved successful as synthetic receptors for *N*-methylammonium salts as reported for receptor **24**.[67] (Scheme 11) Interestingly, the authors could also show that for tetraphosphonate cavitands the depth of insertion of an *N*-Me group into a host cavity – determined by X-ray crystallography – can be correlated to the binding constant. The affinity for the N^+-Me group originates from three types of non-covalent interactions between host and guest: a) cation···dipole interactions (N^+···O=P), b) cation···π interactions of the methyl group with the aromatic cavity and c) two hydrogen bonds between the two nitrogen protons and two adjacent P=O bridges. By way of interest, these features are also found in the X-ray structures of respective tetraphosphonates with monomethyllysines as reported by Geremia, Dalcanale and co-workers,[68] viz. **25** in its complex with methyllysine · 2 HCl · 6.4 CF_3CH_2OH · H_2O (CSD code: OJISEH) and **26** in its complex with methyllysine · HCl · H_2O (CSD code: IKOZUF) (Scheme 12).



24: R^1=Pr, R^2=Ph
25: R^1=H, R^2=Me
26: R^1=$CH_2CH_2CH_2Py^+Cl^-$, R^2=Et
27: R^1=Pr, R^2=Et

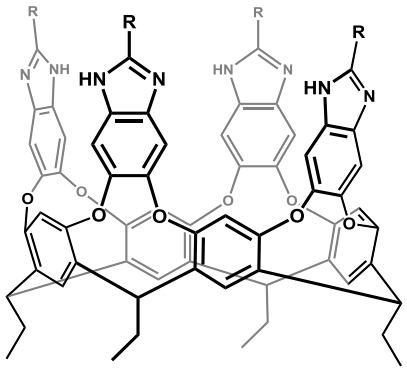
Scheme 11. Tetraphosphonate cavitands **24**-**27** feature different alkyl residues at the methylene bridge (R^1) and the phosphonate ester (R^2).



Scheme 12. X-ray structure of **25** in its complex with methyllysine · 2 HCl, CF_3CH_2OH and water (1:6.4:1), CSD code: OJISEH. (Only the cavitand and the methyllysine guest are shown.)

For cavitands **26** and **27** the dissociations constants of their complexes with monomethyllysine have been determined. Unsurprisingly, the complex in MeOH (**27**: K_d=0.9 μM) has a higher stability than those in aqueous media (**26**: K_d=671 μM in H_2O; 885 μM in NaCl-containing sodium phosphate buffer) (Table S5). Data for the complexation of non-methylated or di- and trimethylated is not given in the reference.
In the following tetraphosphonate cavitands have been applied for the development of an analytical platform based on plasmon-free surface enhanced Raman scattering (SERS). Using this approach Alessandri *et al.* have been able to distinguish *N*-methyllysine hydrochloride from lysine hydrochloride in water.[69] Later a tetraphosphonate cavitand for the recognition of monomethyllysine in histone peptides has been introduced. It allows the discrimination of single monomethylated peptides from multi monomethylated ones.[70]

Hooley, Zhong and co-workers introduced a fluorescence-based supramolecular tandem assay for the *in situ* monitoring of a lysine demethylase (JMJD2E) or methyltransferase (PRDM9).[71] This site-selective displacement assay system contains only three resorcinarene-based cavitands (**28**-**30**) (Scheme 13). As all three have different charges their combination allows the simultaneous investigation of different methylation sites, e.g. peptide sequence AR-Kme_3-ST (H3K9me_3) over T-Kme_3-QTA (H3K4me_3) and AAR-Kme_3-S (H3K27me_3). (N.B.: Receptor **30** featuring four positive charges is able to recognize an alkylammonium ion!)

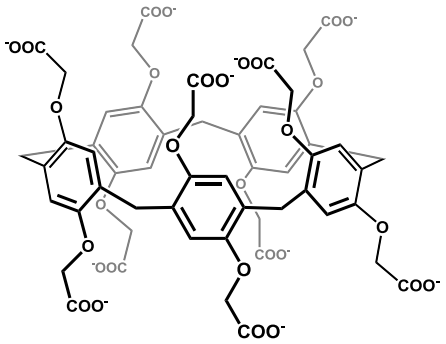


28: R=H
29: R=CH_2COO^-
30: R=NH_3^+

Scheme 13. Structures of cavitands **28**-**30**. Receptor **28** and related hosts[72] are well known for their high affinities towards alkylammonium ions.[73]

4. Pillararenes

Not all electron-rich macrocycles have high affinities to alkylammonium ions and are suited to discriminate methyllysines as demonstrated by pillararene **31** (Scheme 14).[74] It features ten carboxyl groups as well as a stable cavity and is able to bind basic amino acids such as Lys, Arg and His with dissociation constants of 555 μM, 169 μM and 667 μM. Driving forces are electrostatic interactions between COO^- and the cationic side chains and hydrophobic interactions. The complex of **31** with trimethyllysine was found to be less stable as KMe_3 is not able to act as a hydrogen bond donor in contrast to Lys.

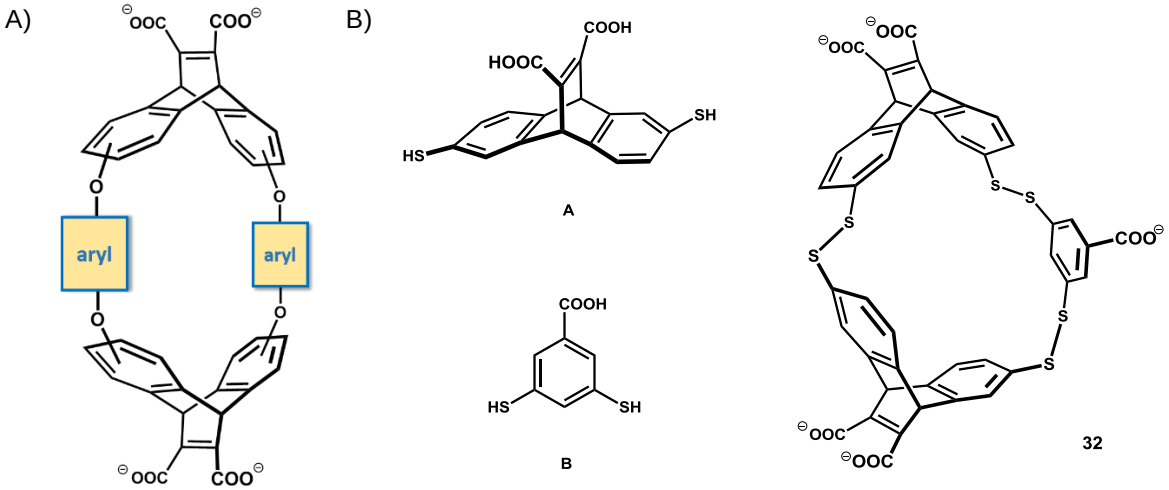


31

Scheme 14. Pillararene **31** has a stable, electron-rich cavity though binds lysine (555 μM) with a higher affinity than trimethyllysine (769 μM).

5. Disulfide cyclophanes

Already back in 1990 Dougherty *et al.* developed a cyclophane consisting of two xylene and two ethenoanthracene units connected via ether bridges. It binds acetylcholine with a dissociation constant of 50 μM, a value comparable to those of biological recognition sites.[75] In the following the xylene units have been replaced by a series of (hetero)aromatic units (Scheme 15a).[76] The resulting receptors favor peralkylated ammonium ions over lower alkylated ammonium ions due to the higher desolvation barriers for the latter. Later, the group demonstrated the importance of cation···π interactions for the complexation of alkylammonium ions.[77] In 2008, Otto and co-workers[78] employed the – appropriately functionalized – building blocks of the first Dougherty receptor for the composition of a dynamic combinatorial library (DCL).[79] The components of this library proved to be disulfide macrocycles with **32** as one example (Scheme 15b). All its members have been studied towards its ability to recognize peralkylated ammonium ions, which were also used as templates in the reaction.



Scheme 15. A) General structure of Dougherty's ethenoanthracene receptors. B) Structure of disulfide cyclophane **32** found in a dynamic combinatorial library (DCL) with **A** and **B** as only building blocks.

Waters and co-workers used **A**, **B** and related aromatic dithiols to create a dynamic combinatorial library similar to the one of Otto, though, in this case dipeptide Ac-Kme$_3$-G-NH$_2$

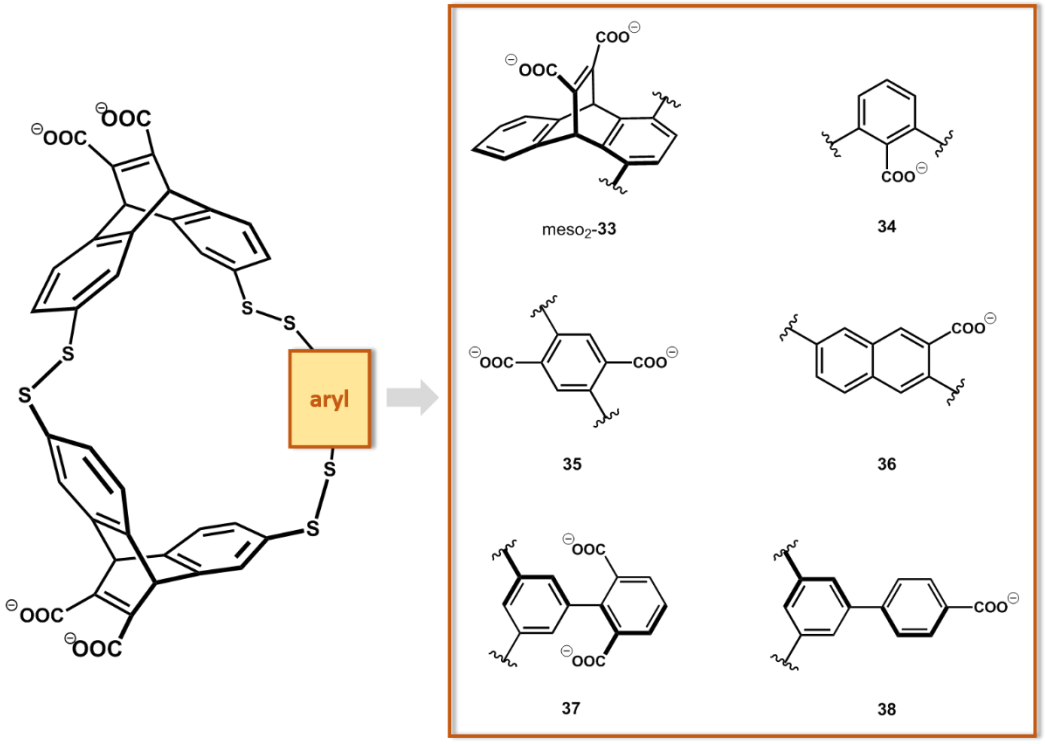
has been used as molecular target. By way of interest, again receptor **32** (Scheme 15c) was the most amplified and was found – after separation and purification – to bind methylated lysines. The binding affinity for histone H3K9me$_3$ is about 25 μM, which is similar to the binding of native HP1 chromodomain, a biological methyllysine receptor (Table 9).[80] Moreover, **32** can discriminate differently methylated lysines.[81] Interestingly, the binding strength and selectivity varies quite drastically depending on the method used for its determination, *e.g.* 25 μM *vs.* 2.6 μM and >48 *vs.* 8.5 for histone H3K9me$_3$ (fluorescence anisotropy *vs.* ITC).[82] The authors explain this with incomplete desalting of the samples measured by fluorescence anisotropy.

Table 9. Dissociation constants for the binding of disulfide macrocycle **32** and HP1 chromodomain to histone peptide H3K9 in its different methylation states

Host	Guest	K_d [μM]	S[a]	solv.	method
rac-**32**[81]	H3K9me$_0$	>1200	-	A[b]	FA
rac-**32**[81]	H3K9me$_1$	166	>7	A	FA
rac-**32**[81]	H3K9me$_2$	58	>20	A	FA
rac-**32**[81]	H3K9me$_3$	25	>48	A	FA
rac-**32**[82]	H3K9me$_0$	22	-	B[c]	ITC
rac-**32**[82]	H3K9me$_1$	13.9	1.6	B	ITC
rac-**32**[82]	H3K9me$_2$	6.3	3.5	B	ITC
rac-**32**[82]	H3K9me$_3$	2.6	8.5	B	ITC
HP1[80]	H3K9me$_0$	>1000	-	C[d]	FA
HP1[80]	H3K9me$_1$	96	>10	C	FA
HP1[80]	H3K9me$_2$	15	>66	C	FA
HP1[80]	H3K9me$_3$	10	>100	C	FA

[a] S: Kme$_0$/Kme$_x$ selectivity
[b] A: 10 mM phosphate buffer; 300 K; *pH* 8.5
[c] B: 10 mM borate buffer; 299 K; *pH* 8.5
[d] C: phosphate buffer, 25 mM NaCl, 1 mM DTT; 288 K; *pH* 7.5

In order to increase and vary the binding affinities and selectivity of receptor **32**, the Waters group used the DCL approach to generate macrocycles **33-38** (Scheme 16). They have impressively high affinities to histone peptide H3K9 in its different methylation forms ranging from 0.13-2.6 μM (trimethylation), 0.18-13.2 μM (dimethylation), 1.0-40 μM (mono-methylation) and 1.8-58 μM (unmethylated peptide) (Table S6).[82],[83],[84] For a broad application of disulfide receptors as trimethyllysine sensors a late stage modification has been suggested.[85] As other synthetic receptors present here the hosts from dynamic combinatorial libraries have been employed in fluorogenic sensor platforms using an indicator displacement system.[86]



Scheme 16. The variation of the aryl unit in the Waters receptor (**32**) produce disulfide macrocycles **33**-**38**. Two COOH functions at the exchangeable aromatic moiety as in **33**, **35** and **37** result in lower dissociation constants (0.13-0.30 µM) and – in most cases – better selectivity (8.1-35) compared to **32** (K_d=2.6 µM; 8.5).

Disulfide cyclophanes **32**-**38** show interesting structure-activity relationships primarily based on the location and number of COO^- functions at the individual aryl units. When the position of the carboxylate in **32** is changed from *meta* to *ortho* (**34**) or a phenylene residue is inserted (**38**) the affinity towards methyllysines is scarcely affected (change from 2.6 to 2.3 and 2.6/2.2, respectively). In case of **34** the Kme_0/Kme_3 selectivity stays more or less the same (change from 8.5 to 9.6), for **38** it decreases significantly (2.7/4.7). The exchange of the benzene moiety in **32** against a naphthalene ring – leading to host **36** – results in a much higher Kme_0/Kme_3 selectivity (>58) than observed for **32** (8.5), though a more or less unchanged affinity for trimethyllysine (1.4 µM). An explanation deliver improved C-H···π-interactions facilitated by the naphthalene ring, which benefit the higher methylated lysines. Cyclophanes **33**, **35** and **37** are featuring an additional COO^- group each, which support secondary interactions with the peptide chain. These obviously lead to much lower dissociation constants for the complex with trimethyllysine (0.13-0.30 µM) and very good (**33**, **35**) or good (**37**) Kme_0/Kme_3 selectivity (35, 34 and 16/8.1, respectively).

Most receptors discussed here have been designed to achieve high affinities and selectivity towards trimethylation. Though, one isomer of macrocycle **37** shows a slightly higher affinity for dimethylated histone protein H3K9 (K_d=0.20 µM) than for the trimethylated one (K_d=0.22 µM). This has also been demonstrated for peptide Ac-Kme_x GGY-NH_2 with K_d=3.32 µM (dimethylation) and 4.30 µM (trimethylation), respectively.[84]

Macrocycle **33** is a good example for a synthetic receptor with comparably high affinities for trimethyllysines in different peptides, though different Kme_0/Kme_x selectivity. For histone peptides H3K9 (Ac-WGGG-QTAR-Kme_x-STG-NH_2) and H3K36 (Ac-WGGG-TGGV-Kme_x-KPH-NH_2) the trimethylated lysine is complexed with the same affinity (K_d=0.3 µM).

However, the affinities of **33** for the unmethylated histone peptides are 10.5 μM (H3K9) and ≈70 μM (H3K36).[82]

For receptor **33** the Waters group also studied the binding affinity towards lysine and trimethyllysines under the influence of neighboring arginine and lysine residues, i.e. possible secondary interactions between host and guest. Their research revealed that, in general, the binding improves when additional cationic side chains (Arg or Lys) are present in the peptide. However, the position of the additional cationic amino acid residue has only a little influence on the binding affinities (K_d=0.31-0.62 μM) and the Kme_0/Kme_3 selectivity decreases in most cases (20-58), except for a lysine in direct neighborhood (100) (Table 10).[87]

Table 10. Dissociation constants for the binding of **33** to non- and trimethylated model peptides featuring varying distances of a neighboring Arg and Lys. (299 K, ITC, 10 mM borate buffer, *pH* 8.5)[87]

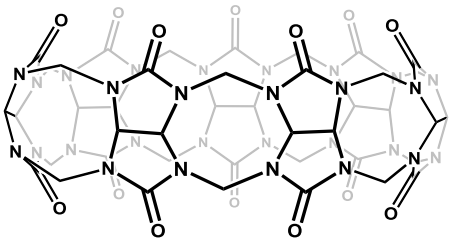
Ac-WGGGG-Z_{i-3}-Z_{i-2}-Z_{i-1}-Kme_x-GGG-NH_2

Z_{i-3}	Z_{i-2}	Z_{i-1}	x	K_d [μM]	Kme_0/Kme_3 selectivity
G	G	G	0	140	67
G	G	G	3	2.1	
G	G	R	0	12	20
G	G	R	3	0.62	
G	R	G	0	13	29
G	R	G	3	0.46	
R	G	G	0	17	34
R	G	G	3	0.50	
G	G	K	0	31	100
G	G	K	3	0.31	
G	K	G	0	23	58
G	K	G	3	0.40	

6. Cucurbiturils

Cucurbit[n]urils are synthetic macrocycles consisting of glycoluril monomers linked by methylene bridges. In contrast to all other synthetic hosts discussed here, cucurbiturils do not contain aromatic moieties. Nevertheless, they have proven as versatile receptors in supramolecular chemistry.[88] Unlike underivatized calixarenes and other cyclophanes cucurbiturils are water-soluble. In the last years mainly cucurbit[7]uril (Scheme 17) has been studied with respect to its ability to form complexes with lysine and other amino acids[89] as well as tetraalkylammonium ions (K_d=1.0-8.3 μM).[90],[91] In 2013, Macartney and co-worker studied the selective molecular recognition of methylated lysines by cucurbiturils with dissociations constants of 0.5 μM, 17 μM, 556 μM, 1,887 μM for tri-, di- and monomethylation as well as lysine, respectively.[92] Interestingly, **39** binds to trimethylated lysine over 3,500 times better than to lysine. This is the highest selectivity observed so far for the recognition of methylated lysines. Thus, cucurbit[7]uril has even a higher selectivity than natural protein receptors (ING2 = 1,500[93], ADDATRX = 7.4[94] / 28[95], HP1 > 100[80]) (Table S7). In cucurbituril complexes the high affinities can clearly not be explained by cation···π and C-H···π interactions, though by the release of high energy water from the hydrophobic cavity (non-classical hydrophobic effect)[96] and ion-dipole interactions.

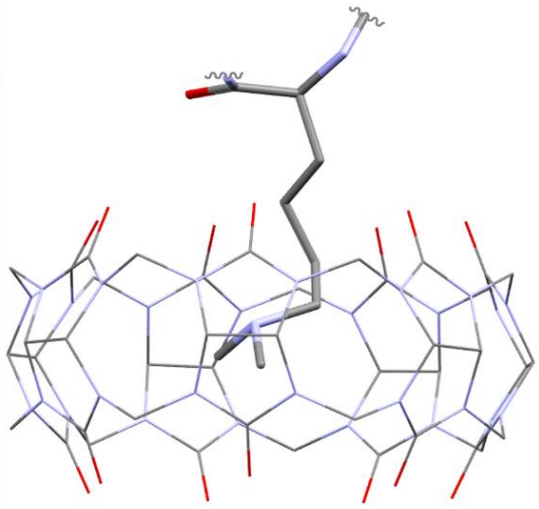
10.1002/cbic.201800398



39

Scheme 17. Cucurbit[7]uril (**39**) is the synthetic receptor with the highest selectivity for trimethylated lysine observed so far.

Recently, Crowley and co-workers described the complex of dimethylated *Ralstonia solanacearum* lectin with cucurbit[7]uril (**39**) (PDB codes: 6F7W/6F7X).[97] In the X-ray structure they found three different modes of binding, suggesting an incomplete filled host cavity (Scheme 18), which could also explain the rather low affinity in solution (K_d≈1,000 μM). Like calixarene **1**, cucurbit[7]uril (**39**) has also been reported to recognize proteins not containing methyllysines with rather high affinity. An example is the complex with human insulin (PDB code: 3Q6E), in which the *N*-terminal phenylalanine residue is preferentially recognized by the cucurbituril over other amino acid sidechains (K_d=0.7 μM).[98]



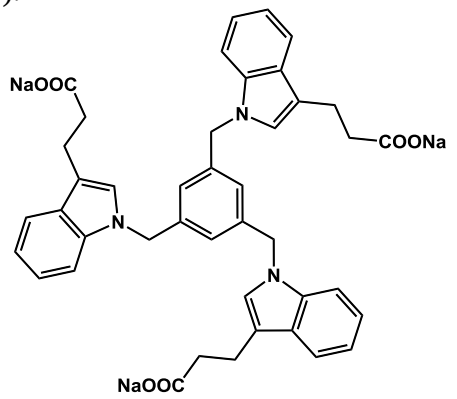
Scheme 18. Detail of the complex of cucurbit[7]uril (**39**) and dimethylated *Ralstonia solanacearum* lectin (PDB code: 6F7W).[97] (Only one host/guest entity is shown. Hydrogen atoms are omitted for clarity.)

Zong and co-workers employed sulfonatocalixarenes **1** and **5** as well as cucurbituril **39** in the separation of methylated histone proteins by host-mediated capillary electrophoresis.[99] The molecular recognition event changes the electrophoretic mobility of the differently methylated peptides. The addition of calixarenes **1** and **5** to the background electrolyte led to their successful separation; cucurbituril **39** has been less effective.

10.1002/cbic.201800398

7. Acyclic receptors

The recognition of amino acids by synthetic receptors is not restricted to cyclic host systems. Early examples for acyclic receptors comprise guanidium receptors[100] introduced by Schmuck and the molecular tweezers[101] of Klärner and Schrader. Other examples of acyclic hosts are tripodal receptors on the base of substituted trimethyl- and triethylbenzenes. These have proven as very successful hosts especially for the recognition of sugars.[102] So far only one tripodal receptor, *viz.* trisindol **40** (Scheme 19), has been studied towards its capacity to complex methylated lysines. However, as **40** shows a low degree of preorganization only unsatisfying affinities towards trimethyllysine have been observed (15,873 μM[103]/4,000 μM[104]). Due to their higher hydrophobicity tetraalkylammonium ions with long alkyl chains are complexed with much higher affinities, *e.g.* tetrabutylammonium chloride with a dissociations constant of 37 μM[103]/142 μM (Table S8).[104]



40

Scheme 19. Trisindol receptor **40** complexes long chain tetraalkyl-ammonium ions with high affinitiy, though, is less suited for the binding of trimethyllysine.

8. Summary and outlook

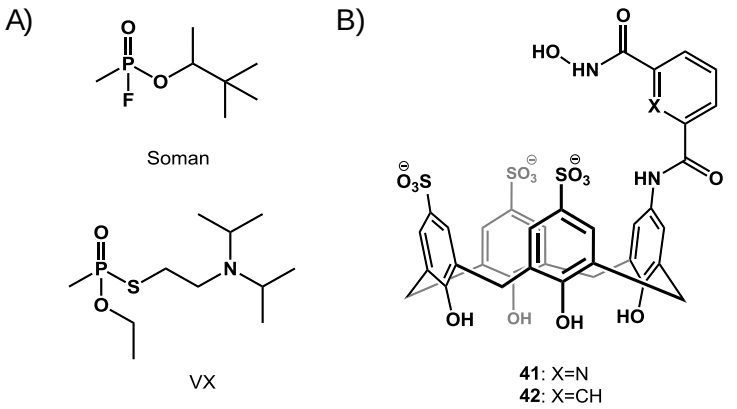
The article on methyllysine recognition by tetrasulfonatocalixarene **1** by Hof and his group in 2010 demonstrated for the first time the feasibility of small molecule receptors as potential hosts for post-translationally modified lysines as free amino acids, in peptides and in proteins. Since then quite a number of receptors for differently methylated lysines has been described. Some of them were already studied before with respect to their ability to bind alkylammonium ions, some were newly introduced. Interestingly, only one family of methyllysine hosts, *viz.* the ethenoanthraceno receptors, is chiral despite the methyllysine guests occur only as single enantiomer – at least in natural peptides and proteins.

In general, it can be stated that methyllysine hosts have higher affinities to methyllysine-containing peptides and proteins than to methyllysines as amino acids. An explanation deliver attractive secondary interactions between the synthetic hosts and the peptide guests. However, in some cases it is problematic to directly compare the binding constants of the various receptors due to varying experimental techniques and conditions in the different references. Hence, it seems not constructive to choose the methyllysine host with the lowest dissociation constant and nominate a 'winner'. Different applications will require different receptor

properties. Depending on the envisaged use overall charge or pH dependency[105] may be more important than selectivity. In some cases competing guests such as N_α-methylamino acids[67] or arginine[106] need to be considered; some receptors may not be stable under the assessed conditions, *e.g.* due to S-S bonds.[107] In some cases the low solubility in water may restricted the applications of the 'synthetic antibodies'. Another possible challenge is the provision of the artificial methyllysine host: tetrasulfonatocalix[4]arene (**1**) and cucurbit[7]uril (**39**) are both commercially available, many others require elaborate synthetic procedures making it more difficult for users beyond chemistry to work with them.

Interestingly, only a rather low number of X-ray structures of methyllysine complexes have been described. Thereby, the binding preferences of the synthetic receptor in solution is not always directly comparable with the situation in solid state – as shown for the rather promiscuous binding of some title compounds. However, X-ray structures of methyllysine inclusion compounds and complexes give valuable information on non-covalent interactions between host and guest and can help to further improve the selectivity and specificity (generic vs. specific receptors).

On the long run, research on hosts for methyllysines will also benefit the development of 'artificial antibodies' for similar post-translational modifications such as methylarginine,[108] methylhistidine,[109] and methyladenin.[110] The general concepts found for methyllysine binding will further stimulate the recognition of related guest species such as methylamines like ecstasy,[111] amphetamines[112] or other illegal substances.[113] Already, synthetic methyllysine receptors have prompted the development of macrocycles for the detoxification of organophosphonates, which can be – and are – misused as chemical warfare agents. By way of example, Kubik and co-workers reported on a series of trisulfonatocalixarenes for the catalytic hydrolysis of V-type nerve agents and soman demonstrating once again the variability of calixarene hosts (Scheme 20).[114]



Scheme 20. Structures of soman and VX (A) as well as catalytic-active trisulfonatocalixarenes **41** and **42** for their accelerated hydrolysis (B).

9. Future research directions and challenges

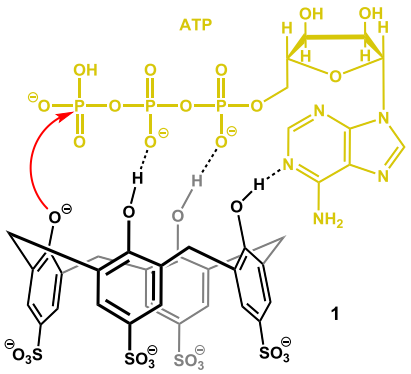
The most successful candidates for the recognition of methylated lysines are macrocycles. They have clearly defined cavities featuring high grades of preorganization paired with controlled flexibility. Macrocycles are deployable as platforms for multivalent systems and functional groups can be introduced at specific sites allowing fine-tuned hosts. The field of methyllysine

recognition is still in its infancies and the author of this review can only speculate about future developments. It can be assumed that prospective receptors will be designed either to be specific or generic in their recognition behavior – in both cases with high Kme_0/Kme_x selectivity. As proposed by the author and co-workers macrocycles with intrinsic fluorescence could be one way forward.[59] Furthermore, bioinformatics combined with molecular modelling and crystallographic approaches will help to design new generations of 'artificial antibodies' for the recognition of the methyllysine motif. A challenge for the future is clearly the better comparability between the different synthetic hosts. The reviewed literature uses six different analytical techniques to determine the dissociations constant of respective complexes, which makes authentic structure-activity relationships rather difficult. It would also be helpful to screen not only lysine and trimethyllysine, but also mono- and dimethyllysines.
As shown in Scheme 2 synthetic receptors for trimethyllysines offer a broad range of applications. Significant progress has been made for the analysis and purification of methyllysines. In several cases the target compounds have been employed in assays using demethylases and/or methyltransferases. Moreover, the biological activity of some methyllysine receptors is a promising start for new therapeutic agents. In no way synthetic receptors will be able to fully replace antibodies, though the reviewed literature delivers promising points of contacts for amendatory applications.
The use of synthetic methyllysine receptor as therapeutic agents is only possible with an acceptable toxicity. Single injected doses of tetrasulfonatocalixarene **1** (equivalent to 2-5 g in humans) showed no acute toxicity[115] and no toxicity towards various tumor cell lines has been observed.[116] It develops no haemolytic toxicity observed for concentrations up to 5 mM[117] and is not activating neutrophils, hence, does not induce an immune response[118]. Also cucurbiturils have demonstrated a rather low toxicity.[119] Despite they are able to cross the cell membrane of mouse embryo cells,[120] cucurbiturils lack of cytotoxicity in mammalian cells (up to 1 mM).[121] Intravenously administered cucuribt[7]uril (**39**) has demonstrated no toxicity at doses up to 200 mg/kg,[122] though has a measureable cardiotoxicity at concentrations > 500 µM.[123] Other synthetic receptors for methyllysine may have much higher toxicity levels, whose determination should be a point of contact for further studies.

Most of the synthetic receptors for methyllysines discussed here also recognize a broad range of other alkylammonium ions. In some cases the latter are complexed with even higher affinities than methyllysines. As most applications use the recognition of methylated lysines in biological media naturally occurring alkylammonium ions may disturb the desired recognition event. An example are cucurbiturils, that – when transferred into the cell – would not only recognize methyllysines, but may also interact with spermine and spermidine with the resulting complexes affecting DNA-modifying enzymes.[124] More studies are necessary to rule out cross-interactions, especially as hosts **1** and **14-16** have been described to disrupt the interaction of histone peptides with their natural protein binder, hence, already demonstrated their influence on DNA activity.

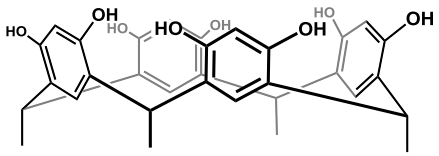
Another broad topic for future research are off-target effects of synthetic receptors for methylated lysines. A possible example is tetrasulfonatocalixarene **1** and its ability to hydrolyze ATP (Scheme 21).[125] After application of **1** on cells or on organisms – aiming for the discrimination of methyllysines – the artificially decreased ATP level could lead to artefacts as (ATP)-dependent chromatin remodeling enzymes[126] may be influenced. Would other sulfonatocalixarenes be active as well? Other untried interactions are conceivable as **1** is also a high affinity blocker of chloride channels[127] and volume-regulated anion channels.[128] Furthermore, it is known for its antithrombotic and anticoagulant properties.[129] By way of interest, both calixarene **1** and cucurbituril **39** inhibit amyloid fibrillation by multipoint

hydrophobic interactions,[130] which may lead to competing host/guest interactions in respective systems.



Scheme 21. Proposed complex of **1** with ATP during the catalytic hydrolysis of the guest.[125]

The blood-brain barrier prevents the uptake of large hydrophilic and highly anionic artificial hosts as shown for tetrasulfonatocalixarene **1**.[115] Nevertheless, in the long term the use of synthetic receptors for methyllysines in the brain may be feasible. Several methyllysine receptors (e.g. **23**, **39**, etc.) have also a high affinities to acetylcholine and the complexation of the neurotransmitter could lead to artefacts when screening for methyllysines. Possible are also off-target effects: receptor **43** (Scheme 22) – closely related to tetracyano receptor **23** – was found to inhibit the hydrolysis of acetylcholine,[131] hence, may influence synapse activities. Future research could help to establish if guest hydrolysis is a general characteristic of synthetic trimethylammonium receptors.



43

Scheme 22. Resorcinarene **43** is an inhibitor of acetylcholine hydrolysis.

10. Acknowledgments

Financial support from the School of Pharmacy, University of Lincoln, is gratefully acknowledged. TG wants to thank Manfred Jung (Institute of Pharmaceutical Sciences, University of Freiburg/Breisgau) for introducing him to the fascinating field of epigenetics.

11. References

[1] T. Jenuwein, C. D. Allis, *Science* **2001**, *293*, 1074-1080.

[2] M. Luo, *Chem. Rev.* **2018**, Ahead of Print. **DOI:** 10.1021/acs.chemrev.8b00008

[3] a) W. Sippl, M. Jung (Eds.), *Epigenetic Targets in Drug Discovery*, Wiley-VCH, Weinheim, **2009**; b) S. Krishnan, S. Horowitz, R. C. Trievel, *ChemBioChem* **2011**, *12*, 254-263; c) R. L. Hancock, M. I. Abboud, T. J. Smart, E. Flashman, A. Kawamura, C. J. Schofield, R. J. Hopkinson, *ChemBioChem* **2018**, *19*, 917-921.

[4] a) C. B. Yoo, P. A. Jones, *Nat. Rev. Drug Discovery* **2006**, *5*, 37-50; b) P. Chi, C. D. Allis, G. G. Wang, *Nat. Rev. Cancer* **2010**, *10*, 457-469; c) M. A. Dawson, T. Kouzarides, *Cell* **2012**, *150*, 12-27.

[5] L. M. Soares, P. C. He, Y. Chun, H. Suh, T. Kim, S. Buratowski, *Mol Cell.* **2017**, *68*, 773-785.

[6] J.-A. Park, A.-J. Kim, Y. Kang, Y.-J. Jung, H. K. Kim, K.-C. Kim, *Mol Cell.* **2011**, *31*, 343-349.

[7] E. Martinez-Garcia, J. D. Licht, *Nat. Gen.* **2010**, *42*, 100-101.

[8] S. Jørgensen, G. Schotta, C. S. Sørensen, *Nucleic Acids Res.* **2013**, *41*, 2797-2806.

[9] Y. Yokoyama, A. Matsumoto, M. Hieda, Y. Shinchi, E. Ogihara, M. Hamada, Y. Nishioka, H. Kimura, K. Yoshidome, M. Tsujimoto, N. Matsuura, *Breast Cancer Research* **2014**, *16*, R66

[10] N. Tsankova, W. Renthal, A. Kumar, E. J. Nestler, *Nat. Rev.* **2007**, *8*, 355-367.

[11] B. D. Smith (Ed.), *Synthetic Receptors for Biomolecules*, Royal Society of Chemistry, **2015**.

[12] S. M. Fuchs, B. D. Strahl, *Epigenomics* **2011**, *3*, 247-249.

[13] a) I. Bock, A. Dhayalan, S. Kudithipudi, O. Brandt, P. Rathert, A. Jeltsch, *Epigenetics* **2011**, *6*, 256-263; b) S. B. Rothbart, B. M. Dickson, J. R. Raab, A. T. Grzybowski, K. Krajewski, A. H. Guo, E. K. Shanle, S. Z. Josefowicz, S. M. Fuchs, C. D. Allis, T. R. Magnuson, A. J. Ruthenburg, B. D. Strahl, *Mol. Cell* **2015**, *59*, 502-511.

[14] Y. Zhao, D. Lou, J. Burkett, H. Kohler, *J. Immunol. Methods* **2001**, *254*, 137-145.

[15] J. R. Couchman, *J. Histochem. Cytochem.* **2009**, *57*, 7-8.

[16] A. Schonbrunn, *Mol. Endocrinol.* **2014**, *28*, 1403-1407.

[17] a) M. Bremang, A. Cuomo, A. M. Agresta, M. Stugiewicz, V. Spadotto, T. Bonaldi, *Mol. BioSyst.* **2013**, *9*, 2231-2247; b) A. Guo, H. Gu, J. Zhou, D. Mulhern, Y. Wang, K. A. Lee, V. Yang, M. Aguiar, J. Kornhauser, X. Jia, J. Ren, S. A. Beausoleil, J. C. Silva, V. Vemulapalli, M. T. Bedford, M. J. Comb, *Mol. Cell. Proteomics* **2014**, *13*, 372-387.

[18] a) B. A. Garcia, S. Mollah, B. M. Ueberheide, S. A. Busby, T. L. Muratore, J. Shabanowitz, D. F. Hunt, *Nature Protocols* **2007**, *2*, 933-938; b) Z. Wu, Z. Cheng, M. Sun, X. Wan, P. Liu, T. He, M. Tan, Y. Zhao, *Mol. Cell. Proteomics* **2015**, *14*, 329-339.

[19] E. S. Witze, W. M. Old, K. A. Resing, N. G. Ahn, *Nat. Methods* **2007**, *4*, 798-806.

[20] a) K. D. Daze, F. Hof, *Acc. Chem. Res.* **2013**, *46*, 937-945; b) A. Shaurya, K. I. Dubicki, F. Hof, *Supramol. Chem*. **2014**, *26*, 583-590; c) F. Hof, *Chem. Commun.* **2016**, *52*, 10093-10108.

[21] A. Späth, B. König, *Beilstein J. Org. Chem*. **2010**, *6*, doi: 10.3762/bjoc.3.32.

[22] E. A. Kataev, C. Müller, *Tetrahedron* **2014**, *70*, 137-167.

[23] a) T. H. Rehm, C. Schmuck, *Chem. Soc. Rev*. **2010**, *39*, 3597-3611; b) E. Persch, O. Dumele, F. Diederich, *Angew. Chem. Int. Ed.* **2015**, *54*, 3290-3327.

[24] S. D. Taverna, H. Li, A. J. Ruthenburg, C. D. Allis, D. J. Patel, *Nat. Struct. Biol.* **2007**, *14*, 1025-1040.

[25] C. Rapp, E. Goldberger, N. Tishbi, R. Kirshenbaum, *Proteins* **2014**, *82*, 1494-1502.

[26] P. Thordarson, *Chem. Soc. Rev*. **2011**, *40*, 1305-1323.

[27] P. Kuzmič, M. L. Moss, J. L. Kofron, D. H. Rich, *Anal. Biochem*. **1992**, 65-69.

[28] T. D. Pollard, *Mol. Biol. Cell* **2010**, *21*, 4061-4067.

[29] Y. Tang, X. Zeng, J. Liang, *J. Chem. Educ.* **2010**, *87*, 742-746.

[30] R. Gleiter, H. Hopf (Eds.), *Modern Cyclophane Chemistry*, Wiley: Weinheim, **2004**.

[31] C. D. Gutsche, *Calixarenes*, The Royal Society of Chemistry, **2008**.

[32] a) M. Gruner, C. Fischer, T. Gruber, E. Weber, *Supramol. Chem*. **2010**, *22*, 256-266; b) T. Gruber, M. Gruner, C. Fischer, W. Seichter, E. Weber, P. Bombicz, *New J. Chem*. **2010**, *34*, 250-259; c) C. Fischer, T. Gruber, D. Eißmann, W. Seichter, E. Weber, *Cryst. Growth Des.* **2011**, *11*, 1989-1994; d) T. Gruber, F. Eißmann, M. Gruner, L. G. Reuter, W. Seichter, E. Weber, *CrystEngComm*. **2014**, *16*, 3730-3736.

[33] a) R. V. Rodik, V. I. Boyko, V. I. Kalchenko, *Curr. Med. Chem*. **2009**, *16*, 1630-1655; b) M. Giuliani, I. Morbioli, F. Sansone, A. Casnati, *Chem Commun.* **2015**, *51*, 14140-141159; c) M. M. Naseer, M. Ahmed, S. Hameed, *Chem. Biol. Drug Des.* **2017**, *89*, 243-256.

[34] a) D.-S. Guo, Y. Liu, *Acc. Chem. Res.* **2014**, *47*, 1925-1934; b) C. Hoskins, A. D. M. Curtis, Nanomed. Res. 2015, 2, 00028.

[35] a) J. L. Atwood, L. J. Barbour, P. C. Junk, G. W. Orr, *Supramol. Chem*. **1995**, *5*, 105-108; b) L. Memmi, A. N. Lazar, A. Brioude, V. Ball, A. W. Coleman, *Chem. Commun.* **2001**, 2474-2475; c) H.-J. Buschmann, L. Mutihac, E. Schollmeyer, *J. Incl. Phenom. Macrocyclic Chem.* **2003**, *46*, 133-137; d) F. Perret, A. N. Lazar, A. W. Coleman, *Chem. Commun.* **2006**, 2425-2438; e) G. Arena, A. Casnati, A. Contino, A. Magrì, F. Sansone, D. Sciotto, R. Ungaro, *Org. Biomol. Chem*. **2006**, *4*, 243-249.

[36] C. Bonaccorso, S. Gentile, F. G. Gulino, D. Sciotto, *Lett. Org. Chem*. **2009**, *6*, 598-603.

[37] E. da Silva, F. Nouar, A. W. Coleman, B. Rather, M. J. Zaworotko, *J. Incl. Phenom. Macrocyclic Chem*. **2003**, *46*, 161-166.

[38] N. Kon, N. Iki, S. Miyano, *Org. Biomol. Chem*. **2003**, *1*, 751-755.

[39] O. Danylyuk, K. Suwinska, *Chem. Commun*. **2009**, 5799-5813.

[40] C. S. Beshara, C. E. Jones, K. D. Daze, B. J. Lilgert, F. Hof, *ChemBioChem* **2010**, *11*, 63-66.

[41] M. Florea, S. Kudithipudi, A. Rei, M. J. Gonzalez-Alvarez, A. Jeltsch, W. M. Nau, *Chem. Eur. J*. **2012**, *18*, 3521-3528.

[42] K. D. Daze, T. Pinter, C. S. Beshara, A. Ibraheem, S. A. Minaker, M. C. Ma, R. J. M. Courtemanche, R. E. Campbell, F. Hof, *Chem. Sci*. **2012**, *3*, 2695-2699.

[43] K. D. Daze, C. E. Jones, B. J. Lilgert, C. S. Beshara, F. Hof, *Can. J. Chem*. **2013**, 1072-1076.

[44] H. F. Allen, K. D. Daze, T. Shimbo, A. Lai, C. A. Musselman, J. K. Sims, P. A. Wade, F. Hof, T. G. Kutateladze, *Biochem. J*. **2014**, *459*, 505-512.

[45] M. Ali, K. D. Daze, D. E. Strongin, S. B. Rothbart, H. Rincon-Arano, H. F. Allen, J. Li, B. D. Strahl, F. Hof, T. G. Kutateladze, *J. Biol. Chem*. **2015**, **290**, 22919-22930.

[46] a) R. E. McGovern, A. A. McCarthy, P. B. Crowley, *Chem. Commun*. **2014**, *50*, 10412-10415; b) R. E. McGovern, B. D. Snarr, J. A. Lyons, J. McFarlane, A. L. Whiting, I. Paci, F. Hof, P. B. Crowley, *Chem. Sci*. **2015**, *6*, 442-449.

[47] R. E. McGovern, H. Fernandes, A. R. Khan, N. P. Power, P. B. Crowley, *Nat. Chem*. **2012**, *4*, 527-533.

[48] M. Selkti, A. W. Coleman, I. Nicolis, N. Douteau-Guevel, F. Villain, A. Tomas, C. De Rango, *Chem. Commun*. **2000**, 161-162.

[49] K. D. Daze, M. C. Ma, F. Pineux, F. Hof, *Org. Lett*. **2012**, *14*, 1512-1515.

[50] S. Tabet, S.F. Douglas, K. D. Daze, G. A. E. Garnett, K. J. H. Allen, E. M. M. Abrioux, T. T. H. Quon, J. E. Wulff, F. Hof, *Bioorg. Med. Chem*. **2013**, *21*, 7004-7010.

[51] A. M. Doolan, M. L. Rennie, P. B. Crowley, *Chem. Eur. J*. **2018**, *24*, 984-991.

[52] Y. Kimura, N. Saito, K. Hanada, J. Liu, T. Okabe, S. A. Kawashima, K. Yamatsugu, M. Kanai, *ChemBioChem* **2015**, *16*, 2599-2604.

[53] J. E. Beaver, M. L. Waters, *ACS Chem Biol*. **2016**, *11*, 643-653.

[54] C. A. Musselman, J. Ramírez, J. K. Sims, R. E. Mansfield, S. S. Oliver, J. M. Denu, J. P. Mackay, P. A. Wade, J. Hagman, T. G. Kutateladze, *PNAS* **2012**, *109*, 787-792.

[55] G. A. E. Garnett, M. J. Starke, A. Shaurya, J. Li, F. Hof, *Anal. Chem*. **2016**, *88*, 3697-3703.

[56] F. Hof, G. A. E. Garnett, U.S. Pat. 20170052154 A1, **2017**.

[57] F. Hof, S. Minaker, K. Daze, S. Tabet, M. Ma, PCT Int. Appl. WO 2013091074 A1 20130627, **2013**.

[58] M. A. Beatty, J. Borges-González, N. J. Sinclair, A. T. Pye, F. Hof, *J. Am. Chem. Soc.* **2018**, *140*, 3500-3504.

[59] M. Bürger, F. Katzsch, E. Brendler, T. Gruber, *J. Org. Chem.* **2015**, *80*, 4882-4892.

[60] I. N. Gober, M. L. Waters, *J. Am. Chem. Soc.* **2016**, *138*, 9452-9459.

[61] R. Lalor, H. Baillie-Johnson, C. Redshaw, S. E. Matthews, A. Mueller, *J. Am. Chem. Soc.*, **2008**, *130*, 2892-2893.

[62] T. Hanauer, R. J. Hopkinson, K. Patel, Y. Li, D. Correddu, A. Kawamura, V. Sarojini, I. K. H. Leung, T. Gruber, *Org. Biomol. Chem.* **2017**, *15*, 1100-1105.

[63] M. L. Rennie, A. M. Doolan, C. L. Raston, P. B. Crowley, *Angew. Chem. Int. Ed.* **2017**, *56*, 5517-5521.

[64] W.-H. Chen, M. Nishikawa, S.-D. Tan, M. Yamamura, A. Satake, Y. Kobuke, *Chem. Commun.* **2004**, 872-873.

[65] W.-H. Chen, Y. Wei, S.-D. Tan, B. Wang, Z.-L. Xu, *Supramol. Chem.* **2005**, *17*, 469-473.

[66] H. Peacock, C. C. Thinnes, A. Kawamura, A. D. Hamilton, *Supramol. Chem.* **2016**, *28*, 575-581.

[67] M. Dionisio, G. Oliviero, D. Menozzi, S. Federici, R. M. Yebeutchou, F. P. Schmidtchen, E. Dalcanale, P. Bergese, *J. Am. Chem. Soc.* **2012**, *134*, 2392-2398.

[68] R. Pinalli, G. Brancatelli, A. Pedrini, D. Menozzi, D. Hernández, P. Ballester, S. Geremia, E. Dalcanale, *J. Am. Chem. Soc.* **2016**, *138*, 8569-8580.

[69] I. Alessandri, E. Biavardi, A. Gianoncelli, P. Bergese, E. Dalcanale, *ACS Appl. Mater. Interfaces* **2016**, *8*, 14944-14951.

[70] N. Bontempi, E. Biavardi, D. Bordiga, G. Candiani, I. Alessandri, P. Bergese, E. Dalcanale, *Nanoscale* **2017**, *9*, 8639-8646.

[71] Y. Liu, L. Perez, A. D. Gill, M. Mettry, L. Li, Y. Wang, R. J. Hooley, W. Zhong, *J. Am. Chem. Soc.* **2017**, *139*, 10964-10967.

[72] a) E. Menozzi, H. Onagi, A. L. Rheingold, J. Rebek Jr., *Eur. J. Org. Chem.* **2005**, *17*, 3633-3636; b) Y. Kubo, K. Tsuruzoe, S. Okuyama, R. Nishiyabu, T. Fujihara, *Chem. Commun.* **2010**, *46*, 3604-3606.

[73] a) M. P. Schramm, R. J. Hooley, J. Rebek Jr., *J. Am. Chem. Soc.* **2007**, *129*, 9773-9779; b) B. Soberats, E. Sanna, G. Martorell, C. Rotger, A. Costa, *Org. Lett*. **2014**, *16*, 840-843.

[74] C. Li, J. Ma, L. Zhao, Y. Zhang, Y. Yu, X. Shu, J. Li, X. Jia, *Chem. Commun.* **2013**, *49*, 1924-1926.

[75] D. A. Dougherty, D. A. Stauffer, *Science* **1990**, *250*, 1558-1560.

[76] P. C. Kearney, L. S. Mizoue, R. A. Kumpf, J. E. Forman, A. McCurdy, D. A. Dougherty, *J. Am. Chem. Soc.* **1993**, *115*, 9907-9919.

[77] a) S. M. Ngola, D. A. Dougherty, *J. Org. Chem*. **1998**, *63*, 4566-4567; b) S. M. Ngola, P. C. Kearney, S. Mecozzi, K. Russel, D. A. Dougherty, *J. Am. Chem. Soc.* **1999**, *121*, 1192-1201.

[78] P. T. Corbett, J. K. Sanders, S. Otto, *Chem. Eur. J.* **2008**, *14*, 2153-2166.

[79] W. Zhang, Y. Jin (Eds.), *Dynamic Covalent Chemistry: Principles, Reactions, and Applications*, John Wiley & Sons: Chichester, **2018**.

[80] R. M. Hughes, K. R. Wiggins, S. Khorasanizadeh, M. L. Waters, *PNAS* **2007**, *104*, 11184-11188.

[81] L. A. Ingerman, M. E. Cuellar, M. L. Waters, *Chem. Commun.* **2010**, *46*, 1839-1841.

[82] N. K. Pinkin, M. L. Waters, *Org. Biomol. Chem*. **2014**, *12*, 7059-7067.

[83] J. E. Beaver, B. C. Peacor, J. V. Bain, L. I. James, M. L. Waters, *Org. Biomol. Chem*. **2015**, *13*, 3220-3226.

[84] I. N. Gober, M. L. Waters, *Org. Biomol. Chem*. **2017**, *15*, 7789-7795.

[85] N. K. Pinkin, A. N. Power, M. L. Waters, *Org. Biomol. Chem*. **2015**, *13*, 10939-10945.

[86] B. C. Peacor, C. M. Ramsay, M. L. Waters, *Chem. Sci.* **2017**, *8*, 1422-1428.

[87] N. K. Pinkin, I. Liu, J. D. Abron, M. L. Waters, *Chem. Eur. J.* **2015**, *21*, 17981-17986.

[88] S. Dinesh, K. Jayshree, K. Khedkar, P. Min, K. Kimoon, *Chem. Soc. Rev*. **2015**, *44*, 8747-8761.

[89] a) H. Zhang, M. Grabenauer, M. T. Bowers, D. V. Dearden, *J. Phys. Chem. A* **2009,** *113*, 1508-1517; b) J. W. Lee, H. H. L. Lee, Y. H. Ko, K. Kim, H. I. Kim, *J. Phys. Chem. B* **2015**, **119**, 4628-4636; c) E. Kovalenko, M. Vilaseca, M. Díaz-Lobo, A. N. Masliy, C. Vicent, V. P. Fedin, *J. Am. Soc. Mass Spectrom*. **2016**, *27*, 265-276.

[90] A. D. St-Jacques, I. W. Wyman, D. H. Macartney, *Chem. Common.* **2008**, 4936-4938.

[91] I. W. Wyman, D. H. Macartney, *Org. Biomol. Chem*. **2010**, 8, 253-260.

[92] M. A. Gamal-Eldin, D. H. Macartney, *Org. Biomol. Chem*. **2013**, *11*, 488-495.

[93] P. V. Peña, F. Davrazou, X. Shi, W. L. Walter, V. V. Verkhusha, O. Gozani, R. Zhao, T. G. Kutateladze, *Nature* **2006**, *442*, 100-103.

[94] S. Iwase, B. Xiang, S. Ghosh, T. Ren, P. W. Lewis, J. C. Cochrane, C. D. Allis, D. J. Picketts, D. J. Patel, H. Li, Y. Shi, *Nat. Struct. Mol. Biol.* **2011**, *18*, 769-776.

[95] S. Eustermann, J.-C. Yang, M. J. Law, R. Amos, L. M. Chapman, C. Jelinska, D. Garrick, D. Clynes, R. J. Gibbons, D. Rhodes, D. R. Higgs, D. Neuhaus, *Nat. Struct. Mol. Biol.* **2011**, *18*, 777-782.

[96] F. Biedermann, W. M. Nau, H.-J. Schneider, *Angew. Chem. Int. Ed.* **2014**, *53*, 11158-11171.

[97] F. Guagnini, P. M. Antonik, M. L. Rennie, P. O'Byrne, A. R. Khan, R. Pinalli, E. Dalcanale, P. B. Crowley, *Angew. Chem. Int. Ed.* **2018**, *57*, 7126-7130.

[98] J. M. Chinai, A. B. Taylor, L. M. Ryno, N. D. Hargreaves, C. A. Morris, P. J. Hart, A. R. Urbach, *J. Am. Chem. Soc.* **2011**, *133*, 8810-8813.

[99] J. Lee, L. Perez, Y. Liu, H. Wang, R. J. Hooley, W. Zhong, *Anal. Chem.* **2018**, *90*, 1881-1888.

[100] a) C. Schmuck, *Chem. Commun.* **1999**, 843-844; b) C. Schmuck, V. Bickert, *Org. Lett.* **2003**, *5*, 4579-4581; c) C. Schmuck, S. Graupner, *Tetrahedron Lett.* **2005**, *46*, 1295-1298; d) C. Schmuck, L. Geiger, *J. Am. Chem. Soc.* **2005**, *127*, 10486-10487; e) C. Schmuck, U. Machon, *Chem. Eur. J.* **2005**, *11*, 1109-1118.

[101] a) S. Dutt, C. Wilch, T. Gersthagen, P. Talbiersky, K. Bravo-Rodriguez, M. Hanni, E. Sánchez-García, C. Ochsenfeld, F.-G. Klärner, T. Schrader, *J. Org. Chem.* **2013**, *78*, 6721-6734; b) F.-G. Klärner, T. Schrader, *Acc. Chem. Res.* **2013**, *46*, 967-978; c) S. Acharya, B. M. Safaie, P. Wongkongkathep, M. I. Ivanova, A. Attar, F.-G. Klärner, T. Schrader, J. A. Loo, G. Bitan, L. J. Lapidus, *J. Biol. Chem.* **2014**, *289*, 10727-10737; d) M. Mallon, S. Dutt, T. Schrader, P. B. Crowley, *ChemBioChem* **2016**, *17*, 774-783; e) T. Schrader, G. Bitan, F.-G. Klärner, *Chem. Commun.* **2016**, *52*, 11318-11334.

[102] a) A. Vacca, C. Nativi, M. Cacciarini, R. Pergoli, S. Roelens, *J. Am. Chem. Soc.* **2004**, *126*, 16456-16465; b) M. Cacciarini, E. Cordiano, C. Nativi, S. Roelens, *J. Org. Chem.* **2007**, *72*, 3933-3936; c) M. Mazik, M. Kuschel, *Chem. Eur. J.* **2008**, *14*, 2405-2419; d) J.-R. Rosien, W. Seichter, M. Mazik, *Org. Biomol. Chem.* **2013**, *11*, 6569-6579.

[103] A. L. Whiting, N. M. Neufeld, F. Hof, *Tetrahedron Lett.* **2009**, *50*, 7035-7037.

[104] A. L. Whiting, F. Hof, *Org. Biomol. Chem.* **2012**, *10*, 6885-6892

[105] S.-D. Tan, W.-H. Chen, A. Satake, B. Wang, Z.-L. Xu, Y. Kobuke, *Org. Biomol. Chem* **2004**, *2*, 2719-2721.

[106] L. I. James, J. E. Beaver, N. W. Rice, M. L. Waters, *J. Am. Chem. Soc.* **2013**, *135*, 6450-6455.

[107] M. Karimi, M. T. Ignasiak, B. Chan, A. K. Croft, L. Radom, C. H. Schiesser, D. I. Pattison, M. J. Davies, *Scientific Reports* 6:38572 | DOI: 10.1038/srep38572.

[108] J. M. Leiper, P. Vallance, *Eur. J. Clin. Pharmacol.* **2006**, *62*, 33-38.

[109] M. G. Zeece, Q. Chu, S. J. Jones, T. L. Woods, W. J. Reville, *J. Cap. Electrophor.* **1996**, *3*, 55-59.

[110] W.-S. Aik, J. S. Scotti, H. Choi, L. Gong, M. Demetriades, C. J. Schofield, M. A. McDonough, *Nucleic Acids Res.* **2014**, *42*, 4741-4754.

[111] D. Masseroni, E. Biavardi, D. Genovese, E. Rampazzo, L. Prodi, E. Dalcanale, *Chem. Commun.* **2015**, *51*, 12799-12802.

[112] a) B. Botta, A. Tafi, F. Caporuscio, M. Botta, L. Nevola, I. D'Acquarica, C. Frachetti, M. Speranza, *Chem. Eur. J.* **2008**, *14*, 3585-3595; b) E. Biavardi, S. Federici, C. Tudisco, D. Menozzi, C. Massera, A. Sottini, G. G. Condorelli, P. Bergese, E. Dalcanale, *Angew. Chem. Int. Ed.* **2014**, *53*, 9183-9188; c) Y. Jang, M. Jang, H. Kim, S. J. Lee, E. Jin, J. Y. Koo, I.-C. Hwang, Y. Kim, Y. H. Ko, I. Hwang, J. H. Oh, K. Kim, *Chem* **2017**, *3*, 641-651.

[113] S. Ganapati, S. D. Grabitz, S. Murkli, F. Scheffenbichler, M. I. Rudolph, P. Y. Zavalij, M. Eikermann, L. Isaacs, *ChemBioChem* **2017**, *18*, 1583-1588.

[114] C. Schneider, A. Bierwisch, M. Koller, F. Worek, S. Kubik, *Angew. Chem., Int. Ed.* **2016**, *55*, 12668-12672.

[115] A. W. Coleman, S. Jebors, S. Cecillon, P. Perret, D. Garin, D. Marti-Battle, M. Moulin, *New J. Chem.* **2008**, *32*, 780-782.

[116] A. W. Coleman, A. N. Lazar, L. G. Bagetto, S. Magnard, M. H. Michaud, *Fr. Pat.*, FR 2899900, **2007**.

[117] E. Da Silva, P. Shahgaldian, A. W. Coleman, *Int. J. Pharm.* **2004**, *273*, 57-62.

[118] M. H. Paclet, C. F. Rousseau, C. Yannick, F. Morel, A. W. Coleman, *J. Incl. Phenom. Macrocycl. Chem.* **2006**, *55*, 353-357.

[119] S. Walker, R. Oun, F. J. McInnes, N. J. Wheate, *Israel J. Chem.* **2011**, *51*, 616-624.

[120] P. Montes-Navajas, M. Gonzalez-Bejar, J. C. Scaiano, H. Garcia, *Photochem. Photobiol. Sci.* **2009**, *8*, 1743-1747.

[121] N. J. Wheate, C. Limantoro, *Supramol. Chem.* **2016**, *28*, 849-856.

[122] V. D. Uzunova, C. Cullinane, K. Brix, W. M. Nau, A. I. Day, *Org. Biomol. Chem.* **2010**, *8*, 2037-2042.

[123] H. Chen, J. Y. W. Chan, X. Yang, I. W. Wyman, D. Bardelang, D. H. Macartney, S. M. Y. Lee, R. Wang, *RSC Adv.* **2015**, *5*, 30067-30074.

[124] C. Parente Carvalho, A. Norouzy, V. Ribeiro, W. M. Nau, U. Pischel, *Org. Biomol. Chem.* **2015**, *13*, 2866-2869.

[125] Y. Tian-Ming, Y. Zhi-Feng, W. Li, G. Jin-Ying, Y. Si-De, S. Xian-Fa, *Spectrochim. Acta A.* **2002**, *58*, 3033-3038.

[126] M. Vignali, A. H. Hassan, K. E. Neely, J. L. Workman, *Mol. Cell. Biol.* **2000**, *20*, 1899-1910

[127] A. K. Singh, C. J. Venglarik, R. J. Bridges, *Kidney Int.*, 1**995**, 48, 985-993.

[128] G. Droogmans, J. Prenen, J. Eggermont, T. Voets, V. Nilius, *Am. J. Physiol.* **1998**, *275*, c646-c652.

[129] K. M. Hwang, Y. M. Qi, S.-Y. Liu, B. Thomas, C. Lee, S. Bruno, W. Choy, J. Chen, *Patent US5409959*, **1995**.

[130] a) H. H. Lee, T. S. Choi, S. J. C Lee, J. W. Lee, J. Park, Y. H. Ko, W. J. Kim, K. Kim, H. I. Kim, *Angew. Chem. Int. Ed.* **2014**, *53*, 7461-7465; b) Z. Wang, S. Tao, X. Dong, Y. Sun, *Chem. Asian J.* **2017**, *12*, 341-346.

[131] H.-J. Schneider, U. Schneider, *J. Org. Chem.* **1987**, *52*, 1613-1615.