A Journal of the Gesellschaft Deutscher Chemiker

Angewandte

GDCh

International Edition

Chemie

www.angewandte.org

Accepted Article

Title: Achieving High Affinity and Selectivity for Asymmetric Dimethylarginine by Putting a Lid on a Box

Authors: Alexandria Mullins, Nicholas Pinkin, Joshua Hardin, and Marcey L. Waters

To be cited as: *Angew. Chem. Int. Ed.* 10.1002/anie.201814645
Angew. Chem. 10.1002/ange.201814645

Link to VoR: http://dx.doi.org/10.1002/anie.201814645
http://dx.doi.org/10.1002/ange.201814645

WILEY-VCH

10.1002/anie.201814645

WILEY-VCH

COMMUNICATION

Achieving High Affinity and Selectivity for Asymmetric Dimethylarginine by Putting a Lid on a Box

Alexandria G. Mullins, Nicholas K. Pinkin, Joshua A. Hardin, Marcey L. Waters*[a]

Abstract: The methylation states of Lys and Arg represent a particularly challenging set of targets to distinguish selectively in water using synthetic receptors. To date, Kme3 is the only post translational modification (PTM) of the eight possible methylation states of Lys and Arg that can be recognized selectively. Here we report the first synthetic receptor capable of selectively recognizing Rme2a. This was achieved using a biased DCC library to generate a receptor mimicking the 5-sided box-like shape of Rme2 reader proteins, a feature that has been hypothesized to impart selectivity. Additionally, we have synthesized a thioether-linked analogue of the resulting receptor to provide a novel scaffold with maintained selectivity, but greater stability. This work introduces strategies that can be applied towards achieving selectivity based on subtle differences in hydrophilic guests in aqueous solutions.

Supramolecular chemistry is increasingly being applied to problems in chemical biology. However, an ongoing fundamental challenge that limits its application to biological problems is achieving selective molecular recognition of hydrophilic guests in water.[1–4] Methylated lysine (Kme) and arginine (Rme) (Fig. 1a) are protein post-translational modifications (PTMs) that represent a noteworthy example of difficult targets to distinguish in water. These PTMs are of significant interest due to their role in mediating gene expression and their dysregulation in numerous diseases[5,6]. Much work has gone into developing synthetic receptors that bind PTMs selectively in aqueous solution, with applications in sensing, tagging, enrichment, and enzymatic assays.[7–12] However, achieving selectivity for one methylation state of Lys or Arg has been a persisting challenge, as all eight possible methylation states can participate in electrostatic interactions, cation-π interactions, and H-bonding, with the exception of trimethyllysine (Kme3) which cannot H-bond. Indeed, the lower desolvation cost of Kme3 may be why it has been the easiest to bind with high affinity (< 1 μM) and selectivity (> 10-fold over Kme2).[13] Attempts at selectively binding lower methylation states of lysine[14] or dimethyl arginine (Rme2)[15] have been met with limited success (see Fig. 2b,c). Herein, we describe the first selective receptor for asymmetric dimethylarginine (Rme2a), which is associated with a number of diseases, including cancer, neurogenic, metabolic, and muscular disorders.[5] This receptor, **N_2G_2**, binds more tightly than all reported reader proteins for Rme2.[16] Both high affinity and selectivity were achieved by mimicking the 5-sided box-like shape of Rme2 reader proteins (Fig. 1b) through a unique receptor design with a two-flapped lid (Fig. 2d).

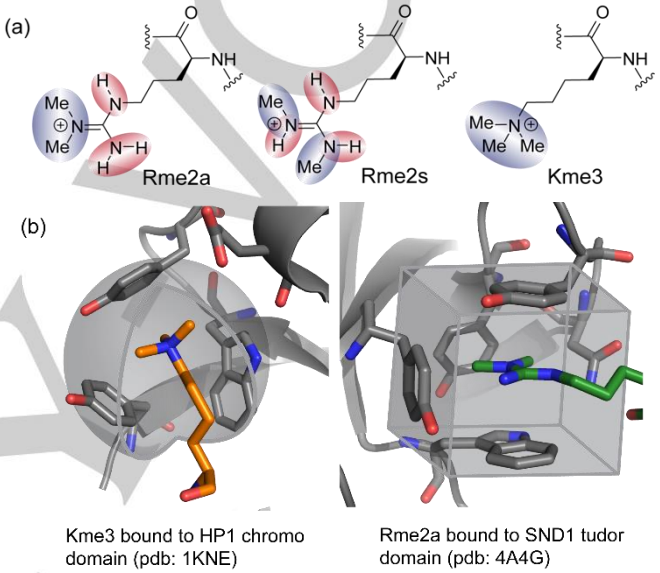


Figure 1. (a) Rme2a, Rme2s, and Kme3, highlighting regions capable or hydrogen-bonding (red) or cation-π interactions (blue). (b) A crystal structure of Kme3 (orange) and an NMR structure of Rme2a (green) in their respective binding pockets.

Previously our group has utilized dynamic combinatorial chemistry (DCC) [17] to develop novel receptors for methylated Lys and Arg using disulfide exchange, and identified a number of **A_2X** receptors, where **X** is a varied monomer imposing selectivity (Fig. 2a,c). Switching **X** from a phenyl to a naphthyl monomer resulted in the first synthetic receptor for Rme2a, **A_2D**.[15] While **A_2D** remains the tightest binding receptor for Rme2a,[18] it does not exhibit any selectivity over Kme3, binding both equally in the low micromolar range.[15] Inspection of the binding pockets of reported synthetic receptors for methylated Lys and Arg as compared to those of their respective reader proteins suggested a feature that may explain the lack of selectivity: **A_2D** has a spherical binding pocket, similar to Kme3 reader proteins, but reader proteins for Rme2 have a conserved 5-sided box with an open face for Rme2 to thread in (Fig. 1b, S1).[16] The Rme2 reader proteins provide excellent selectivity over Kme3, a feat that has been unachievable using synthetic receptors to date.[15,18] It has been proposed that the narrower cleft formed by the top and bottom aromatic faces of what has been described as a "rectangle cuboid" pocket provides selectivity over Kme3, although this has not been experimentally demonstrated.[19] Additionally, we envisioned that the sides and

[a] Department of Chemistry, CB 3290, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, United States
E-mail: mlwaters@email.unc.edu

Supporting information for this article is given via a link at the end of the document.

10.1002/anie.201814645

back of the box could further control selectivity for Rme2a over Rme2s by constraining the size of the binding pocket and providing additional CH-π interactions. This may disfavor Rme2s, which has H-bond donors on all three N atoms, resulting in more disperse hydrophilic regions (Fig 1a).

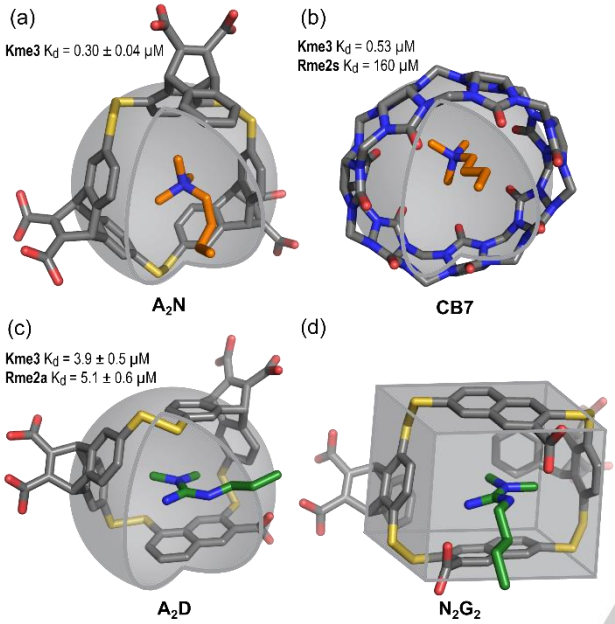


Figure 2. Gas-phase models of previously reported receptors (a) **A_2N**,[13] (b) **CB7**,[18] (c) **A_2D**,[15] and the proposed box-like receptor (d) **N_2G_2** shown bound to Kme3 (orange) or Rme2a (green). The reported affinities were measured with free amino acids for **CB7** and with histone peptides for **A_2N** and **A_2D**.

To evaluate these hypotheses, we screened libraries consisting of different combinations of three previously reported building blocks **E**, **G**, and **N** (Fig. 3a).[13,20] Compound **G** was included in all libraries, as it has been reported to form rectangular receptors with **E**,[20] and the extended width of **G** was expected to be necessary to accommodate methylated Arg. Monomers **E** and **N** both introduce a -2 charge and contain a para-substituted dithiobenzene, predicted to give the correct height to a receptor for Rme2 but not Kme3. However, **N** provides a second benzene ring that can act as a "lid", providing the fifth face of the receptor, as is observed in Rme2 reader proteins. Moreover, **N** has previously been shown to impart significant selectivity for binding an NMe over an NH group in a receptor for methylated Lys by providing two CH-π interactions and creating a deeper aromatic pocket.[13] Thus, three libraries consisting of **E+G**, **E+N+G**, and **N+G**, were investigated to determine which features, if any, are important for Rme2a selectivity. Amplification of a species in a dynamic combinatorial library (DCL) generally correlates with binding, although statistical preferences must be kept in mind.[17] These libraries were screened against tetrameric peptides, Ac-XGGY-NH_2, containing a variable residue X (=R, Rme2a, Rme2s, and Kme3), a Gly-Gly spacer, and a Tyr tag for concentration determination (Figure 3b-d). Additional libraries templated with the lower methylation states of Arg and Lys are reported in the SI and demonstrate no selective amplification.

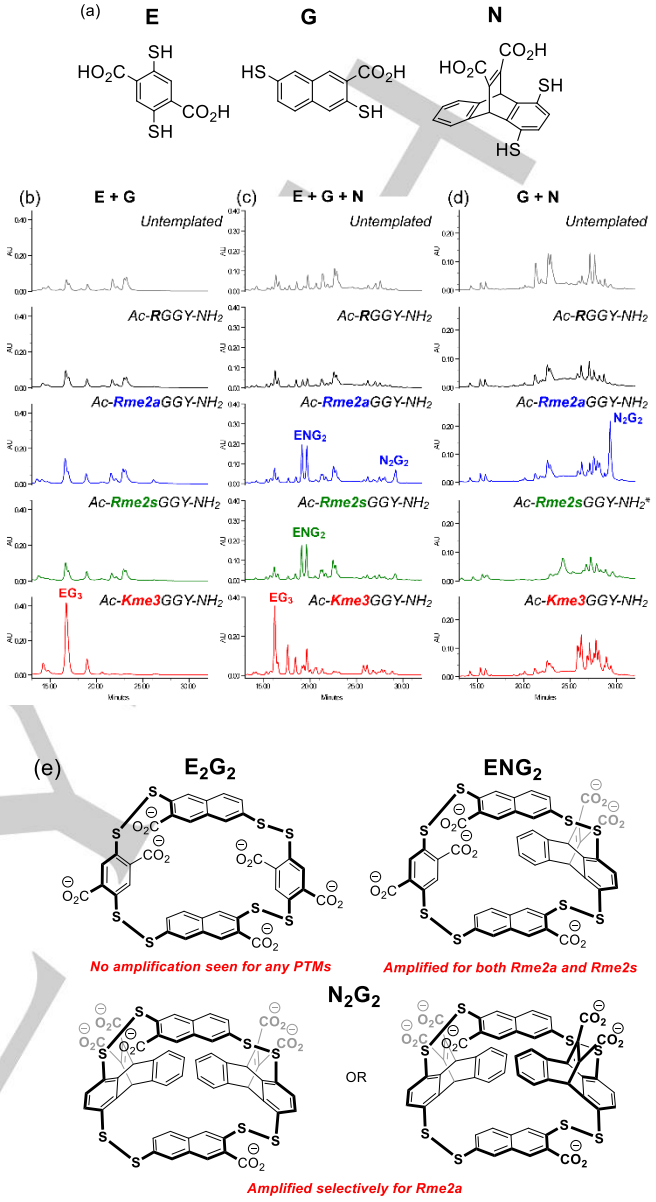


Figure 3. (a) Monomers **E**, **G**, and **N** (b-d) HPLC traces of a libraries at 254 nm after equilibrating in pH 8.5 50 mM sodium borate buffer; [peptide] = 1.35 mM; (b) [G] = 0.67 mM, [E] = 0.67 mM; (c) [E] = 0.34 mM, [G] = 0.67 mM, [N] = 0.34 mM; (d) [G] = 0.67 mM, [N] = 0.67 mM. (Shifted peaks due to analysis on a different day)*. (e) Proposed box-like receptors.

Interestingly, despite the fact that all three libraries have the potential to form the same rectangular scaffold, the amplified species varied significantly between PTMs and libraries. In the **E+G** library, the rectangular **E_2G_2** macrocycle was not amplified by any PTMs, suggesting that a box-like shape alone is not sufficient for Arg recognition (Figure 3b and e). Indeed, Kme3 resulted in the only amplified species, **EG_3,** which has a larger pocket than **E_2G_2**. Since **E_2G_2** is statistically preferred over **EG_3**, lack of **E_2G_2** amplification is consistent with preferential binding.

Introducing one equivalent of monomer **N** provided amplification of two isomers of **ENG_2** in the presence of both

10.1002/anie.201814645

Rme2a and Rme2s (Figure 3c and e). Thus, these results support selectivity for Rme2 over Kme3, presumably because of a narrow cleft that cannot accommodate Kme3 (see Fig. S10). However, **ENG$_2$** appears to lack selectivity between Rme2s and Rme2a. Monomer **E** introduces additional carboxylates on both edges of the pocket that can H-bond with the NH groups of both Rme2a and Rme2s, which may allow for binding both Rme2a and Rme2s.

In contrast to the **E+N+G** library, the **N+G** library resulted in a single amplified peak, corresponding to **N$_2$G$_2$**, in the presence of Rme2a only. Note that this peak was also weakly amplified in the **E+N+G** library with Rme2a but not Rme2s. While the other three isomers of **N$_2$G$_2$** were observed, none were amplified (see Fig. S14). Replacing monomer **E** with another unit of monomer **N** to give **N$_2$G$_2$** allows for CH-π interactions with Rme2a, regardless of whether both **N** units are parallel or antiparallel (Fig. 3e and S11). In contrast, Rme2s would have to thread an NH into the aromatic cage in order to bind, which would have a higher desolvation cost. Thus, the deeper aromatic pocket may give rise to the selectivity between Rme2a and Rme2s. As with **ENG$_2$**, we expect that the cleft between the two naphthalenes is too narrow for Kme3, with a width of 7.8-7.9 Å, which compares well to the distances seen in Rme2 reader proteins of 6.7-7.2 Å.[19] The lack of amplification of any rectangular species by Kme3 in any of the libraries further supports the hypothesis that a narrower cleft excludes Kme3 binding.

Because of its apparent selectivity, we isolated **N$_2$G$_2$** by RP-HPLC for additional characterization. We undertook tandem mass spectrometry experiments to determine the connectivity of monomers **N** and **G** in the macrocycle, specifically looking for whether our modeled alternating connectivity of monomers was accurate.[21] We reasoned that the disulfide bonds of an **NG**, **NN**, and **GG** species should all fragment and ionize equivalently and therefore we could determine their connectivity by the population of fragmented species observed. The lack of **NN** or **GG** species supported the modeled constitutional isomer of **N$_2$G$_2$** with alternating **N** and **G** monomers necessary for forming the predicted box (Fig. S18).

Isothermal calorimetry (ITC) was used to characterize the binding of **N$_2$G$_2$** to a short model peptide sequence representative of the histone H3 tail (H3R8me$_X$K9me$_Z$ = Ac-YGGQTARme$_x$Kme$_z$STG-NH$_2$, where X, Z = 0; X = 0, Z = 3; X = 1, Z = 0; or X = 2; Z = 0) (Table 1). Binding affinities indicate that **N$_2$G$_2$** is indeed selective, exhibiting >10-fold selectivity over both Rme2s and Kme3, consistent with amplification results. Additionally, the receptor is approximately 20-fold selective over Rme and unmethylated Arg. The binding constant for **N$_2$G$_2$** to H3R8me2a (K_d = 1.2±0.3 µM) is tighter than *all known* Rme2 reader proteins (5 to >1000 µM, see Table S1). and demonstrates 4-fold greater affinity than the previously reported receptor **A$_2$D** to the same peptide (K_d = 5.1±0.6 µM) as well as improved selectivity over Rme2s (7-fold for **A$_2$D**).[15] Thus, creating a narrower cleft increased selectivity over Kme3 more than 10-fold and improved the absolute binding affinity. Furthermore, inclusion of the aromatic "lids" appears to be required for selectivity over Rme2s.

We further explored whether we could use **N$_2$G$_2$** as a template to create a receptor that would be more stable in the reducing conditions of biological assays, by investigating a thioether analogue of **N$_2$G$_2$** (**thN$_2$G$_2$**). Thioether bonds maintain a similar dihedral angle as disulfides and would be stable in the presence of thiols such as glutathione, so we hypothesized an analogue of the DCC-identified receptor would maintain selectivity over Kme3 with a box-like geometry while providing a more stable linkage to access a larger breadth of future applications. The binding of thioether cyclophanes has not previously been explored as a class of synthetic receptors for biomolecules.

Table 1. Binding Data N_2G_2 with for Rme and Kme Peptides.

Peptide	Kd (µM) [a]	Selectivity
H3R8me2a	1.2±0.3	-
H3K9me3	13±1	11
H3R8me2s	14±2	11
H3R8me1	21±3	17
H3R8K9	23±2	19

[a] Conditions: 25° C, 10 mM Na_2HBO_3/NaH_2BO_3, pH 8.5. Average of two runs.

A variant of monomer **N** in which the thiols were replaced with bromomethyl groups was synthesized[22] to provide a site for nucleophilic substitution with the thiol of monomer **G** (Scheme S1). This synthetic route provided the four possible isomers of thioether-linked **N$_2$G$_2$** (**thN$_2$G$_2$**) in a statistical distribution. Preliminary binding studies revealed that the fourth isomer (**thN$_2$G$_2$-4**) demonstrated modestly tighter binding to Rme2a than the others. Therefore this isomer was used for continued studies. It is worth noting that while **thN$_2$G$_2$-4** and the isomer of **N$_2$G$_2$** that was amplified both displayed the longest retention times, we were unable to confirm that they are the same isomer (*vide infra*).

ITC binding studies of **thN$_2$G$_2$-4** revealed weaker binding affinity for Rme2a (K_d = 25±3 µM) compared to the disulfide-linked **N$_2$G$_2$**, although still tighter than any reported Rme2a reader proteins, which have K_d's ranging from 42 to >1000 µM (Table S1).[23] Selectivity over Kme3 and the other methylation states of arginine was maintained (Table S2). The weaker binding may be the result of several factors, including a smaller binding pocket due to shorter C-S bonds, which could make binding less favorable, or could allow for greater edge-face interactions between the naphthyl groups in the unbound state, which would need to be disrupted to bind Rme2a. Additionally, while the disulfides in **N$_2$G$_2$** are included as a result of disulfide exchange, they may contribute favorably to binding via dispersion interactions, as sulfur is quite polarizable.

To further probe the recognition of Rme2a, the change in the ^{1}H NMR spectrum of the tetrameric peptide Ac-Rme2aGGY-NH_2 in the bound state with both **N$_2$G$_2$** and **thN$_2$G$_2$-4** was investigated under saturating conditions. While **N$_2$G$_2$** appeared to aggregate under these conditions, **thN$_2$G$_2$-4** was well behaved. In the presence of **N$_2$G$_2$** and **thN$_2$G$_2$-4** the methyl groups on Rme2a were upfield shifted by 1.9 ppm and 1.7 ppm, respectively, indicating significant CH-π interactions, consistent with binding within the aromatic pocket. The alkyl chain of Rme2a, in contrast, is only minimally shifted, suggesting that binding is primarily at the guanidinium group. Furthermore, the peptide backbone and other

sidechains are not significantly shifted (Fig. S38). When compared to the reported NMR binding data for **A_2D** (Table S3), it is notable that the methyl groups experience almost a 2-fold greater Δδ for **thN_2G_2-4**, but less shifting for all of the protons on the carbon chain. Thus, **thN_2G_2-4** is a well behaved selective receptor for Rme2a with protein-like binding affinity and linkages that are stable to a reducing environment.

In conclusion, we have developed the first selective receptor for Rme2a, **N_2G_2**, which is tighter than any of the known Rme2 reader proteins.[23] Selectivity was achieved by using principles of molecular recognition gleaned from native reader proteins, including mimicking the cuboid shape which provides steric occlusion of Kme3 and formation of a deeper aromatic pocket through the incorporation of aromatic "lids", which disfavors binding of Rme2s. Selective molecular recognition of hydrophilic guests in water is an ongoing challenge, and this approach allowed us to select for a more hydrophilic guest, Rme2a, over a less hydrophilic guest, Kme3. In a general sense, this work advances the strategies that can be applied to achieve selectivity for subtle differences in hydrophilic guests in aqueous solution.

Additionally, we have demonstrated that thioether-linked receptors provide a novel scaffold capable of selectively recognizing PTMs. The potential to develop receptors selected from DCC studies into analogues that are stable to the reducing environment of a cell is particularly appealing. While some degree of binding affinity was lost relative to the disulfide-linked receptor, binding was nonetheless tighter than any reported native reader protein for Rme2a[23] and selectivity over Kme3 and unmethylated arginine was observed. Further functionalization of this receptor is planned to expand applications for characterizing methylated arginines.

Acknowledgements

This material is based upon work supported by the National Science Foundation (CHE-1608333).

Keywords: dimethylarginine • histone post-translational modifications • molecular recognition in water • epigenetics • synthetic receptors

(1) Uhlenheuer, D. A.; Petkau, K.; Brunsveld, L. *Chem. Soc. Rev.* **2010**, *39* (8), 2817.

(2) Davis, A. P.; Kubik, S.; Dalla Cort, A. *Org. Biomol. Chem.* **2015**, *13*, 2499–2500.

(3) Persch, E.; Dumele, O.; Diederich, F. *Angew. Chemie - Int. Ed.* **2015**, *54* (150), 3290–3327.

(4) Urbach, A. R.; Ramalingam, V. *Isr. J. Chem.* **2011**, *51* (5–6), 664–678.

(5) Blanc, R. S.; Richard, S. *Mol. Cell* **2017**, *65* (1), 8–24.

(6) Hyun, K.; Jeon, J.; Park, K.; Kim, J. *Exp. Mol. Med.* **2017**, *49* (4), e324–e324.

(7) Florea, M.; Kudithipudi, S.; Rei, A.; González-Álvarez, M. J.; Jeltsch, A.; Nau, W. M. *Chem. - A Eur. J.* **2012**, *18* (12), 3521–3528.

(8) Liu, Y.; Perez, L.; Gill, A. D.; Mettry, M.; Li, L.; Wang, Y.; Hooley, R. J.; Zhong, W. *J. Am. Chem. Soc.* **2017**, *139* (32), 10964–10967.

(9) Guagnini, F.; Antonik, P. M.; Rennie, M. L.; O'Byrne, P.; Khan, A. R.; Pinalli, R.; Dalcanale, E.; Crowley, P. B. *Angew. Chemie Int. Ed.* **2018**, *57* (24), 7126–7130.

(10) Beatty, M. A.; Borges-González, J.; Sinclair, N. J.; Pye, A. T.; Hof, F. *J. Am. Chem. Soc.* **2018**, *140* (10), 3500–3504.

(11) Gruber, T. *ChemBioChem* **2018**, *19* (22), 2324–2340.

(12) Zamora-Olivares, D.; Kaoud, T. S.; Dalby, K. N.; Anslyn, E. V. *J. Am. Chem. Soc.* **2013**, *135* (39), 14814–14820.

(13) Pinkin, N. K.; Waters, M. L. *Org. Biomol. Chem.* **2014**, *12* (36), 7059–7067.

(14) Gober, I. N.; Waters, M. L. *Org. Biomol. Chem.* **2017**, *15* (37), 7789–7795.

(15) James, L. I.; Beaver, J. E.; Rice, N. W.; Waters, M. L. *J. Am. Chem. Soc.* **2013**, *135* (17), 6450–6455.

(16) Liu, K.; Guo, Y.; Liu, H.; Bian, C.; Lam, R.; Liu, Y.; Mackenzie, F.; Rojas, L. A.; Reinberg, D.; Bedford, M. T.; Xu, R.-M.; Min, J. *PLoS One* **2012**, *7* (2), e30375.

(17) Corbett, P. T.; Leclaire, J.; Vial, L.; West, K. R.; Wietor, J.-L.; Sanders, J. K. M.; Otto, S. *Chem. Rev.* **2006**, *106* (9), 3652–3711.

(18) Gamal-Eldin, M. A.; Macartney, D. H. *Org. Biomol. Chem.* **2013**, *11* (3), 488–495.

(19) Liu, K.; Chen, C.; Guo, Y.; Lam, R.; Bian, C.; Xu, C.; Zhao, D. Y.; Jin, J.; MacKenzie, F.; Pawson, T.; Min, J. *Proc. Natl. Acad. Sci.* **2010**, *107* (43), 18398–18403.

(20) Hamieh, S.; Saggiomo, V.; Nowak, P.; Mattia, E.; Ludlow, R. F.; Otto, S. *Angew. Chemie Int. Ed.* **2013**, *52* (47), 12368–12372.

(21) Chung, M.-K.; Severin, K.; Lee, S. J.; Waters, M. L.; Gagné, M. R. *Chem. Sci.* **2011**, *2* (4), 744.

(22) Iyengar, B. S.; Dorr, R. T.; Alberts, D. S.; So, M.; Remers, W. A. *J. Med. Chem.* **1997**, *2623* (97), 3734–3738.

(23) Beaver, J. E.; Waters, M. L. *ACS Chem. Biol.* **2016**, *11* (3), 643–653.

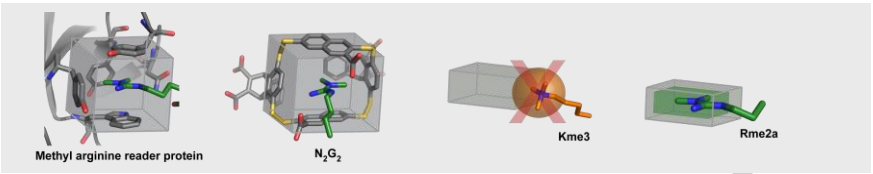
10.1002/anie.201814645

WILEY-VCH

Entry for the Table of Contents

Layout 2:

COMMUNICATION



Tailor-made: By introducing “lids” on a box-like receptor, a perfect fit for asymmetric dimethylarginine is achieved, to the exclusion of symmetric dimethylarginine and trimethyllysine.

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

Page No. – Page No.

Title