FULL PAPER

Binding methylarginines and methyllysines as free amino acids: a comparative study of multiple host classes

Zoey Warmerdam,[1, a] Bianca E. Kamba,[1, b] My-Hue Le,[c] Thomas Schrader,[c] Lyle Isaacs,[d*] Peter Bayer,[b*] and Fraser Hof[a*]

Dedicated to François Diederich, enthusiastic promotor of collaboration and partnerships

[a] Department of Chemistry and the Centre for Advanced Materials and Related Technology
University of Victoria
3800 Finnerty Rd, V8W 3V6 Victoria, BC, Canada
* E-mail: fhof@uvic.ca
[b] Department of Structural and Medicinal Biochemistry
Universität Duisburg Essen
Universitätstrasse 2, 45141 Essen, Germany
* E-mail: peter.bayer@uni-due.de
[c] Department of Chemistry
Universität Duisburg Essen
Universitätstrasse 7, 45117 Essen, Germany
[d] Department of Chemistry and Biochemistry
University of Maryland
College Park, MD 20742, United States
* E-mail: lisaacs@umd.edu
1 These authors contributed equally to the work in this paper

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

Abstract: Methylated free amino acids are an important class of targets for host-guest chemistry that have recognition properties distinct from those of methylated peptides and proteins. We present comparative binding studies for three different host classes that are each studied with multiple methylated arginines and lysines to determine fundamental structure-function relationships. The hosts studied are all anionic and include three calixarenes, two acyclic cucurbiturils, and two other cleft-like hosts, a clip and a tweezer. We determined the binding association constants for a panel of methylated amino acids using indicator displacement assays. The acyclic cucurbiturils display stronger binding to the methylated amino acids, and some unique patterns of selectivity. The two other cleft-like hosts follow two different trends, shallow host (clip) following similar trends to the calixarenes, and the other more closed host (tweezer) binding certain less-methylated amino acids stronger than their methylated counterparts. Molecular modeling sheds some light on the different preferences of different hosts. The results identify hosts with new selectivities and with affinities in a range that could be useful for biomedical applications. The overall selectivity patterns are explained by a common framework that considers the geometry, depth of binding pockets, and functional group participation across all host classes.

Introduction

The selective binding of free amino acids in physiologically relevant solutions is difficult to achieve. In neutral aqueous solution arginine and lysine are zwitterionic and have a relatively small hydrophobic surface area, combined with relatively high charge, which means that they are strongly solvated by water. This results in unfavourable binding energy upon complexation with the host, as several water molecules need to be released from strong hydrogen bonds to the amino acid's charged groups. The common presence of salts and other competing co-solutes creates further challenges.[1]

There are dozens of diseases caused by disordered amino acid metabolism.[2] Directly capturing and sequestering disease-related amino acids is a new approach that might be useful for diagnosis, disease monitoring, and possibly for direct therapeutics.[3] Antibodies are known that bind amino acids but have some inherent shortcomings[4] that could be overcome by the creation of synthetic organic binding tools for amino acids. Synthetic hosts are a great starting point to develop new tools. The hosts have a concave binding pocket where the molecular recognition takes place via non-covalent interactions (e.g. charge-charge interactions, hydrogen bonding, van der Waals forces, π-π interactions, etc.) and the hydrophobic effect. These binding pockets come with different shapes, sizes and chemical properties. The recognition of amino acids by hosts has recently been reviewed by Basílio et al.[5] Examples of hosts binding to free amino acids include calixarenes[6], pillararenes[7], cucurbiturils[8], cyclodextrins.[9]

We studied a set of seven hosts from three different classes to get a better fundamental understanding of the structure-function relationship for amino acid binding by multiple different kinds of hosts. Calixarenes are relative shallow but easily functionalized molecules, where the functionalization often directly lines the binding pocket.[10] We chose a parent anionic calixarene, **sCx4**, and the functionalized analogs **sCx4-NO$_2$** and **sCx4-CHO** (Figure 1a-c) to include in this study after a preliminary screen (not shown) ruled out some weaker-binding calixarenes. Cucurbituril (CB) hosts have a deeper and more rigid binding pocket, but the functional group additions to CBs often do not directly influence the binding properties as they happen on the outside of the host. Acyclic CB analogs such as **M1** and **M2** (Figure 1d,e) can be functionalized along the edge of their binding

surfaces, and are increasingly being used in biomedical applications.[11] Another more rigid cleft-like host family includes the 'tweezers' and 'clips' introduced by Klärner and Schrader.[12] The "phosphate clip" **PC** carries planar aromatic sidewalls ideal for aromatic cations, whereas the "tweezers" **CLR01** form a torus-shaped cavity that is selective for arginine and lysine (Figure 1f, g).

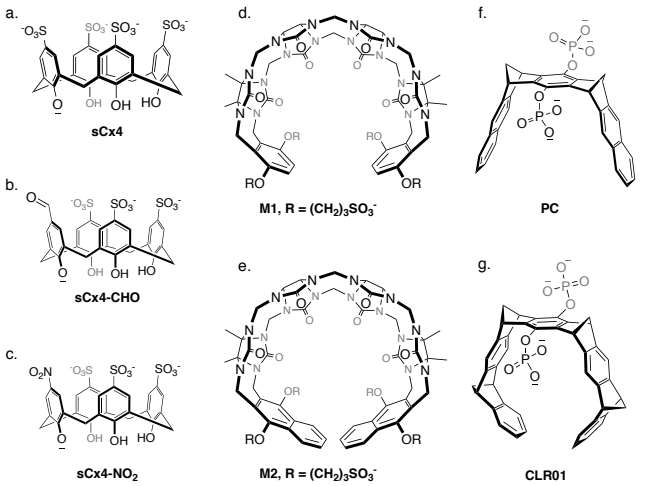


Figure 1. Hosts studied in this report. Calixarenes a) **sCx4**, b) **sCx4-CHO,** and c) **sCx4-NO₂.** Cleft-like hosts d) **M1**, e) **M2**, f) **PC**, and g) **CLR01**.

For this fundamental, comparative study we have selected methylated amino acids as binding targets. Our test set (Figure 2) includes arginine and lysine, and each of their physiologically relevant methylated states: (monomethyl arginine (MMA), asymmetric dimethyl arginine (ADMA), symmetric dimethyl arginine (SDMA), monomethyl lysine (MML) dimethyl lysine (DML) and trimethyl lysine (TML)). They have subtle structural differences among them,[13] which make them an interesting test case for molecular recognition and selectivity. They participate in diverse biological pathways that are relevant to multiple pathological states. Methylated arginines have been linked to a wide array of diseases such as cardiovascular disease,[14] renal disease,[15] and others.[16] While the exact values are sometimes controversial, meta-analyses show that the ADMA plasma concentration in healthy adults is ~0.7 μM, and higher in various disease states (with some studies showing plasma [ADMA] for disease conditions as high as 5-7 μM).[17] SDMA plasma levels are lower (<1 μM).[18] The plasma concentration of the methylated lysines has not significantly been linked to disease, but studying them will provide us insight into this family of complexes.

Lots of research has been focused on binding methylated lysines,[6c, 19] while the selective binding of methylated arginine has received relatively less attention.[20] Most prior literature has focused on binding these methylated residues in the context of proteins and peptides. The current study on free amino acids is motivated by a body of literature[14a, 21] demonstrating that some of the free methylated amino acids are metabolites that play critical biological roles in pathways that are distinct from those involving whole proteins, protein tails, and peptides with post-translationally methylated residues.

Results and Discussion

The binding constants were determined for the complexes formed by each member of the host library with each of the guests, using indicator displacement assays (IDAs). We adapted previously reported IDAs[22] for **sCx4** (using lucigenin, Figure 2b)[23] and **M1** and **M2** (using Rhodamine 6G, Figure 2b).[24] We established IDAs for **PC** and **CLR01** during this work. Initial experiments using Rhodamine 6G, Proflavine, and Neutral Red revealed either non-ideal stoichiometries of binding or inadequate intensity changes upon binding. Studies using 4-ASP (Figure 2b) as the indicator gave reliable results for **PC** and **CLR01** with the whole panel of guests. Each host-indicator dissociation constant (K_{ind}) was determined by a direct titration of the host into indicator, and the host-guest equilibrium dissociation constants (K_d) were then determined using competitive titrations of guests into a pre-formed host-indicator complex. The calixarene and the acyclic CB IDAs gave a turn-on signal, where displacement of the indicator by the guest results in an increase of fluorescence emission. The **PC** and **CLR01** host provided a turn-off signal where displacement of 4-ASP quenches its fluorescence emission. All the titrations were optimized to work in a 10 mM Na_2HPO_4 buffer at pH 7.4. This choice of buffer rules out the ability to measure weak (>mM) binding (which would require analyte concentrations that would overwhelm the buffer), but it ensures that the trends for stronger-binding guests can be compared across different host types. The studied host-guests show a range of affinities (Table 1), which we categorize as follows for convenient presentation: strong binding (K_d < 200 μM), medium to strong binding (K_d 200–1500 μM) and weak to no binding (K_d > 1500 μM). We also presented the data as a 3D bar graph (Figure 3), to help visualise the binding trends.

For the calixarene hosts, strong binding is only observed for a few host-guest combinations. A binding trend is observed for the calixarene-guest complexes in which higher methylation states generally result in stronger binding. Most notably, trimethyllysine (TML) is the guest that shows strongest binding with each of the sulfonatocalixarenes. This trend is less straightforward for the arginine guests although the lower methylation states show no binding within the limits of this assay condition. Comparing the symmetric and asymmetric isomers of dimethylarginine, ADMA displays a ~2-fold stronger binding. Interestingly, **sCx4-NO₂** selectively binds ADMA over the other arginine guests, for which no binding is observed. Overall, the sCx-hosts bind the free amino acids weakly compared to physiological concentration ranges for these amino acids.

Strong binding of the amino acids is consistently observed for **M1** and **M2**. **M1** binds all three methylated arginines, while ignoring unmethylated arginine. **M1** is slightly selective for ADMA (K_d = 10 μM) over MMA (K_d = 45 μM) and SDMA (K_d = 35 μM). It has a similar selectivity for the methylated lysines showing no binding to lysine and strong binding for the methylated lysines, being selective for TML (K_d = 15 μM) over DML (K_d = 70 μM) and MML (K_d = 160 μM). **M2** displays a sharp drop off between strongly binding highly methylated guests (ADMA, SDMA, DML, TML) and the weak binding of lower methylation states. CB[n]-type receptors are known to prefer quaternary over primary ammoniums.[25] Among the stronger binding guests, **M2** has a >3-fold selectivity for SDMA (K_d = 20 μM) over ADMA (K_d = 60 μM), and a >10-fold selectivity for TML (K_d = 12 μM) over DML (K_d = 130 μM).

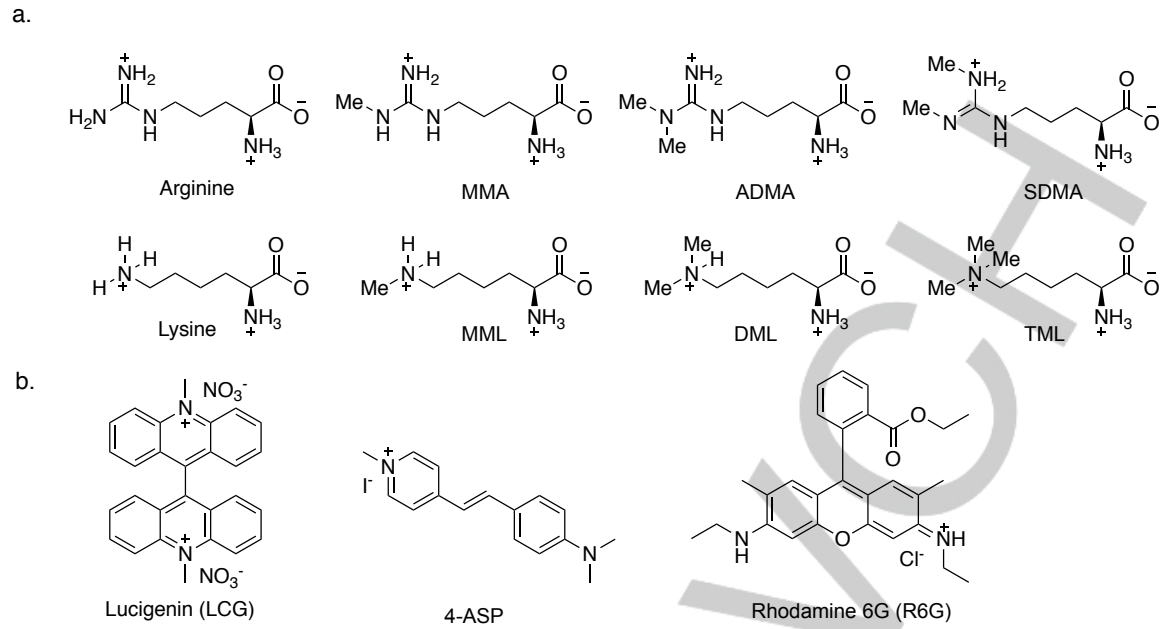


Figure 2. Guests and indicators studied in this report. a) Guests arginine, monomethylarginine (MMA), asymmetric dimethylarginine (ADMA), symmetric dimethylarginine (SDMA), Lysine, monomethyllysine (MML), dimethyllysine (DML), and trimethyllysine (TML), b) Indicators LCG, 4-ASP, and R6G.

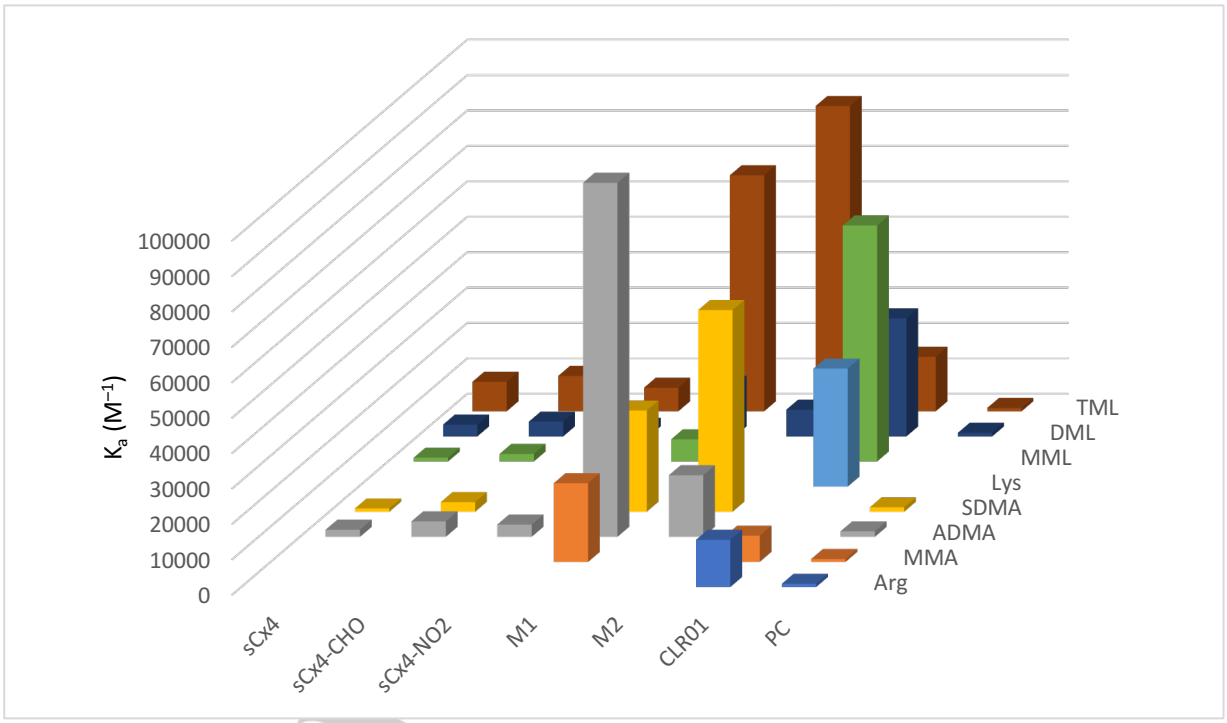


Figure 3. *3D bar graph to visualize the binding trends, represented as the K_a values in units M^{-1} (K_a = 1/K_d). See Table 1 for experimental details.*

Table 1. *K_d values determined by IDA for each host-guest complex in 10 mM phosphate buffer.* [a]

	sCx4[b]	**sCx4-CHO**[b]	**sCx4-NO₂**[b]	**M1**[c]	**M2**[c]	**CLR01**[d]	**PC**[d]
	K_d (μM)	K_d (μM)	K_d (μM)	K_d (μM)	K_d (μM)	K_d (μM)	K_d (μM)
Arginine	>1500	>1500	>1500	>1500	>1500	75 ± 20	1100 ± 630
MMA	>1500	>1500	>1500	45 ± 10	>1500	135 ± 30	1230 ± 540
ADMA	500 ± 50	230 ± 50	290 ± 60	10 ± 5	60 ± 20[e]	>1500	650 ± 130
SDMA	1090 ± 300	370 ± 150	>1500	35 ± 10	20 ± 5	>1500	810 ± 270
Lysine	>1500	>1500	>1500	>1500	>1500	30 ± 5	>1500
MML	930 ± 120	470 ± 110	610 ± 130	160 ± 60	>1500	15 ± 5	>1500
DML	300 ± 40	240 ± 50	410 ± 60	70 ± 10	130 ± 90[e,f]	30 ± 5	980 ± 710
TML	120 ± 20	100 ± 30	150 ± 30	15 ± 5	12 ± 3	65 ±20	1040 ± 360

[a] All titrations were carried out in 10 mM NaH_2PO_4 buffer at pH 7.4. See the Supporting Information for titration curves experimental details, and fitting details. All K_d values arise from fits with $R^2 \geq 0.95$ except where indicated. [b] Lucigenin (LCG) was used as indicator. [c] Rhodamine 6G (R6G) was used as indicator. [d] 4-(4-Diethylaminostyryl)-1-methylpyridinium iodide (4-ASP) was used as indicator. [e] $R^2 \geq 0.92$. [f] Value arises from one triplicate.

Despite their functional group similarity, **PC** and **CLR01** have different behaviors. **PC** has relatively weak binding across the set of guests, but some trends can be observed. The higher the methylation state of the arginine guests the stronger the binding. When looking at the lysine guests, the binding is weak with relatively large uncertainties, and no strong conclusions can be drawn relating to binding trends. **CLR01** shows different binding trends than the other hosts. The binding is strong, and unlike the other hosts this also includes <100 μM binding of each unmethylated amino acid.[26] Lysine binds with a K_d = 30 μM, and the slightly more hydrophobic MML binds more strongly at K_d = 15 μM. Besides this one exception, each other amino acid binds progressively weaker with increasing methylation. Unlike for all other hosts, dimethylarginines are not measurably bound by **CLR01** under the conditions of the experiment.

To determine the effect of a higher salt concentration, IDAs on the amino acids were also done in a more concentrated 50 mM phosphate buffer with the three strongest binding hosts: **M1**, **M2** and **CLR01**. We see that the strong binding and the binding trends are maintained for **M2** and **CLR01**. The binding strength for **M1** becomes 2–4-fold weaker for each guest, but still remains in the "strong" range of K_d values, see Table 2.

Studies with a small set of simpler guests show how the impact of hydrophobicity manifests differently in the different host classes. We did IDA studies with dimethylammonium chloride, tetramethylammonium chloride, and benzyltrimethylammonium chloride as controls that display increasing hydrophobicity without neighboring polar functionality and with relatively simple changes in size and shape (Supp. Info.). Dimethylammonium is the smallest, least hydrophobic, resulting in weak binding (K_d ≥1 mM) across the set of hosts. In shallow hosts like the calixarenes and **PC**, tetramethylammonium generally binds in the range of K_d = 300–900 μM, with approximately the same affinity as benzyltrimethylammonium in each case. The deeper hosts **M1**, **M2**, and **CLR01** show clear preference for benzyltrimethylammonium over tetramethylammonium, as would be expected from the guest's increased size and hydrophobicity. We rationalise that this difference in trends for benzyltrimethylammonium arises because the shallow hosts don't have any additional hydrophobic surface area available to interact with the larger guest.

The hosts studied in this paper (except **sCx4-CHO**, **sCx4-NO2** and **PC**) have previously been studied via ITC with lysine and arginine. [6c, 27] From those studies, we see that enthalpy is the dominant driving force across all hosts (Table S2). ITC studies for **sCx4** with methyllysine and methylarginine analogs show that increasing methylation is accompanied by significant increases in enthalpic driving force, and smaller changes in entropy. As in all similar work, we caution against drawing specific mechanistic conclusions for host-guest ITC studies that are done in salty aqueous solution, because salt effects are often large and unpredictable contributors to the heats of binding.

Table 2. *K_d values determined by IDA for each host-guest complex in 50 mM phosphate buffer.* [a]

	M1[b]	**M2**[b]	**CLR01**[c]
	K_d (μM)	K_d (μM)	K_d (μM)
Arginine	>1500	>1500	90 ± 25
MMA	180 ± 60	>1500	155 ± 30
ADMA	30 ± 10	125 ± 45	>1500
SDMA	55 ± 15	40 ± 10	>1500
Lysine	>1500	>1500	50 ± 15
MML	300 ± 150	>1500	35 ± 5
DML	210 ± 50	190 ± 110[d]	55 ± 10
TML	30 ± 5	25 ± 5	40 ± 5

a] All titrations were carried out in 50 mM NaH_2PO_4 buffer at pH 7.4. See the Supporting Information for titration curves, experimental details, and fitting details. All K_d values arise from fits with $R^2 \geq 0.95$ except where indicated. [b] Rhodamine 6G (R6G) was used as indicator. [c] 4-(4-Diethylaminostyryl)-1-methylpyridinium iodide (4-ASP) was used as indicator. [d] $R^2 \geq 0.93$.

We did molecular modeling for six of the hosts to provide a general view of the differences and similarities between the hosts for certain key complexes. Molecular modeling was done using minimization in explicit water for all the indicated complexes using Maestro (Schrödinger, Inc). Host **sCx4-CHO**, **sCx4-NO_2**, **M1**, **M2**, **PC**, **CLR01** were modeled in complex with ADMA (Figure 4) and hosts **M1**, **M2**, **PC**, **CLR01** were also modeled each in complex with MML and TML (Supporting Information) in order to gain further insight into selectivity among these guests. The models reveal the extent and nature of interactions between hosts and guests and show qualitatively that the different hosts generally fit into two classes: open geometry hosts that engage only the charged side chain of the guests (the calixarenes and **PC**), and closed geometry hosts that almost completely surround their guests (**M1**, **M2**, and **CLR01**). They also reveal differences in host-guest interactions among the different host classes. The connections between these structural features and guest-binding selectivities are discussed below.

By comparing across host classes, we can derive some general lessons about the contributions from electrostatics, hydrophobicity, and geometric shape matching.

While electrostatics are undoubtedly important for molecular recognition, our binding data show that they are not the key determining factor for guest selectivity. We know from other literature that neutral guests do not bind the calixarene hosts as strongly[6c], where the tweezer-type host do have a precedent as strong binding host.[26, 27b] Yet, the comparison across all three host classes makes it clear that the *selective* binding of charged species in this relatively salty environment is not strongly controlled by the charges on hosts or guests. All amino acids studied have a zwitterionic α-amino-acid component, and a net charge of +1. The methylated guests have positive charges that are distributed significantly onto their methyl CH atoms, changing the nature of electrostatic interactions with the hosts.[28] The calixarenes and acyclic CBs have net charges of between –4 and –5, and yet bind guests with affinities that vary over >2 orders of magnitude. The clip-type hosts have net charges of between –2 and –4 (depending on the degree of second ionization of the phosphate groups) and generate K_d values that vary by almost 100-fold between different guests. **M1** and **M2** have eight carbonyl groups that can form ion-dipole interactions, and two sulfonate groups that can form ion-ion interactions. Molecular modeling corroborates that the ion-dipole interactions are the main electrostatic interaction that is taking place for **M1** and **M2**, whereas charged interactions between the guests and the hosts' sulfonate arms are not as prominently seen in the energy-minimized structures (Figure 4c-d). The calixarenes also have

multiple sulfonate groups, and in these complexes host-guest salt bridges are prominent features of the complexes (Figure 4a-b and Supporting Information). **PC** and **CLR01** both have two phosphate groups that can have ion-ion interactions between the charged phosphate groups and the guanidinium, see Figure 4e-f. All the hosts can form cation-pi interactions with their cationic guests, although the geometric details vary.

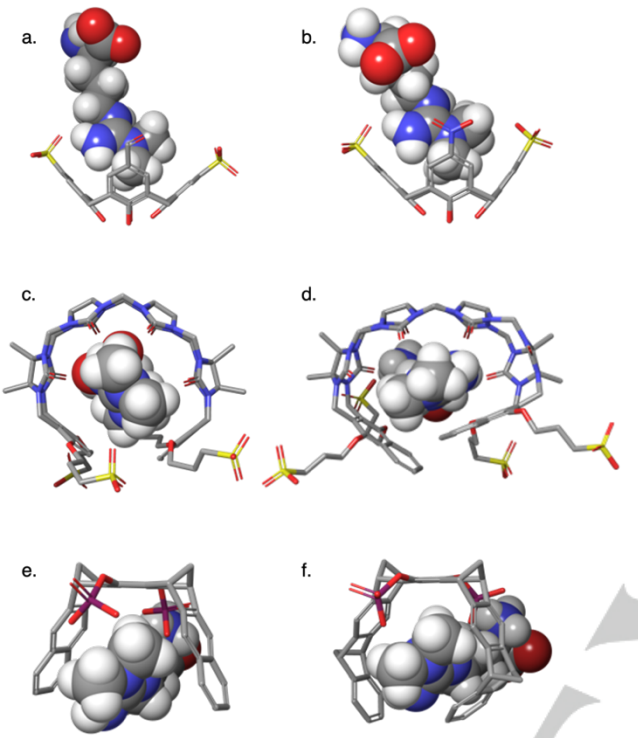


Figure 4. Molecular modelling of each host in complex with ADMA. Calixarenes a) **sCx4-CHO**, b) **sCx4-NO$_2$** . Cleft-like hosts c) **M1**, d) **M2**, e) **PC**, and f) **CLR01**. Molecules were energy-minimized in explicit water (not shown) using OPLS_2005 as implemented in Maestro (Schrödinger, Inc). See Supporting information for more views and for other host-guest complexes.

Guest hydrophobicity is the main determinant for guest selectivity in hosts with more open geometries, including all calixarenes and **PC**. Methylation of arginine and lysine increases the volume of the head groups. This results in an increased hydrophobic surface area, a change to a more diffuse charge distribution, and decrease in the guest's potential to form strong NH hydrogen bonds.[28] This modification is favourable when looking at the calixarenes and **PC**, where an increased number of methyl groups results in a stronger binding. This trend can clearly be observed for the lysine guests, where unmethylated lysine displays weak binding and TML displays strong binding to the hosts (Table 1). When looking at the dimethylated arginines we also see that position of the methyl groups has an influence. The models show that the hydrophobic surface area of ADMA is localized in one patch made up of two geminal methyl groups, whereas the hydrophobic surface area of SDMA is separated with methyls on two distal nitrogen atoms. In open-geometry hosts that don't constrain for molecular shape, this results in a stronger engagement of ADMA in the host pockets relative to SDMA.

Geometric shape matching also contributes to the binding strengths and selectivity of the host-guest complexes. The relative openness of geometry is a key determinant.[29] A shallower or open binding pocket is unfavourable for binding the free amino acids, as would be expected from the arguments made above. This can be clearly observed for the calixarenes and **PC**, where there isn't a large complementary overlap of the surface area, the binding is weak. Hosts with closed geometries can display selectivities that run against the underlying trends caused by hydrophobicity. In general, increasing methylation increases the tendency to bind, but binding can be discouraged when shape and fit are incompatible. This can most clearly be observed for **CLR01** and **M2**. **CLR01** is in the middle of the "openness range" of this library, it has nine rings forming its tweezer shape. Its binding pocket is a perfect fit to bind medium-sized hydrophobic molecules and is the only host in this library with strong binding for both unmethylated arginine and lysine. When looking at the dimethylated arginines, ADMA's geminal methyls can fit into the pocket of most hosts (see Figure 4). The exception is **M2**, where ADMA binding is 3-fold weaker compared to SDMA. This makes **M2** a rare example of an SDMA-selective host molecule. The structural difference between **M1** and **M2** is the naphthalene of **M2**. This creates a bigger, more hydrophobic, and somewhat more closed-off binding pocket. This small change makes a significant difference in the binding properties of the host-guest complexes. When looking at MML versus TML we can see that **M1** keeps the same formation where **M2** must accommodate for its bigger structure and is slightly askew (see Supporting Information). **PC** and **CLR01** are more rigid molecules, and do not have the option to flex to accommodate guests in the same way as the acyclic cucurbiturils.

Conclusion

Our studies have revealed some interesting individual complexes, as well as some broader trends that arise from comparisons. We found that the **CLR01** host was the best size for the non-methylated amino acids, showing a strong binding for arginine and lysine which the other hosts do not. The methylated arginine guests are bound the strongest by the acyclic CB hosts, but in a weaker range of K_d values the host **sCx4-NO$_2$** shows good selectivity for ADMA over all other arginines. TML is bound strongly by all hosts, except **PC**, but is only bound with good selectivity over other related guests by **M2**. Although most of the hosts bind their guests outside of the targeted physiological concentration ranges, the K_d for the complex of **M1**-ADMA is at the upper limit of the target concentration for that amino acid, which motivates further work on this class of hosts. While this study is fundamental and not targeted at applied science, we do identify useful new selectivities, such as the complete selectivity of **CLR01** for methyllysines over methylarginines, and the novel selectivity for SDMA over ADMA that is displayed by **M2**. This study also allows us to see some interesting trends emerge from direct comparison of different host classes under identical conditions. For example, in most ways **PC** behaves more like the **sCx4** hosts than like its close chemical relative **CLR01**. In this work we can tie that similar behaviour across many guests to similarities in host geometry that mostly override the more obvious differences in functional group identity and arrangement that typically dominate our thinking about host-guest binding. We think that additional non-dogmatic, collaborative, open comparisons of host molecules[30] would lead to more such insights.

Acknowledgements

We thank Hendrik Kirschner and Christine Beuck for carrying out preliminary studies. We thank Rebecca Hof for training and Allison Selinger for helpful discussions on the data analysis for indicator displacement assays.

Funding

We thank NSERC (RGPIN-2019-04806), the National Science Foundation (CHE-1807486), Deutsche Forschungsgemeinschaft (DFG) to TS (Schr 604/19-1) and PB (BA1624/15-1) for their financial support.

Keywords: macrocycles • methylated amino acids • host-guest chemistry • indicator displacement assays

[1] M. A. Beatty, F. Hof, *Chem. Soc. Rev.* **2021**.

[2] [a] S. Schlesinger, S. R. Sonntag, W. Lieb, R. Maas, *PLoS One* **2016**, *11*, e0165811; [b] R. J. DeBerardinis, C. B. Thompson, *Cell* **2012**, *148*, 1132-1144.

[3] [a] G. A. Gowda, S. Zhang, H. Gu, V. Asiago, N. Shanaiah, D. Raftery, *Expert. Rev. Mol. Diagn.* **2008**, *8*, 617-633; [b] V. J. Ruigrok, M. Levisson, M. H. Eppink, H. Smidt, J. van der Oost, *Biochem. J.* **2011**, *436*, 1-13.

[4] [a] T. A. Egelhofer, A. Minoda, S. Klugman, K. Lee, P. Kolasinska-Zwierz, A. A. Alekseyenko, M. S. Cheung, D. S. Day, S. Gadel, A. A. Gorchakov, T. Gu, P. V. Kharchenko, S. Kuan, I. Latorre, D. Linder-Basso, Y. Luu, Q. Ngo, M. Perry, A. Rechtsteiner, N. C. Riddle, Y. B. Schwartz, G. A. Shanower, A. Vielle, J. Ahringer, S. C. Elgin, M. I. Kuroda, V. Pirrotta, B. Ren, S. Strome, P. J. Park, G. H. Karpen, R. D. Hawkins, J. D. Lieb, *Nat. Struct. Mol. Biol.* **2011**, *18*, 91-93; [b] G. Kungulovski, A. Jeltsch, *F1000Research* **2015**, *4*, 1160; [c] 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; [d] P. Acharya, A. Quinlan, V. Neumeister, *F1000Research* **2017**, *6*, 851.

[5] J. N. Martins, J. C. Lima, N. Basilio, *Molecules* **2020**, *26*.

[6] [a] N. Douteau-Guevel, A. W. Coleman, J. P. Morel, N. Morel-Desrosiers, *J. Chem. Soc. Perk. Trans. 2.* **1999**, 629-633; [b] N. Douteau-Guevel, A. W. Coleman, J. P. Morel, N. Morel-Desrosiers, *J. Phys. Org. Chem.* **1998**, *11*, 693-696; [c] C. S. Beshara, C. E. Jones, K. D. Daze, B. J. Lilgert, F. Hof, *ChemBioChem* **2010**, *11*, 63-66.

[7] [a] D. Hessz, S. Badogos, M. Bojtar, I. Bitter, L. Drahos, M. Kubinyi, *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.* **2021**, *252*, 119455; [b] C. Li, J. Ma, L. Zhao, Y. Zhang, Y. Yu, X. Shu, J. Li, X. Jia, *Chem. Commun.* **2013**, *49*, 1924-1926.

[8] [a] M. A. Gamal-Eldin, D. H. Macartney, *Org. Biomol. Chem.* **2013**, *11*, 488-495; [b] F. Biedermann, W. M. Nau, *Angew. Chem. Int. Ed.* **2014**, *53*, 5694-5699; [c] D. M. Bailey, A. Hennig, V. D. Uzunova, W. M. Nau, *Chem. Eur. J.* **2008**, *14*, 6069-6077.

[9] M. N. Roy, D. Ekka, S. Saha, M. C. Roy, *RSC Adv.* **2014**, *4*, 42383-42390.

[10] S. Shinkai, K. Araki, O. Manabe, *J. Chem. Soc. Chem. Commun.* **1988**, 187-189.

[11] B. Zhang, L. Isaacs, *J. Med. Chem.* **2014**, *57*, 9554-9563.

[12] [a] I. Hadrovic, P. Rebmann, F. G. Klarner, G. Bitan, T. Schrader, *Front. Chem.* **2019**, *7*, 657; [b] P. Talbiersky, F. Bastkowski, F. G. Klarner, T. Schrader, *J. Am. Chem. Soc.* **2008**, *130*, 9824-9828.

[13] K. McQuinn, L. S. McIndoe, F. Hof, *Chem. Eur. J.* **2008**, *14*, 6483-6489.

[14] [a] J. P. Cooke, *Vasc. Med.* **2005**, *10*, S11-S17; [b] L. Sibal, S. C. Agarwal, P. D. Home, R. H. Boger, *Curr. Cardiol. Rev.* **2010**, *6*, 82-90.

[15] [a] J. T. Kielstein, S. R. Salpeter, S. M. Bode-Boeger, J. P. Cooke, D. Fliser, *Nephrol. Dial. Transpl.* **2006**, *21*, 2446-2451; [b] O. Tutarel, A. Denecke, S. M. Bode-Boger, J. Martens-Lobenhoffer, B. Schieffer, M. Westhoff-Bleck, J. T. Kielstein, *Kidney Blood Press. Res.* **2011**, *34*, 41-45; [c] E. Schepers, D. V. Barreto, S. Liabeuf, G. Glorieux, S. Eloot, F. C. Barreto, Z. Massy, R. Vanholder, *Clin. J. Am. Soc. Nephrol.* **2011**, *6*, 2374-2383.

[16] Y. L. Tain, C. N. Hsu, *Toxins (Basel)* **2017**, *9*.

[17] [a] B. Nemeth, Z. Ajtay, L. Hejjel, T. Ferenci, Z. Abram, E. Muranyi, I. Kiss, *PLoS One* **2017**, *12*, e0177493; [b] G. L. Erre, A. A. Mangoni, F. Castagna, P. Paliogiannis, C. Carru, G. Passiu, A. Zinellu, *Sci. Rep.* **2019**, *9*, 5426.

[18] C. Fleck, A. Janz, F. Schweitzer, E. Karge, M. Schwertfeger, G. Stein, *Kidney. Int. Suppl.* **2001**, *78*, S14-18.

[19] [a] K. D. Daze, T. Pinter, C. S. Beshara, A. Ibraheem, S. A. Minaker, M. C. F. Ma, R. J. M. Courtemanche, R. E. Campbell, F. Hof, *Chem. Sci.* **2012**, *3*, 2695-2699; [b] Lindsey A. Ingerman, M. E. Cuellar, M. L. Waters, *Chem. Commun.* **2010**, *46*, 1839–1841; [c] T. Gruber, *ChemBioChem* **2018**, *19*, 2324-2340.

[20] [a] A. G. Mullins, N. K. Pinkin, J. A. Hardin, M. L. Waters, *Angew. Chem. Int. Ed.* **2019**, *58*, 5282-5285; [b] L. I. James, J. E. Beaver, N. W. Rice, M. L. Waters, *J. Am. Chem. Soc.* **2013**, *135*, 6450-6455.

[21] [a] S. M. Bode-Boger, F. Scalera, L. J. Ignarro, *Pharmacol. Ther.* **2007**, *114*, 295-306; [b] L. Adler-Abramovich, L. Vaks, O. Carny, D. Trudler, A. Magno, A. Caflisch, D. Frenkel, E. Gazit, *Nat. Chem. Biol.* **2012**, *8*, 701-706; [c] P. R. Blackburn, J. M. Gass, F. P. E. Vairo, K. M. Farnham, H. K. Atwal, S. Macklin, E. W. Klee, P. S. Atwal, *Appl. Clin. Genet.* **2017**, *10*, 57-66; [d] I. E. Emrich, A. M. Zawada, J. Martens-Lobenhoffer, D. Fliser, S. Wagenpfeil, G. H. Heine, S. M. Bode-Boger, *Clin. Res. Cardiol.* **2018**, *107*, 201-213.

[22] R. N. Dsouza, U. Pischel, W. M. Nau, *Chem. Rev.* **2011**, *111*, 7941-7980.

[23] Z. Warmerdam, B. E. Kamba, A. Shaurya, X. X. Sun, M. K. Maguire, F. Hof, *Supramol. Chem.* **2021**.

[24] [a] D. Ma, B. Zhang, U. Hoffmann, M. G. Sundrup, M. Eikermann, L. Isaacs, *Angew. Chem. Int. Ed.* **2012**, *51*, 11358-11362; [b] D. Ma, P. Y. Zavalij, L. Isaacs, *J. Org. Chem.* **2010**, *75*, 4786-4795.

[25] L. Cao, M. Sekutor, P. Y. Zavalij, K. Mlinaric-Majerski, R. Glaser, L. Isaacs, *Angew. Chem. Int. Ed.* **2014**, *53*, 988-993.

[26] M. Fokkens, T. Schrader, F. G. Klarner, *J. Am. Chem. Soc.* **2005**, *127*, 14415-14421.

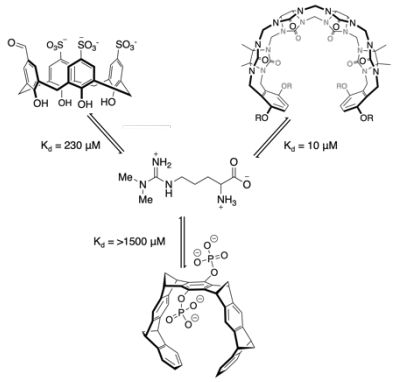
[27] [a] S. Dutt, C. Wilch, T. Gersthagen, P. Talbiersky, K. Bravo-Rodriguez, M. Hanni, E. Sanchez-Garcia, C. Ochsenfeld, F. G. Klarner, T. Schrader, *J. Org. Chem.* **2013**, *78*, 6721-6734; [b] S. A. Z. Ndendjio, L. Isaacs, *Supramol. Chem.* **2019**, *31*, 432-442.

[28] M. Evich, E. Stroeva, Y. G. Zheng, M. W. Germann, *Protein Sci.* **2016**, *25*, 479-486.

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

[30] M. Mallon, S. Dutt, T. Schrader, P. B. Crowley, *ChemBioChem* **2016**, *17*, 774-783.

Entry for the Table of Contents



We present comparative binding studies for three different host classes (calixarenes, acyclic cucurbiturils and other cleft-like hosts), each studied with multiple methylated arginines and lysines to determine fundamental structure-function relationships. Molecular modeling was used to give an insight on the preferences of the different hosts.

Supporting information

for

Binding methylarginines and methyllysines as free amino acids: a comparative study of multiple host classes

Table of Contents

1. General information and materials

The host used in these studies have all been previously published or are commercially available. **sCx4** (4-sulfocalic[4]arene Hydrate) was purchased from Tokyo Chemical Industry (TCI) (CAS 112269-92-8, >94.0%). **sCx4-CHO**[1], **sCx4-NO2**[2], **M1**[3], **M2**[3], **CLR01**[4] and **PC**[5] have been previously published. The dyes used in these studies have been purchased from Sigma-Aldrich, lucigenin (CAS 2315-97-1), 4-ASP (CAS 68971-03-9, 98%), R6G (CAS 989-38-8, ~95 %). The guests were purchased from different suppliers. Arginine was purchased from Calbiochem (CAS 1119-34-2, 99.7%). Lysine was purchased from USB Corporation (CAS 56-87-1). MMA (53308-83-1, ≥99%), and SDMA (30344-00-4, ≥97%) were purchased from ENZO. ADMA (CAS 220805-22-1) and MML (CAS 7622-29-9) were purchased from Toronto research chemicals. DML (CAS 79416-87-8, ≥96%), TML (55528-53-5, ≥97%), dimethylamine hydrochloride (506-59-2, 99%) and tetramethylammonium chloride (75-57-0, >98%) were purchased from Sigma-Aldrich. Benzyltrimethylammonium chloride (56-93-9, >99%) was purchased from Tokyo Chemical Industry Co., Ltd. (TCI).

2. IDA titrations

The IDAs were conducted in 384 well plates (Nunc™ 384-Well, Non-Treated, Flat-Bottom Microplate). The fluorescent signal was read on the Cytation 5 (Software Version 3.05.11) at room temperature. All wells had a final total volume of 50 μL. The IDA data was analysed in GraphPad Prism Version 8.3.0 (328). One exemplary replicate is shown for each host. All of the experiments for the 10 mM phosphate buffer were done as duplicates of triplicates, with the exception of the control compounds which were done as a single set of triplicates. All of the experiments in 50 mM phosphate buffer were done in triplicate.

The fluorescent signal for the calix[4]arene host titrations was read as a fluorescence endpoint measurement. Lucigenin (0.25 μM) was used as the indicator. The settings were as follows: Excitation: 369/20, Emission: 475/20. Optics: Top, Gain: extended. Light Source: Xenon Flash, Lamp Energy: High, Extended Dynamic Range. Read Speed: Normal, Delay: 100 msec, Measurements/Data Point: 10. Read Height: 10.5 mm.

The fluorescent signal for **CLR01** and **PC** titrations was read as a fluorescence endpoint measurement. 4-ASP (0.20 μM) was used as the indicator. The settings were as follows: Excitation: 490/10, Emission: 610/10. Optics: Bottom, Gain: extended. Light Source: Xenon Flash, Lamp Energy: High, Extended Dynamic Range. Read Speed: Normal, Delay: 100 msec, Measurements/Data Point: 10. Read Height: 10.5 mm.

The fluorescent signal for **M1** and **M2** titrations was read as a fluorescence endpoint measurement. R6G (0.10 μM) was used as the indicator. The settings were as follows: Excitation: 510/10, Emission: 550/10. Optics: Bottom, Gain: extended. Light Source: Xenon Flash, Lamp Energy: High, Extended Dynamic Range. Read Speed: Normal, Delay: 100 msec, Measurements/Data Point: 10. Read Height: 10.5 mm.

2.1 Equations

2.1.1. Outliers

Outliers are determined by the Dixon's Q-test, Equation 1, where the gap is the absolute difference between the outlier in question and the closest number to it. With three observations and at 95% confidence, Q > 0.970 = $Q_{95\%,n=3}$, we conclude the data point is an outlier.

Equation 1. Dixon's Q-test.

$$Q = \frac{gap}{range}$$

2.1.2. The standard error

The standard error is calculated for each triplicate, Equation 2.

Equation 2. Standard error

$$SD_{K_d} = \left(ln\ (10)\ * 10^{logK_i}\right) *\ SD_{logK_d}$$

SD_{Kd} = Standard error of K_d
$logK_d$ = Value of the log
SD_{logKd} = Standard error of the $logK_d$

2.1.3. The total standard error

The total standard error between two the triplicates is calculated by Equation 3.

Equation 3. Total standard error.

$$SD_{total} = \sqrt{\frac{(SD_1^2 + SD_2^2)}{n}}$$

SD_{total} = Standard derivation of both experiments
SD_1 = Standard derivation of the first triplicate
SD_2 = Standard derivation of the second triplicate
n = Number of experiments

2.1.4. Curve fit for the direct titration

To curve fit for the direct titration Equation 4, Equation 5 are used.

Equation 4. Curve fit for the direct titration for a turn on signal.

$$F = F_{min} + \frac{(F_{max} - F_{min}) * ([D] + [H] + K_d) - \sqrt{([D] + [H] + K_d)^2 - 4 * [H] * [D]}}{2 * [D]}$$

Equation 5. Curve fit for the direct titration for a turn off signal

$$F = F_{max} - \frac{(F_{max} - F_{min}) * ([D] + [H] + K_d) - \sqrt{([D] + [H] + K_d)^2 - 4 * [H] * [D]}}{2 * [D]}$$

F = Fitted data point
F_{max} = Maximum signal
F_{min} = Minimum signal
[D] = Molar concentration of dye in µM
[H] = Molar concentration of host (titrant)
K_d = Dissociation constant

2.1.5. Curve fit for the competitive titration

$$log_{EC50} = log\ (10^{log_{Ki}} * \left(1 + \frac{[D]}{K_d}\right))$$

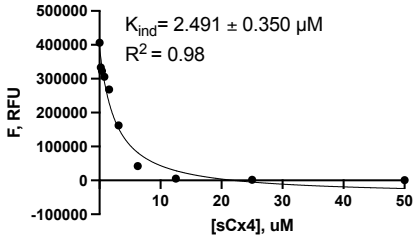
$$F = F_{min} + (F_{max} - F_{min})/(1 + 10^{(X - log_{EC50})})$$

log_{EC50} = log of the concentration of the competitor binding half-way between F_{min} and F_{max}
K_i = Equilibrium dissociation constant in Molar
[D] = Concentration of dye in nM
K_d = Equilibrium dissociation constant of the direct titration
F = Fitted data point
F_{max} = Maximum signal
F_{min} = Minimum signal

2.2. Fluorescence based studies 10 mM buffer

2.2.1. Fluorescence based studies of **sCx4**

a.



b.

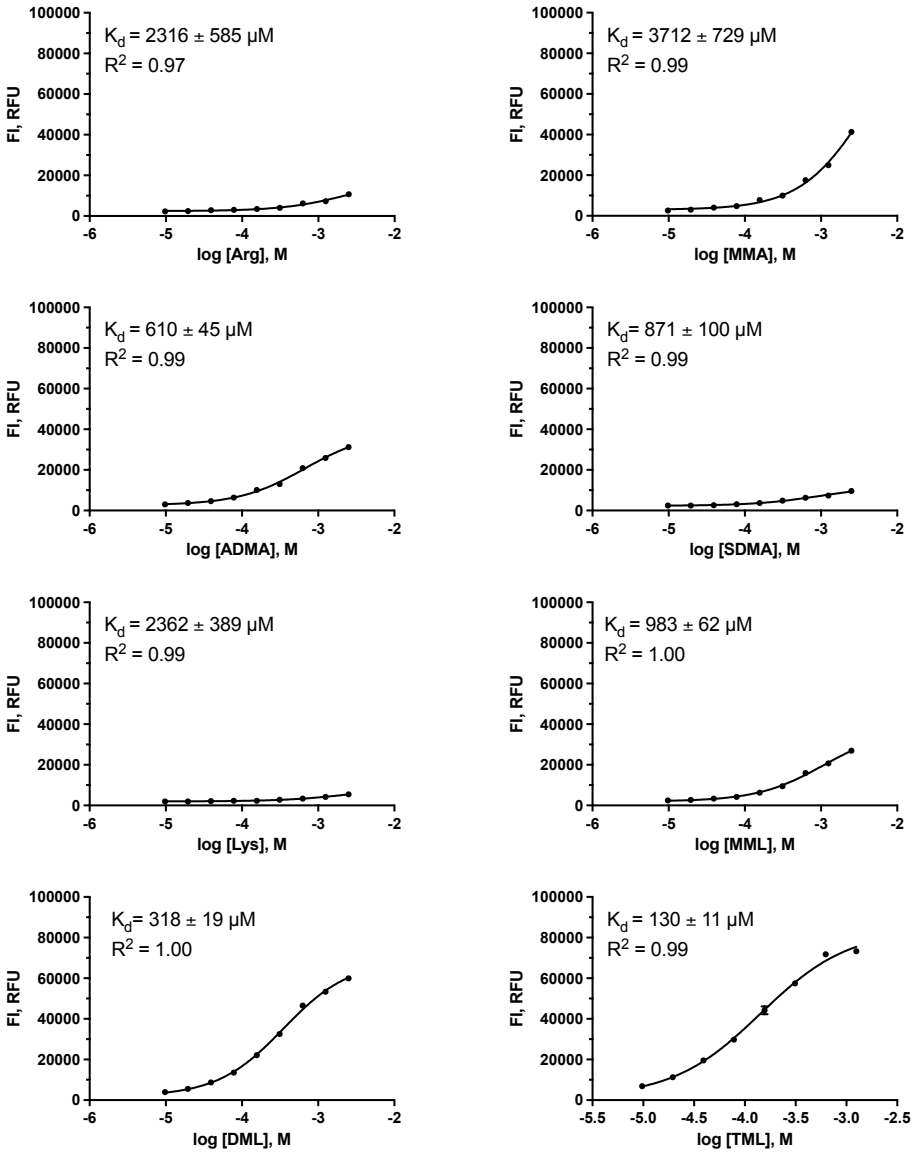


Figure S1: Fluorescence based studies of **sCx4**. a) Direct titration of LCG (0.25 µM) with **sCx4** (0 – 50 µM). b) Competitive titrations of Arginine, MMA, ADMA, SDMA, Lysine, MML, DML and TML (0 – 2.5 mM) individually titrated into the **sCx4**-LCG (5 µM **sCx4**, 0.25 µM LCG) complex.

2.2.2. Fluorescence based studies of **sCx4-CHO**

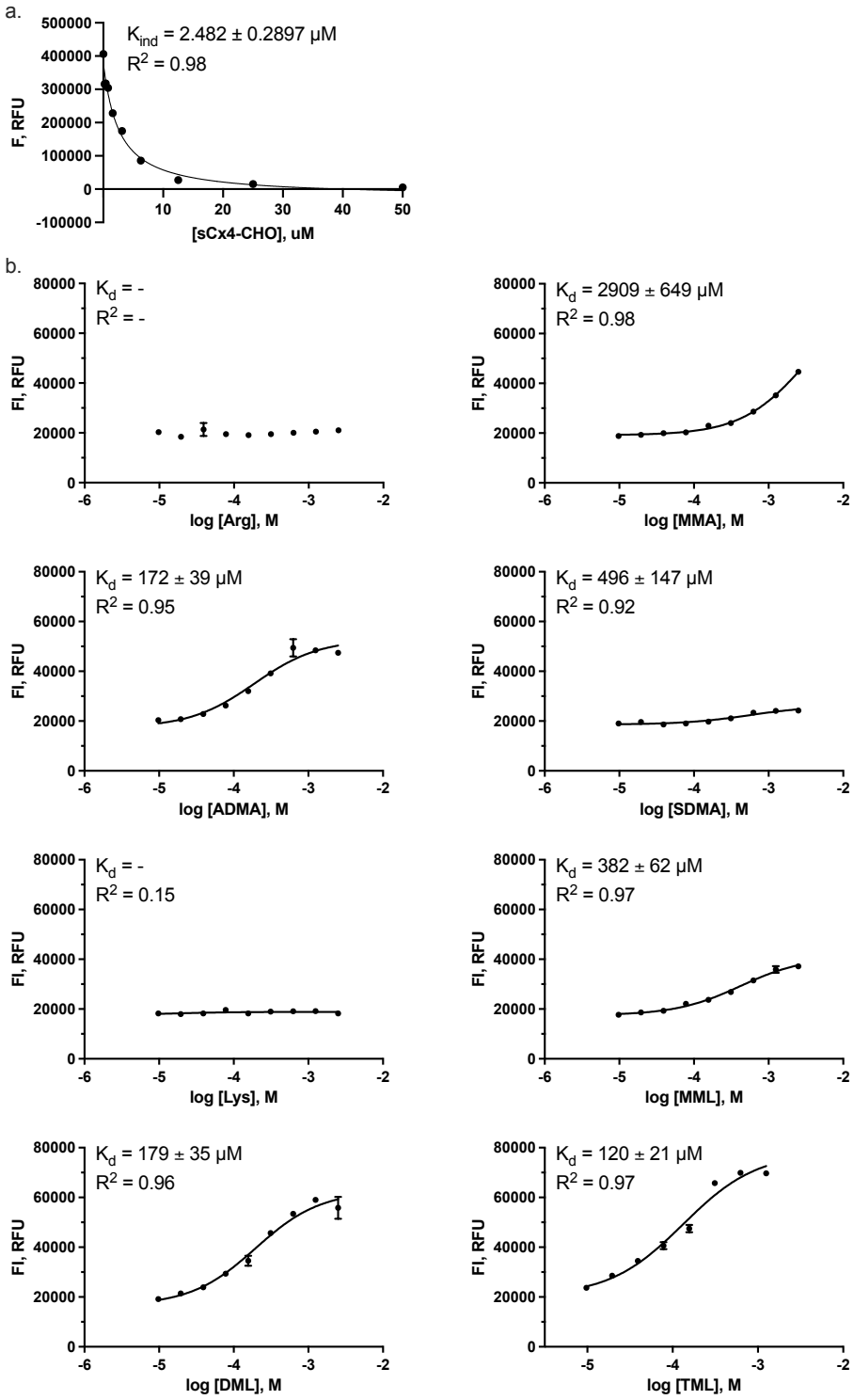


Figure S2: Fluorescence based studies of **sCx4-CHO**. a) Direct titration of LCG (0.25 µM) with **sCx4-CHO** (0 – 50 µM). b) Competitive titrations of Arginine, MMA, ADMA, SDMA, Lysine, MML, DML and TML (0 – 2.5 mM) individually titrated into the **sCx4-CHO**-LCG (5 µM **sCx4-CHO**, 0.25 µM LCG) complex.

2.2.3. Fluorescence based studies of **sCx4-NO2**

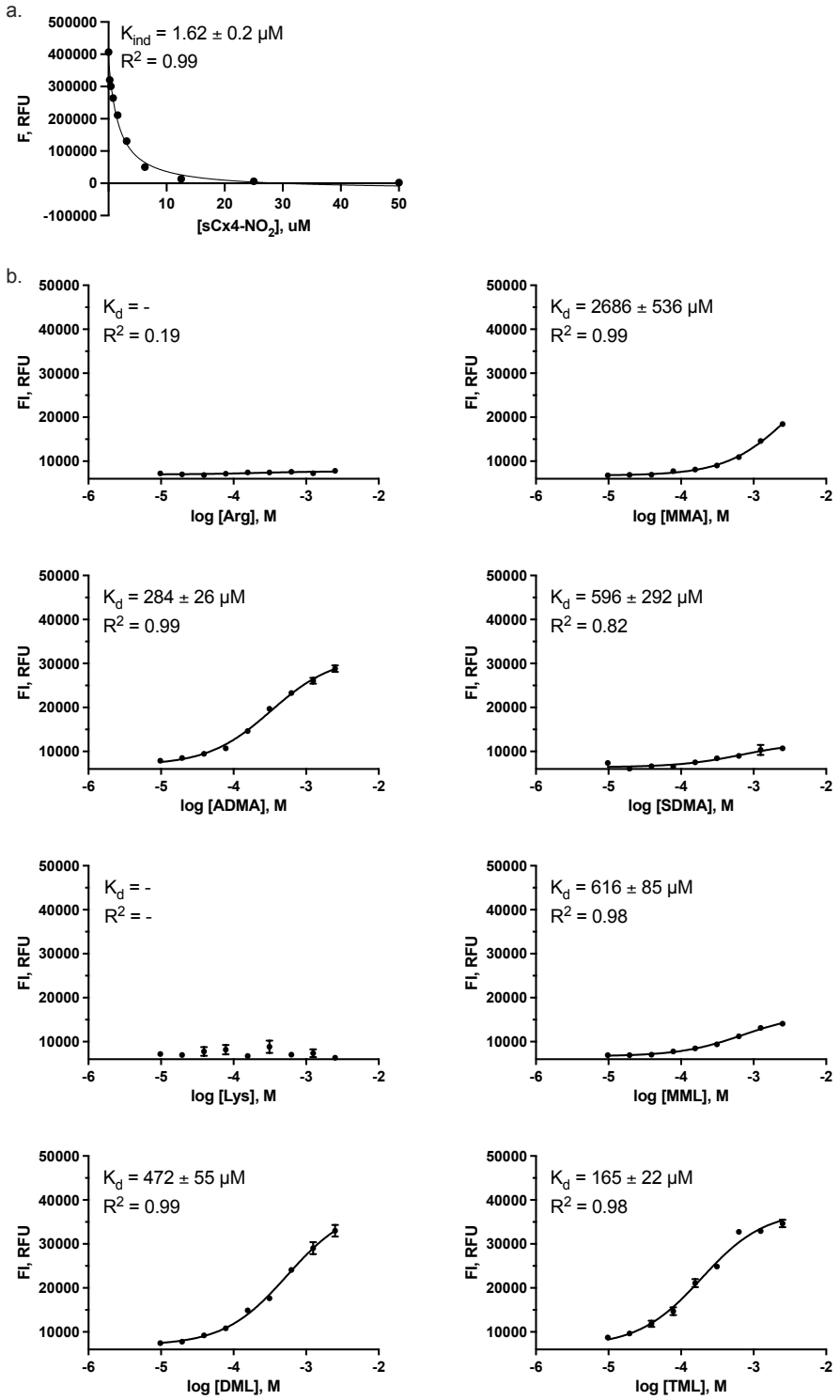


Figure S3: Fluorescence based studies of **sCx4-NO2**. a) Direct titration of LCG (0.25 µM) with **sCx4-NO2** (0 – 50 µM). b) Competitive titrations of Arginine, MMA, ADMA, SDMA, Lysine, MML, DML and TML (0 – 2.5 mM) individually titrated into the **sCx4-NO2**-LCG (5 µM **sCx4-NO2**, 0.25 µM LCG) complex.

2.2.4. Fluorescence based studies of **CLR01**

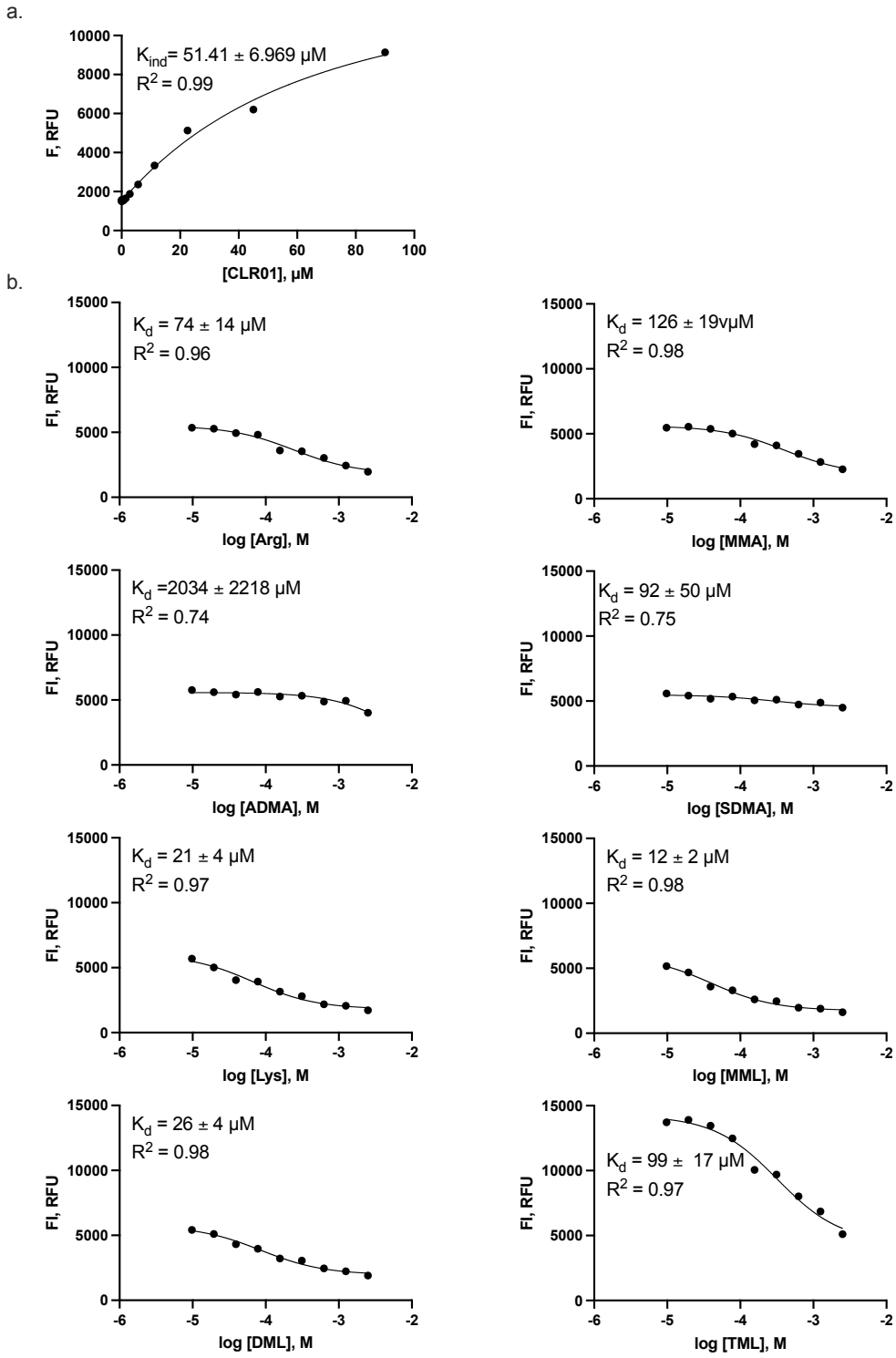


Figure S4: Fluorescence based studies of **CLR01**. a) Direct titration of 4-ASP(20 µM) with **CLR01** (0 – 50 µM). b) Competitive titrations between Arginine, MMA, ADMA, SDMA, Lysine, MML, DML and TML (0 – 2.5 mM) individually titrated into the **CLR01**-4-ASP (70 µM **CLR01**, 20 µM 4-ASP) complex.

2.2.5. Fluorescence based studies of **PC**

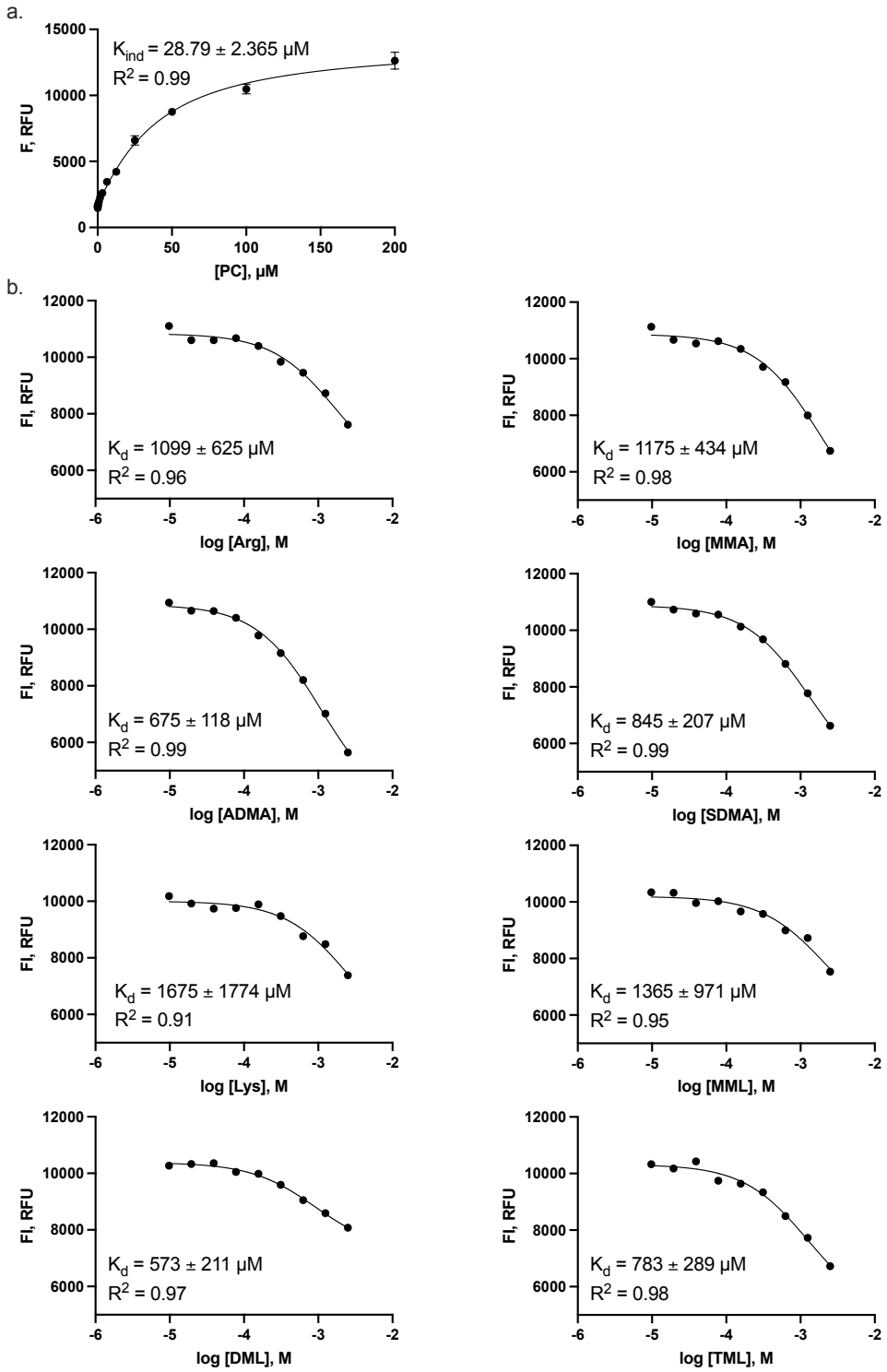


Figure S5: Fluorescence based studies of **PC**. a) Direct titration of 4-ASP (20 µM) with **PC** (0 – 200 µM). b) Competitive titrations between Arginine, MMA, ADMA, SDMA, Lysine, MML, DML and TML (0 – 2.5 mM) individually titrated into the **PC**-4-ASP (30 µM **PC**, 20 µM 4-ASP) complex.

2.2.6. Fluorescence based studies of **M1**

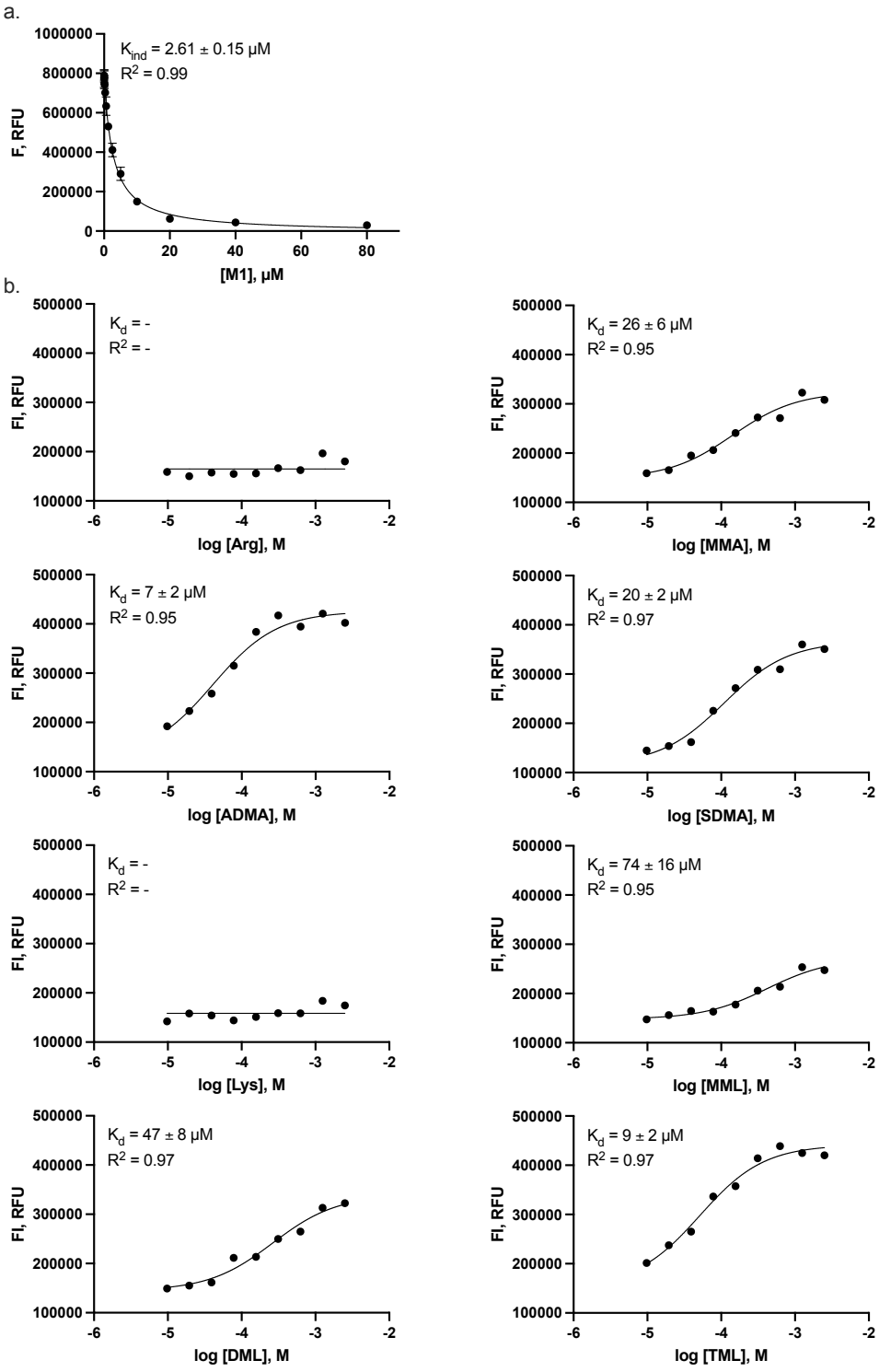


Figure S6: Fluorescence based studies of **M1**. a) Direct titration of R6G (10 µM) with **M1** (0 – 80 µM). b) Competitive titrations between Arginine, MMA, ADMA, SDMA, Lysine, MML, DML and TML (0 – 2.5 mM) individually titrated into the **M1**-R6G (10 µM **M1**, 10 µM R6G) complex.

2.2.7. Fluorescence based studies of **M2**

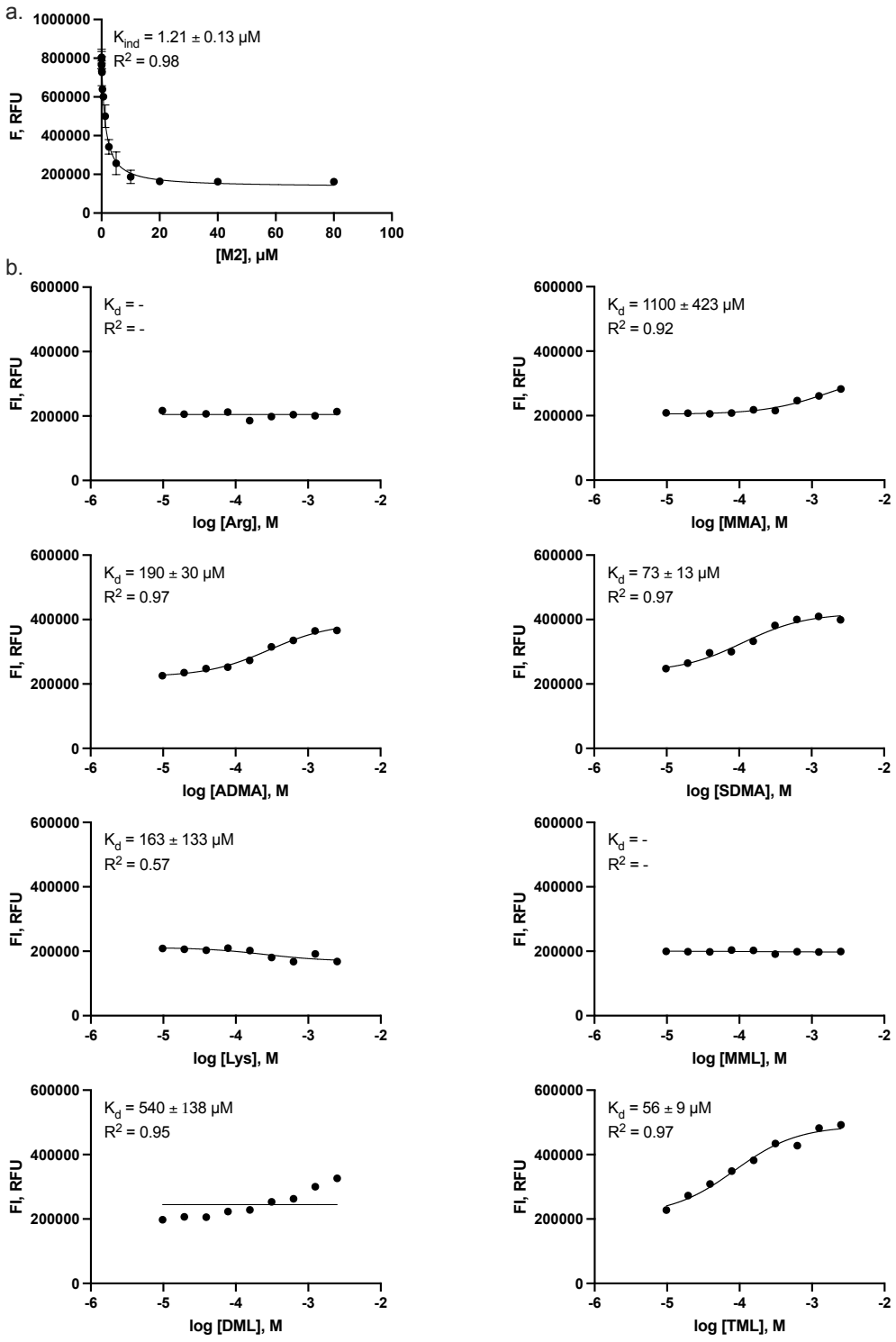


Figure S7: Fluorescence based studies of **M2**. a) Direct titration of R6G (10 µM) with **M2** (0 – 80 µM). b) Competitive titrations between Arginine, MMA, ADMA, SDMA, Lysine, MML, DML and TML (0 – 2.5 mM) individually titrated into the **M2**-R6G (10 µM **M2**, 10 µM R6G) complex.

2.3. Fluorescence based studies with control compounds

We did IDA studies dimethylamine, tetramethylammonium, and benzyltrimethylammonium (Figure S8) as controls for evaluating hydrophobicity and size/shape-matching. **M1** and **M2** displayed such strong binding to benzyltrimethylammonium that we did a second IDA with a lower concentration range (0 to 0.31 mM) to observe the binding curve.

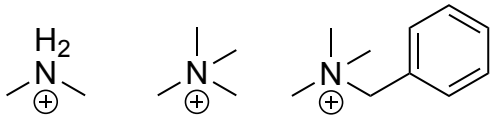


Figure S8: Control compounds for hydrophobicity and shape-matching. Left to right dimethylammonium, tetramethylammonium and benzyltrimethylammonium. Each guest was used as its chloride salt.

Table S1. *K_d values determined by IDA for complexes of control guests in 10 mM phosphate buffer*. [a]

	sCx4[b]	**sCx4-CHO**[b]	**sCx4-NO_2**[b]	**M1**[c]	**M2**[c]	**CLR01**[d]	**PC**[d]
	K_d (μM)	K_d (μM)	K_d (μM)	K_d (μM)	K_d (μM)	K_d (μM)	K_d (μM)
Dimethylamine	>1500	>1500	>1500	>1500	>1500	1100 ± 1100	>1500
Tetramethylammonium	370 ± 70	390 ± 110	660 ± 90	12 ± 5	40 ± 40[e]	460 ± 360[f]	370 ± 440[g]
Benzyltrimethylammonium	940 ± 200	340 ± 60	660 ± 180	3 ± 1	1.5 ± 0.4	170 ± 30	960 ± 560

[a] All titrations were carried out in 10 mM NaH_2PO_4 buffer at pH 7.4. See below for titration curves experimental details, and fitting details. All K_d values arise from fits with $R^2 \geq 0.95$ except where indicated. [b] Lucigenin (LCG) was used as indicator. [c] Rhodamine 6G (R6G) was used as indicator. [d] 4-(4-Diethylaminostyryl)-1-methylpyridinium iodide (4-ASP) was used as indicator. [e] $R^2 \geq 0.83$. [f] $R^2 \geq 0.92$. [g] $R^2 \geq 0.79$.

2.3.1 Fluorescence based studies of **sCx4** and control compounds

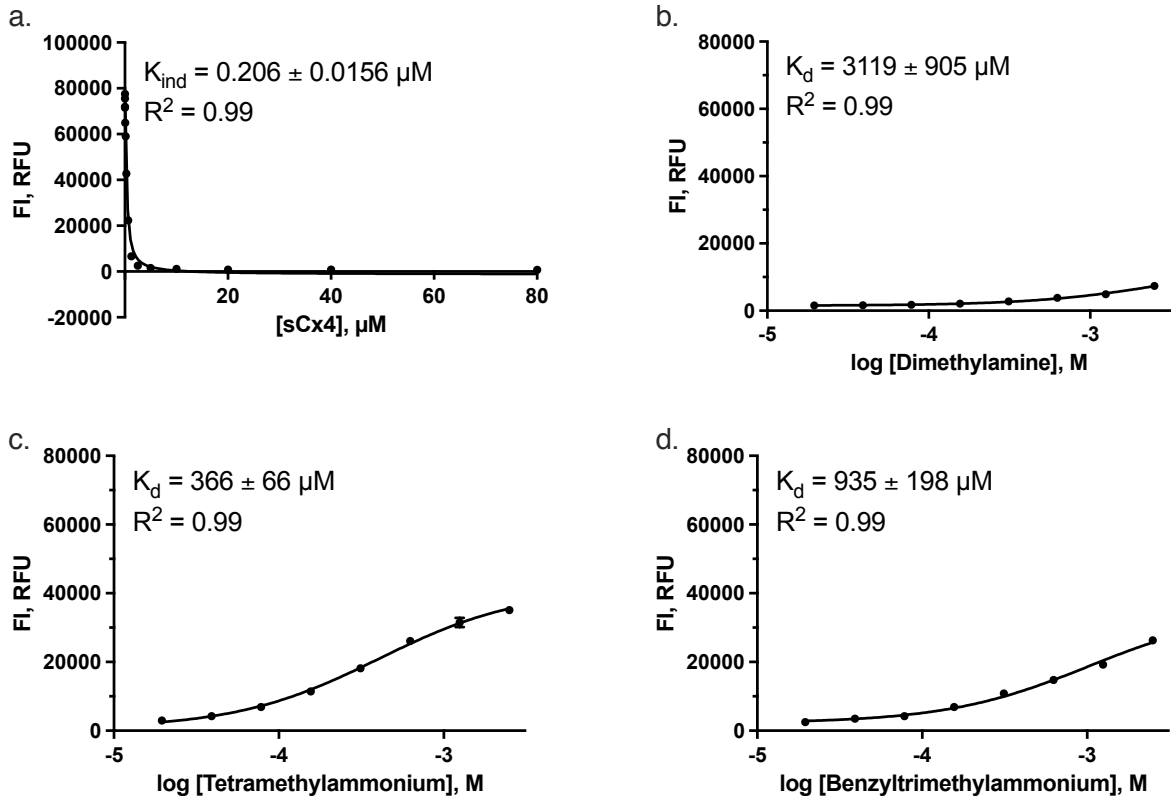


Figure S9: Fluorescence based studies of **sCx4**. a) Direct titration of LCG (0.25 µM) with **sCx4** (0 – 80 µM). b) Competitive titration with dimethylamine (0 – 2.5 mM) titrated into the **sCx4**-LCG (5 µM **sCx4**, 0.25 µM LCG) complex. c) Competitive titration with tetramethylammonium (0 – 2.5 mM) titrated into the **sCx4**-LCG (5 µM **sCx4**, 0.25 µM LCG) complex. d) Competitive titration with benzyltrimethylammonium (0 – 2.5 mM) titrated into the **sCx4**-LCG (5 µM **sCx4**, 0.25 µM LCG) complex. See Table S1 for buffer conditions.

2.3.2 Fluorescence based studies of **sCx4-CHO** and control compounds

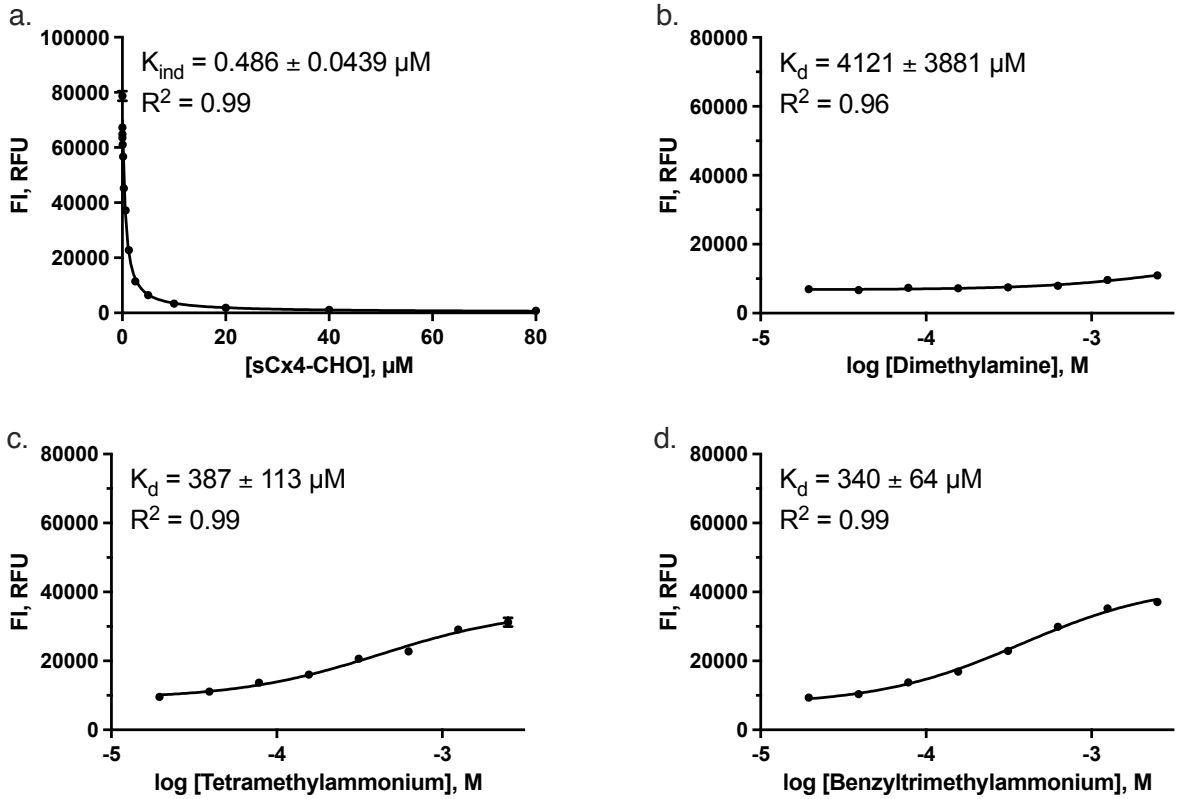


Figure S10: Fluorescence based studies of **sCx4-CHO**. a) Direct titration of LCG (0.25 µM) with **sCx4-CHO** (0 – 80 µM). b) Competitive titration with dimethylamine (0 – 2.5 mM) titrated into the **sCx4-CHO**-LCG (5 µM **sCx4**, 0.25 µM LCG) complex. c) Competitive titration with tetramethylammonium (0 – 2.5 mM) titrated into the **sCx4-CHO**-LCG (5 µM **sCx4-CHO**, 0.25 µM LCG) complex. d) Competitive titration with benzyltrimethylammonium (0 – 2.5 mM) titrated into the **sCx4-CHO**-LCG (5 µM **sCx4-CHO**, 0.25 µM LCG) complex. See Table S1 for buffer conditions.

2.3.3 Fluorescence based studies of **sCx4-NO2** and control compounds

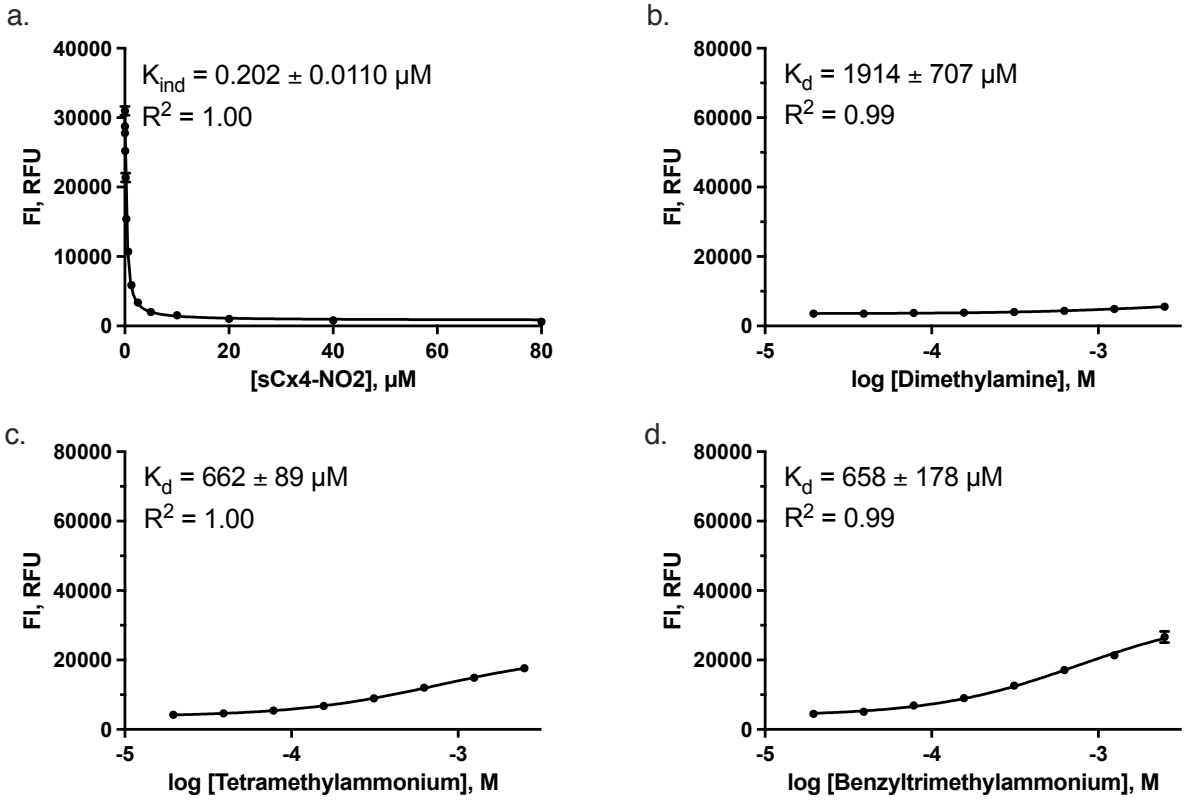


Figure S11: Fluorescence based studies of **sCx4-NO2**. a) Direct titration of LCG (0.25 µM) with **sCx4-NO2** (0 – 80 µM). b) Competitive titration with dimethylamine (0 – 2.5 mM) titrated into the **sCx4-NO2**-LCG (5 µM **sCx4-NO2**, 0.25 µM LCG) complex. c) Competitive titration with tetramethylammonium (0 – 2.5 mM) titrated into the **sCx4-NO2**-LCG (5 µM **sCx4-NO2**, 0.25 µM LCG) complex. d) Competitive titration with benzyltrimethylammonium (0 – 2.5 mM) titrated into the **sCx4-NO2**-LCG (5 µM **sCx4-NO2**, 0.25 µM LCG) complex. See Table S1 for buffer conditions.

2.3.4 Fluorescence based studies of **M1** and control compounds

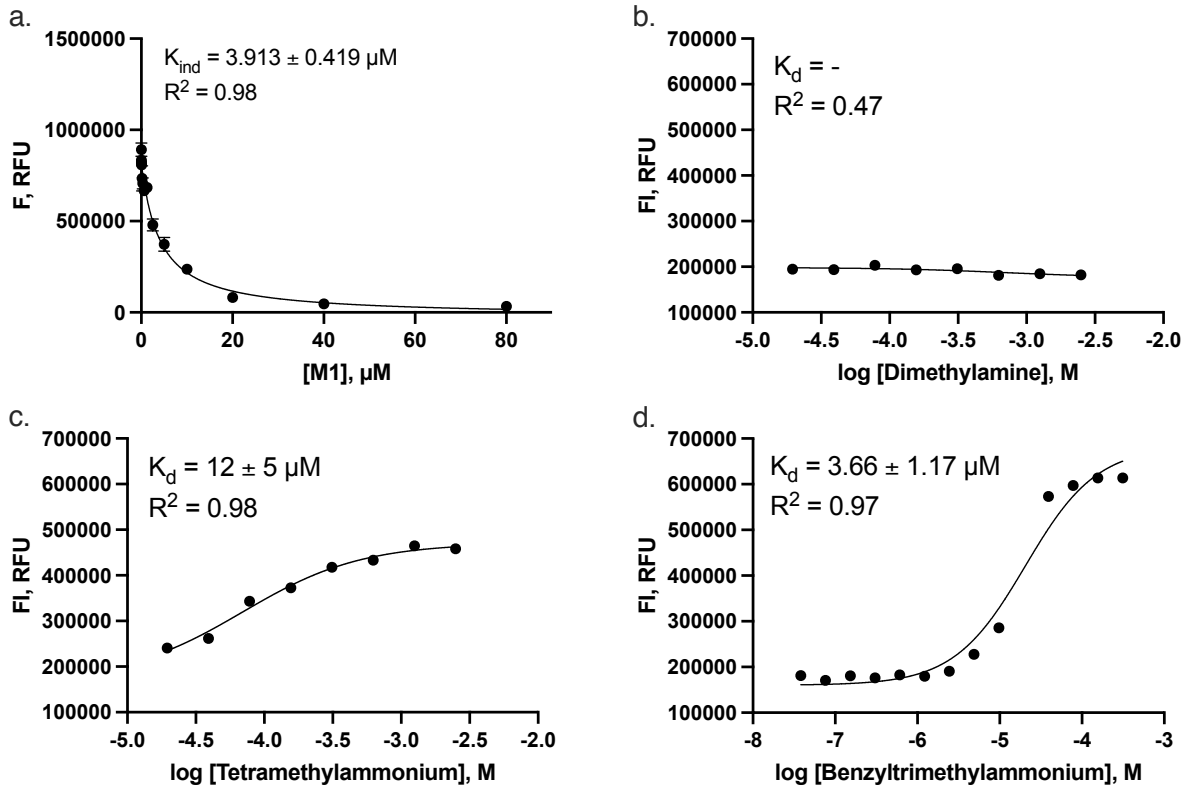


Figure S12: Fluorescence based studies of **M1**. a) Direct titration of R6G (10 µM) with **M1** (0 – 80 µM). b) Competitive titration with dimethylamine (0 – 2.5 mM) titrated into the **M1**-R6G (10 µM **M1**, 10 µM R6G) complex. c) Competitive titration with tetramethylammonium (0 – 2.5 mM) titrated into the **M1**-R6G (10 µM **M1**, 10 µM R6G) complex. d) Competitive titration with benzyltrimethylammonium (0 – 0.31 mM) titrated into the **M1**-R6G (10 µM **M1**, 10 µM R6G) complex. See Table S1 for buffer conditions.

2.3.5 Fluorescence based studies of **M2** and control compounds

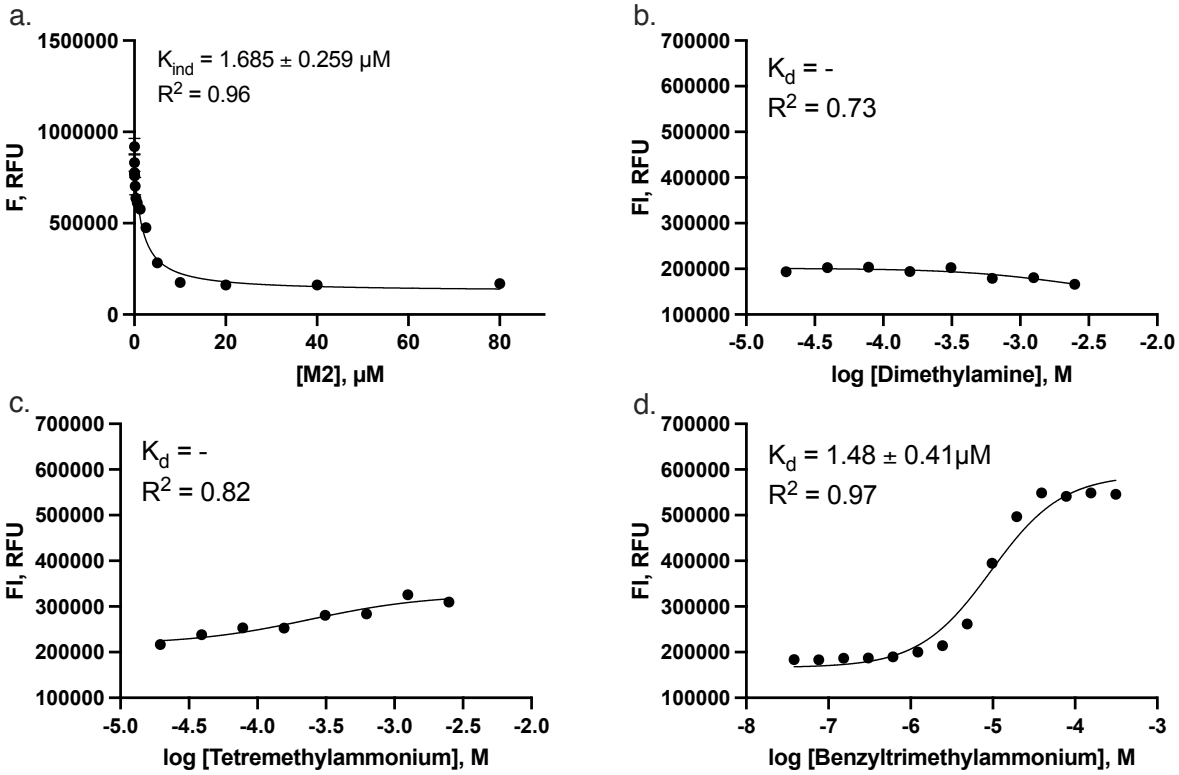


Figure S13: Fluorescence based studies of **M2**. a) Direct titration of R6G (10 µM) with **M2** (0 – 80 µM). b) Competitive titration with dimethylamine (0 – 2.5 mM) titrated into the **M2**-R6G (10 µM **M2**, 10 µM R6G) complex. c) Competitive titration with tetramethylammonium (0 – 2.5 mM) titrated into the **M2**-R6G (10 µM **M2**, 10 µM R6G) complex. d) Competitive titration with benzyltrimethylammonium (0 – 0.31 mM) titrated into the **M2**-R6G (10 µM **M2**, 10 µM R6G) complex. See Table S1 for buffer conditions.

2.3.6 Fluorescence based studies of **PC** and control compounds

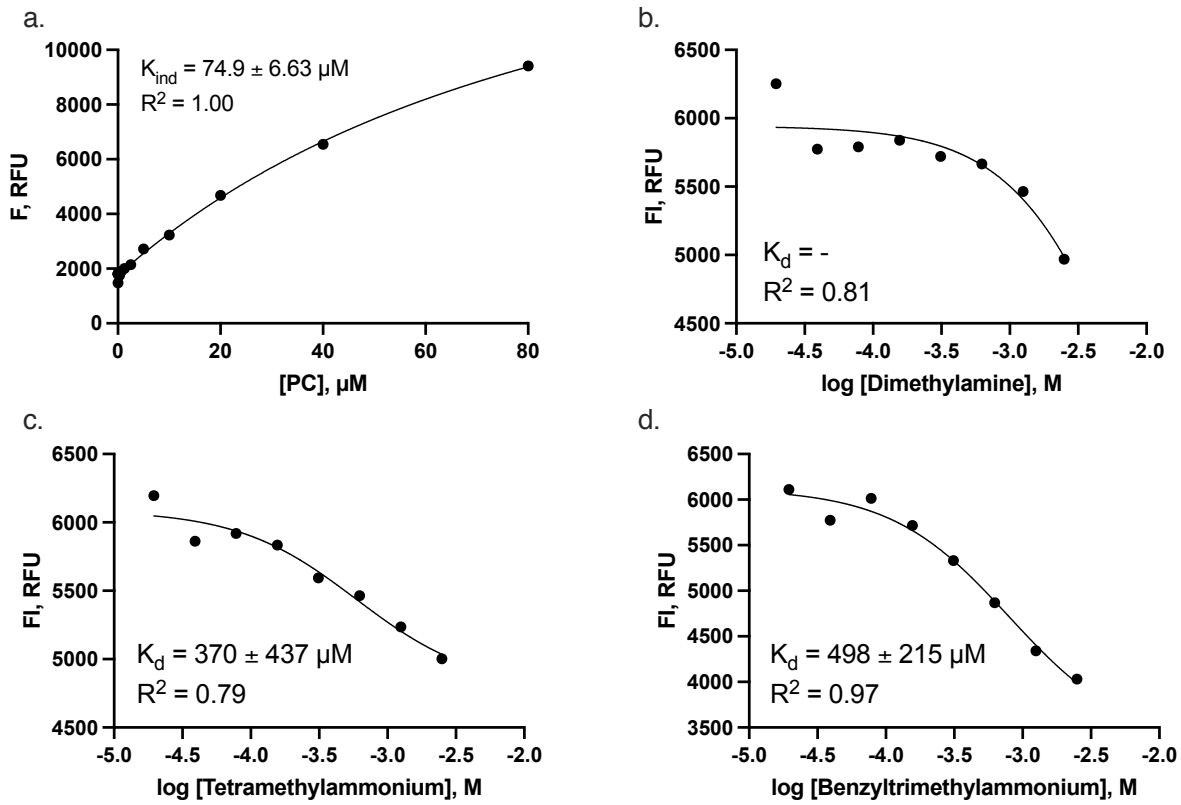


Figure S14: Fluorescence based studies of **PC**. a) Direct titration of 4-ASP (20 µM) with **PC** (0 – 50 µM). b) Competitive titrations with dimethylamine (0 – 2.5 mM) titrated into the **PC**-4-ASP (30 µM **PC**, 20 µM 4-ASP) complex. c) Competitive titrations with tetramethylammonium (0 – 2.5 mM) titrated into the **PC**-4-ASP (30 µM **PC**, 20 µM 4-ASP) complex.d) Competitive titrations with benzyltrimethylammonium (0 – 2.5 mM) titrated into the **PC**-4-ASP (30 µM **PC**, 20 µM 4-ASP) complex. See Table S1 for buffer conditions.

2.3.7 Fluorescence based studies of **CLR01** and control compounds

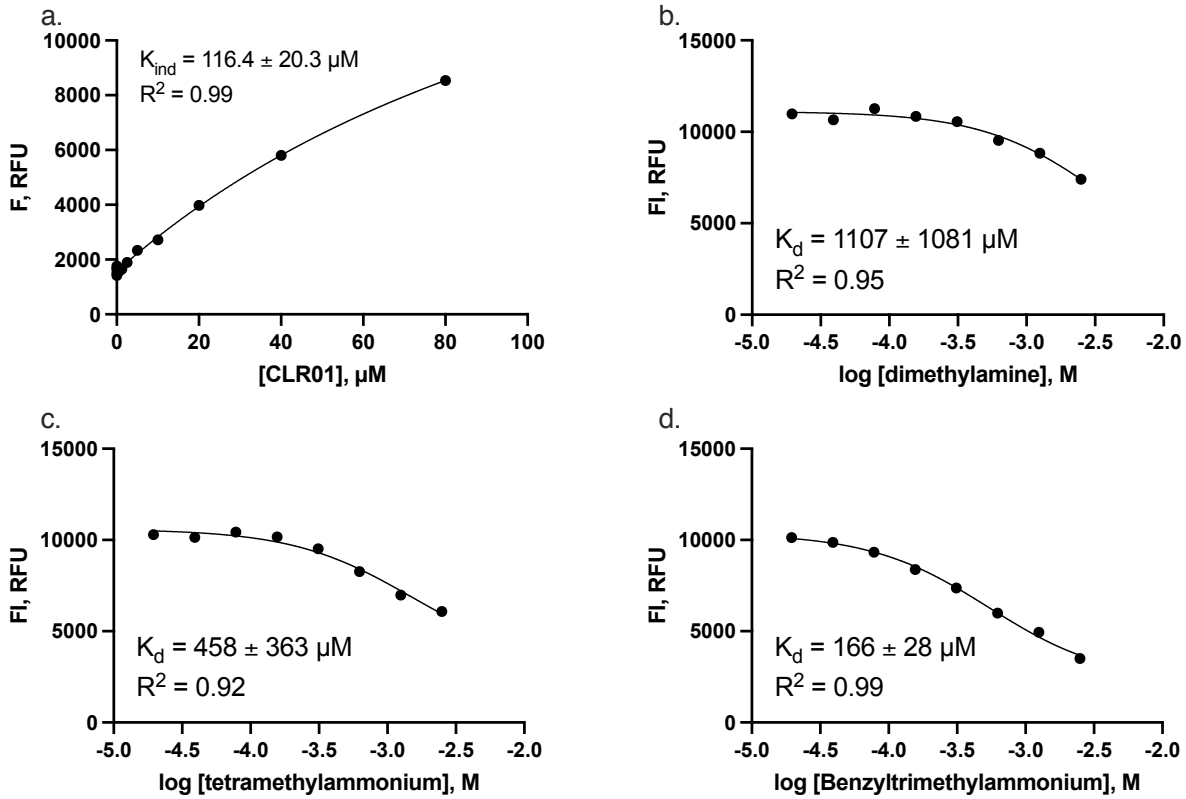


Figure S15: Fluorescence based studies of **CLR01**. a) Direct titration of 4-ASP (20 µM) with **CLR01** (0 – 50 µM). b) Competitive titrations with dimethylamine (0 – 2.5 mM) titrated into the **CLR01**-4-ASP (70 µM **CLR01**, 20 µM 4-ASP) complex. c) Competitive titrations with tetramethylammonium (0 – 2.5 mM) titrated into the **CLR01**-4-ASP (70 µM **CLR01**, 20 µM 4-ASP) complex.d) Competitive titrations with benzyltrimethylammonium (0 – 2.5 mM) titrated into the **CLR01**-4-ASP (70 µM **CLR01**, 20 µM 4-ASP) complex. See Table S1 for buffer conditions.

2.3.8 3D bar graph visualizing binding trends between hosts and dimethylamine, tetramethylammonium and benzyltrimethylammonium

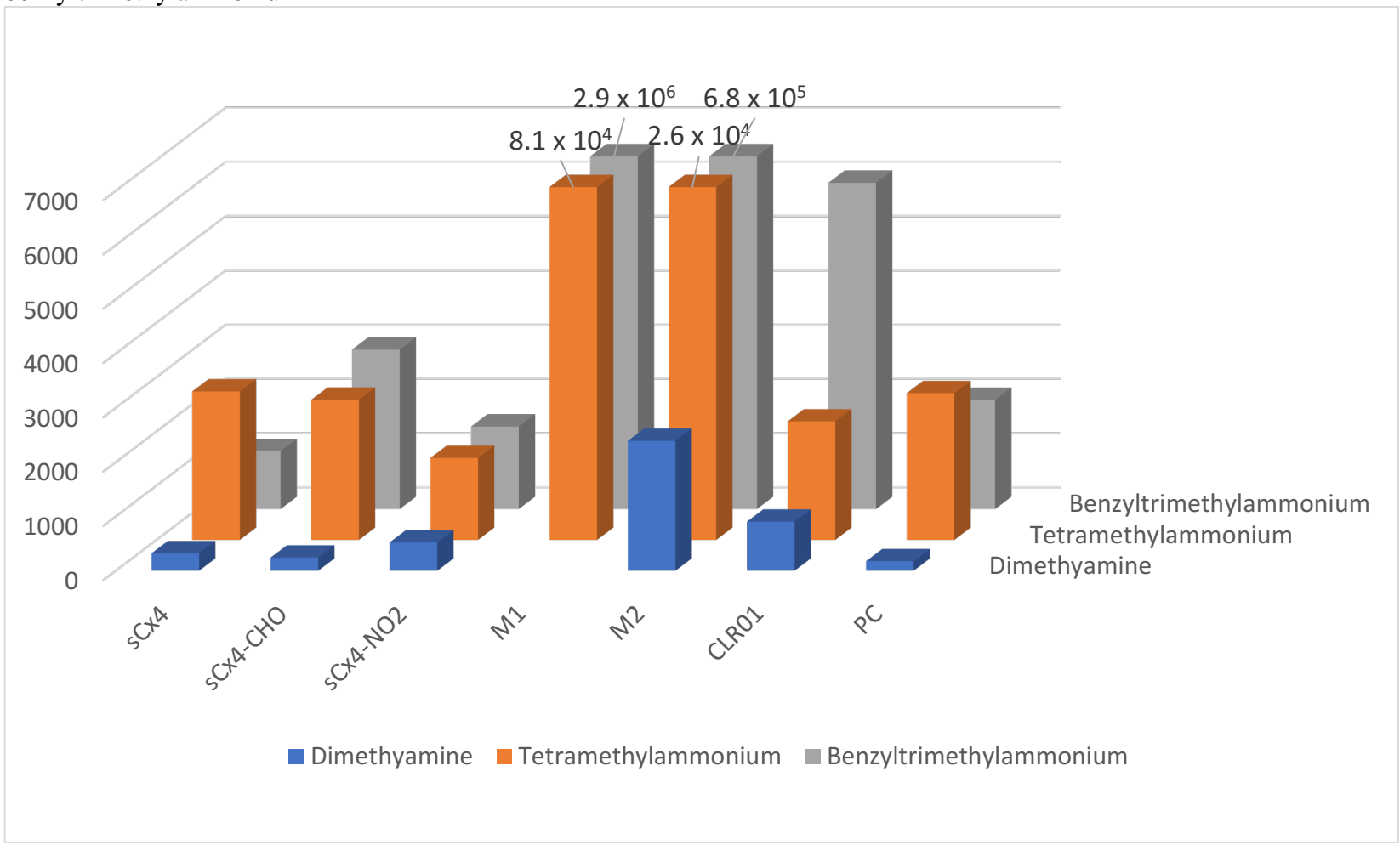


Figure S16: 3D bar graph visualizing binding trends between the hosts and dimethylamine, tetramethylammonium and benzyltrimethylammonium. The binding strengths are presented as the K_a values in units M^{-1} ($K_a=1/K_d$). The bar for **M1** and **M2** with tetramethylammonium and benzyltrimethylammonium are truncated, with the actual K_a values presented above the bars.

2.4. Fluorescence based studies 50 mM buffer

2.4.1. Fluorescence based studies of **M1**

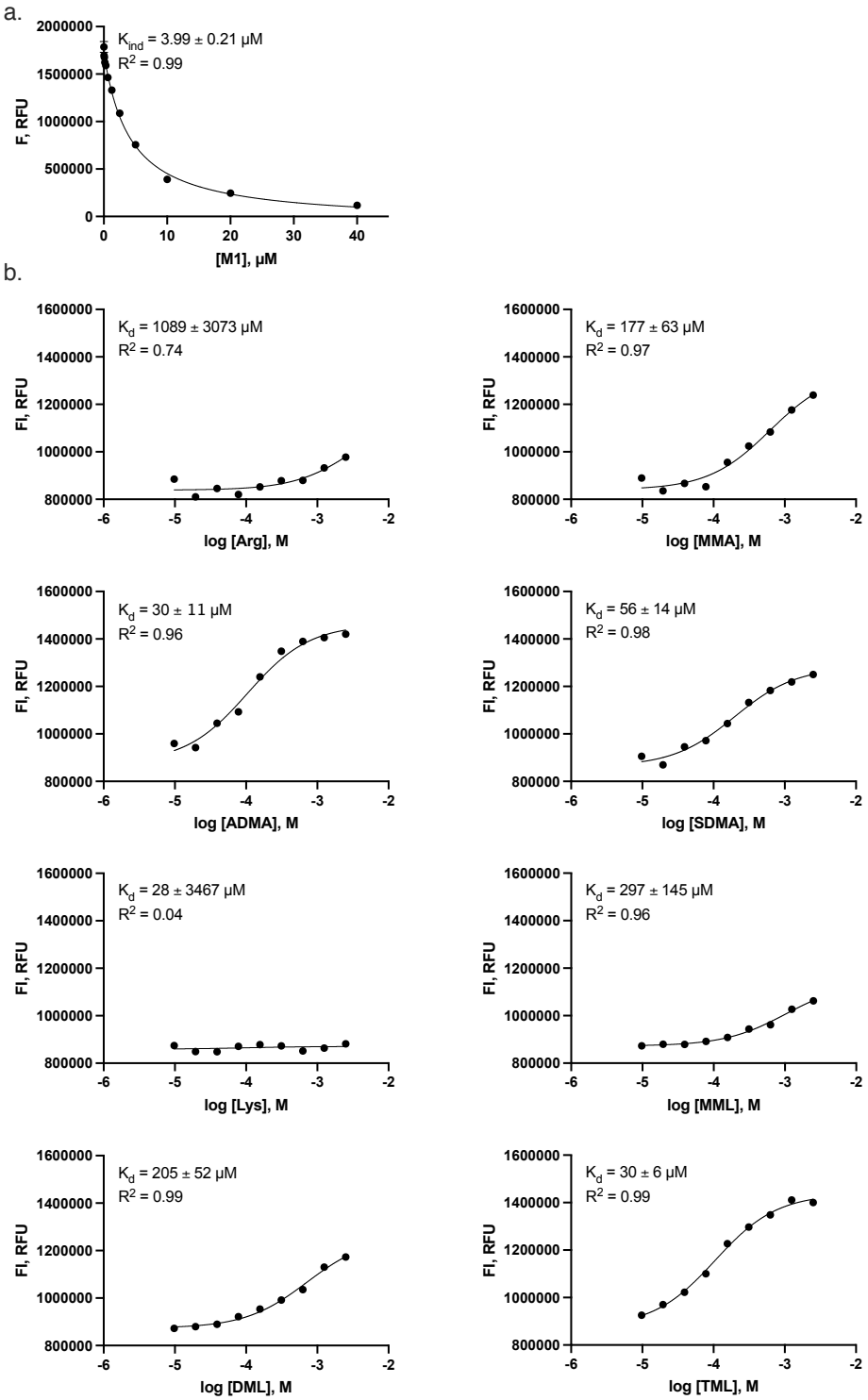


Figure S17: Fluorescence based studies of **M1**. a) Direct titration of R6G (10 µM) with **M1** (0 – 40 µM). b) Competitive titrations between Arginine, MMA, ADMA, SDMA, Lysine, MML, DML and TML (0 – 2.5 mM) individually titrated into the **M1**-R6G (10 µM **M1**, 10 µM R6G) complex.

2.4.2. Fluorescence based studies of **M2**

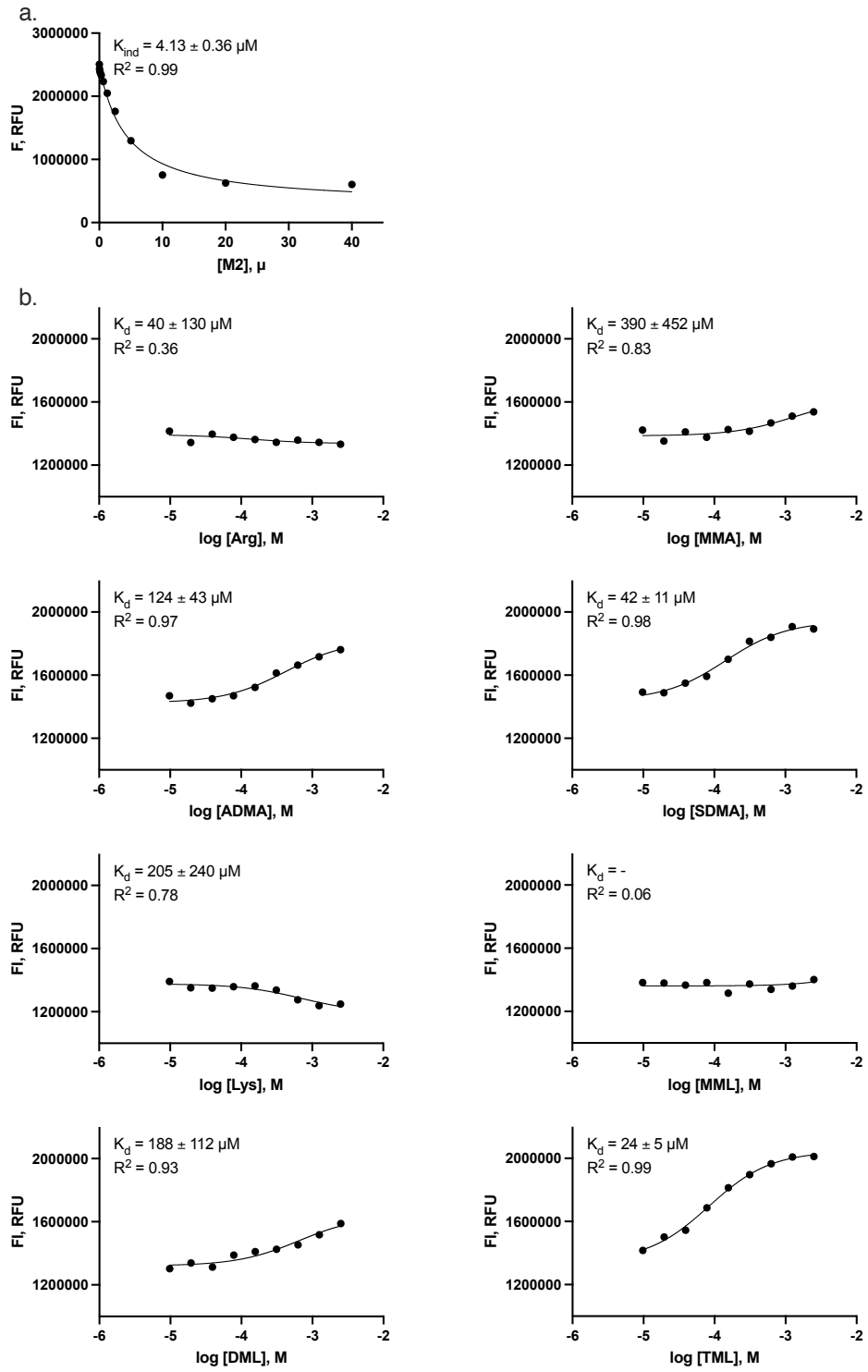


Figure S18: Fluorescence based studies of **M2**. a) Direct titration of R6G (10 µM) with **M2** (0 – 40 µM). b) Competitive titrations between Arginine, MMA, ADMA, SDMA, Lysine, MML, DML and TML (0 – 2.5 mM) individually titrated into the **M2**-R6G (10 µM **M2**, 10 µM R6G) complex.

2.4.3. Fluorescence based studies of **CLR01**

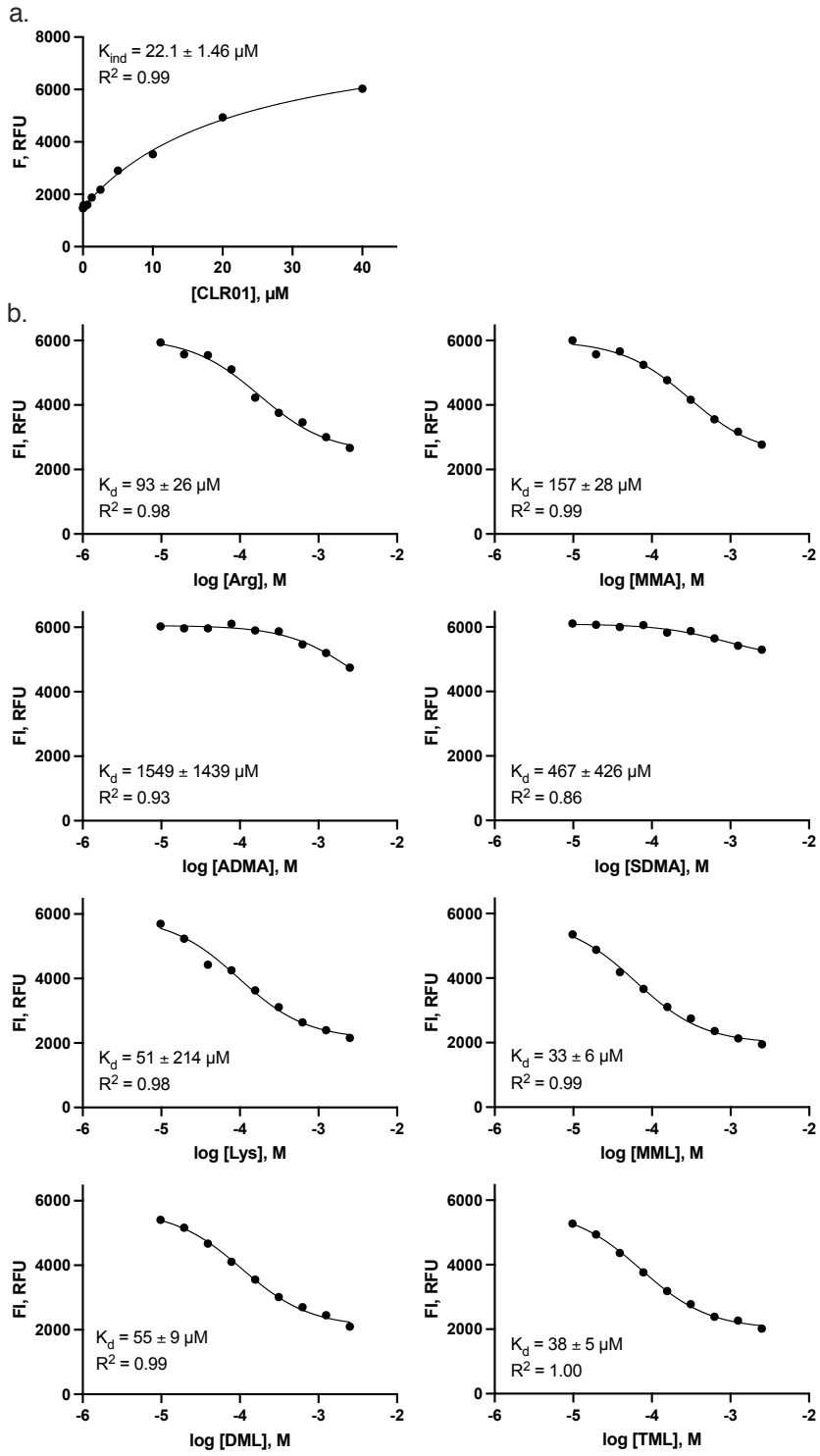


Figure S19: Fluorescence based studies of **CLR01**. a) Direct titration of 4-ASP(20 µM) with **CLR01** (0 – 40 µM). b) Competitive titrations between Arginine, MMA, ADMA, SDMA, Lysine, MML, DML and TML (0 – 2.5 mM) individually titrated into the **CLR01**-4-ASP (70 µM **CLR01**, 20 µM 4-ASP) complex.

3. Isothermal Titration Calorimetry — literature data

Table S2: Summary of previously published ITC data on relevant complexes.

		K_d μM	ΔH kJ/mol	TΔS kJ/mol
sCx4[6]a	Lysine	n.d.	n.d.	n.d.
	MML	335	–16	3.5
	DML	95	–20	3.4
	TML	30	–22	4.3
M1[7]b	H-Lys-COO^-	3	–49	34
	H-Arg-COO^-	6.6	–16	–2
M2[7]b	H-Lys-COO^-	2010	–9	–6
	H-Arg-COO^-	420	–41	22
CLR01[8]c	Ac-Lys-OMe	0.15	–23	–4
	Ac-Arg-OMe	0.34	–29	4

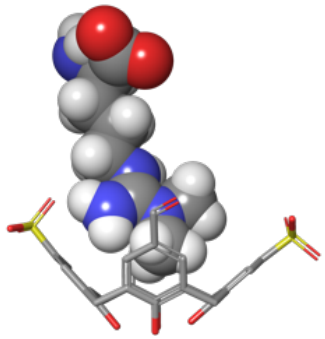
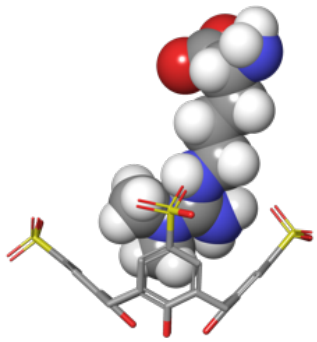
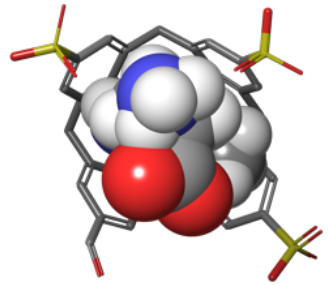
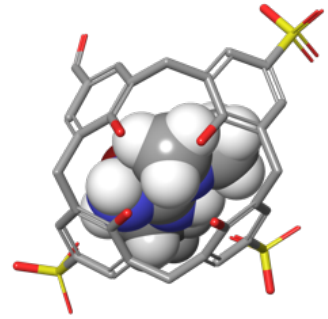
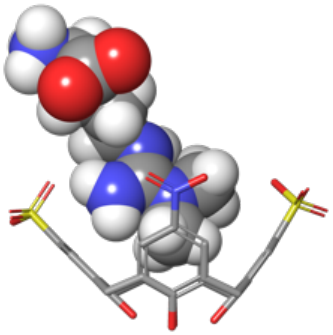
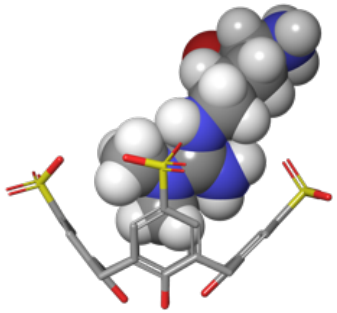
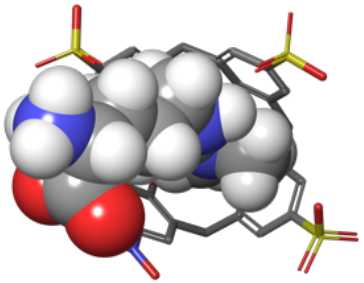
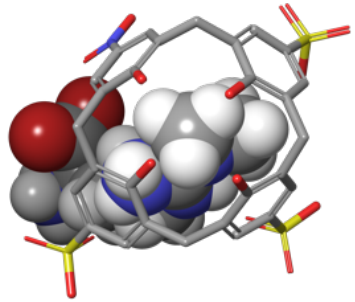
[a] ITC were carried out in 40 mM sodium phosphate buffer (pH 7.4). [b] ITC were carried out in 20 mM sodium phosphate buffer (pH 7.4). [c] ITC were carried out in 10 mM sodium phosphate buffer (pH 7.6).

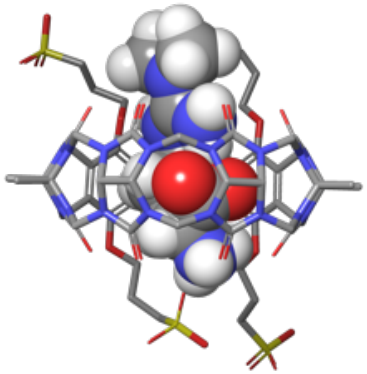
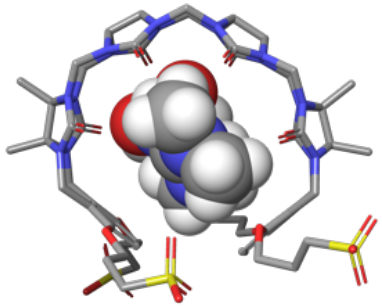
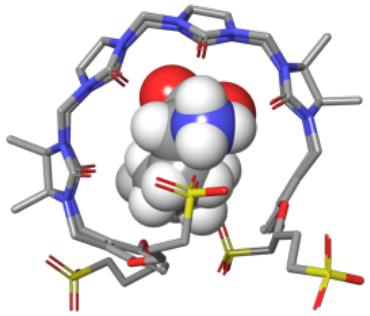
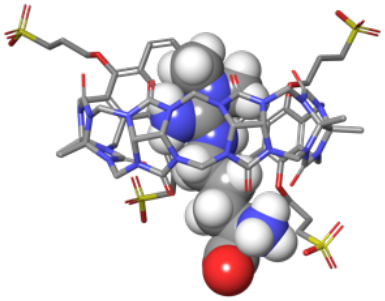
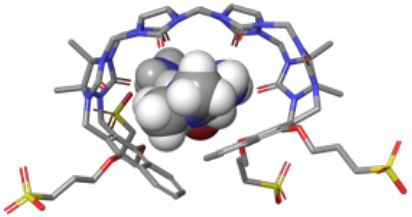
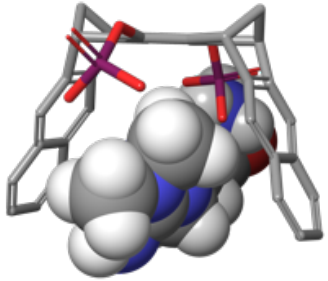
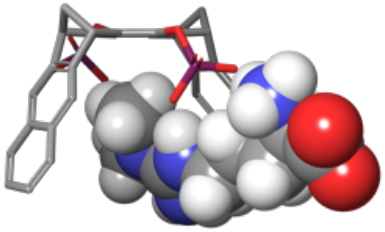
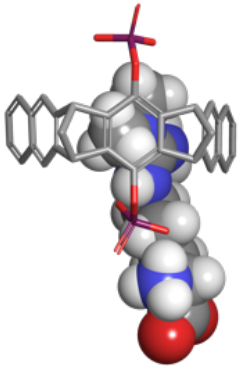
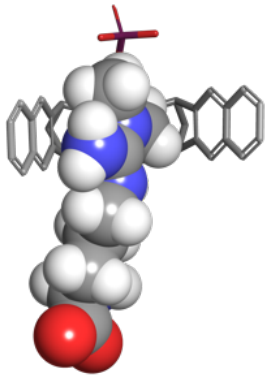
4. Molecular Modeling

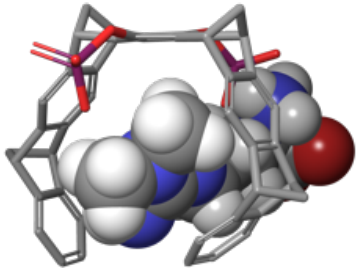
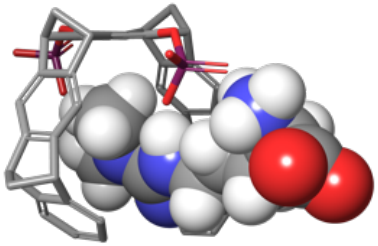
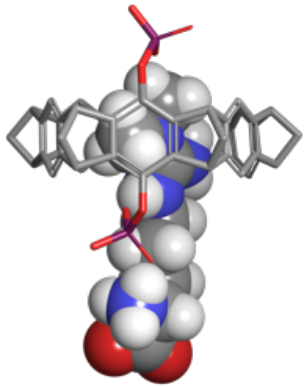
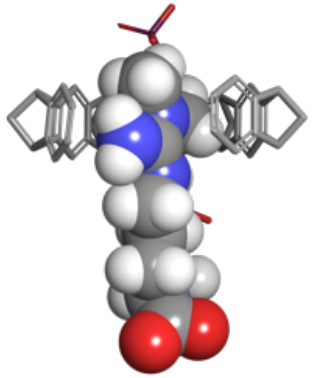
Modeling was done using minimization in explicit water for all the complexes (Force Field: OPLS_2005), in Maestro. Hosts are shown without hydrogen atoms and guests are shown with H atoms. Color Code for atoms: C = gray, N = blue, H = white, O = red, P = purple, S = yellow. Each host is modeled with ADMA (**sCx4-CHO**, **sCx4-NO2**, **M1**, **M2**, **PC**, **CLR01**). Hosts **M1**, **M2**, **PC**, **CLR01** are also modeled with MML and with TML.

4.1. Each host modeled with ADMA. Front, back, top, and bottom views for all complexes.

Table S3: Energy-minimized complexes of hosts with ADMA (OPLS_2005, explicit water)

	Front	Back	Top	Bottom
sCx4-CHO				
sCx4-NO2				

M1				
M2				
PC				

CLR01				

4.2. Hosts **M1**, **M2**, **PC**, **CLR01** each modeled with MML and with TML.

Table S4: Energy-minimized complexes of host **M1** with MML and TML (OPLS_2005, explicit water)

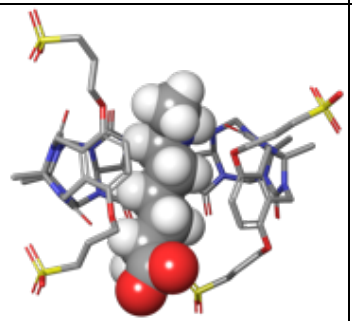
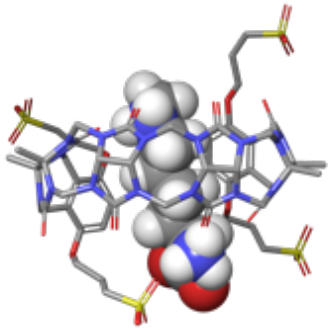
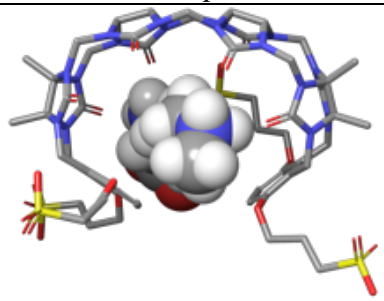
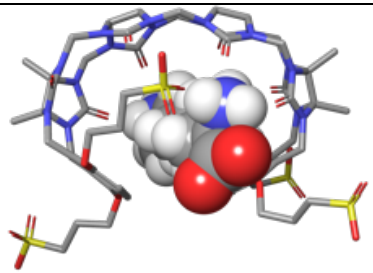
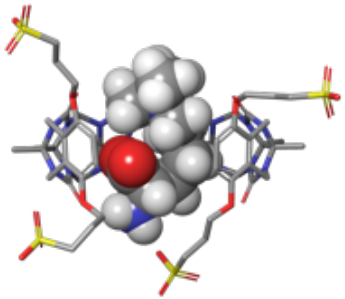
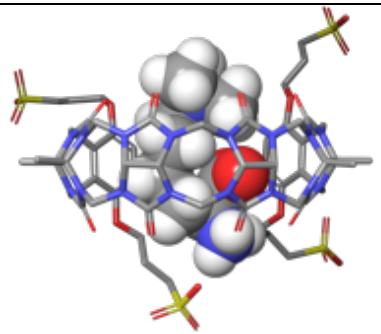
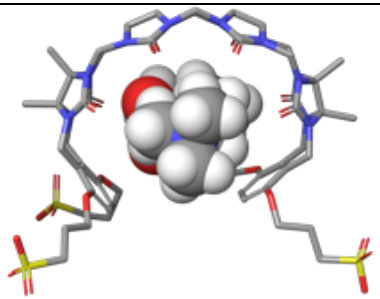
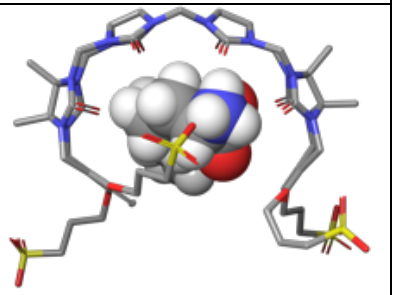
	Front	Back	Top	Bottom
MML				
TML				

Table S5: Energy-minimized complexes of host **M2** with MML and TML (OPLS_2005, explicit water)

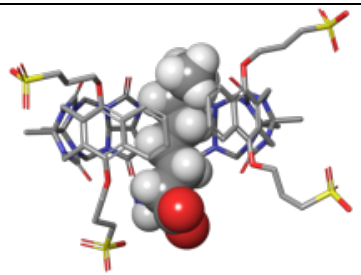
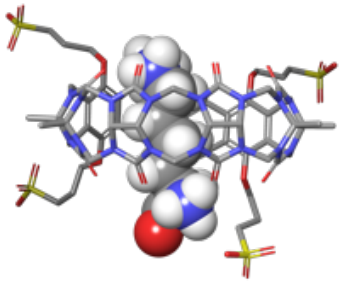
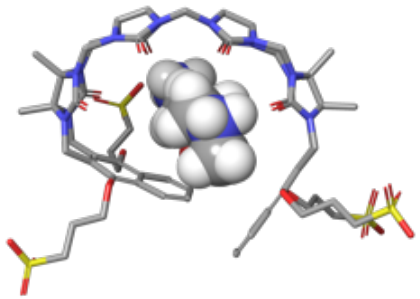
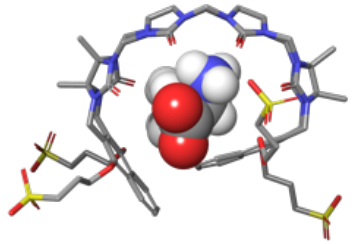
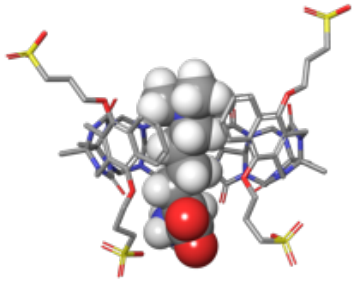
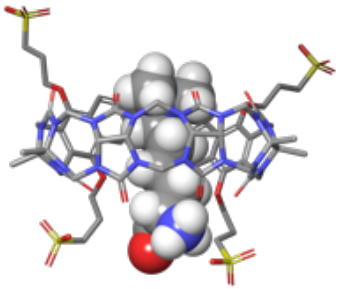
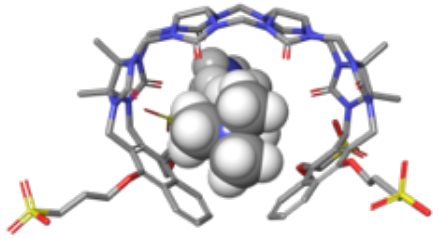
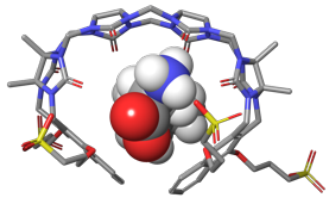
	Front	Back	Top	Bottom
MML				
TML				

Table S6: Energy-minimized complexes of host **PC** with MML and TML (OPLS_2005, explicit water)

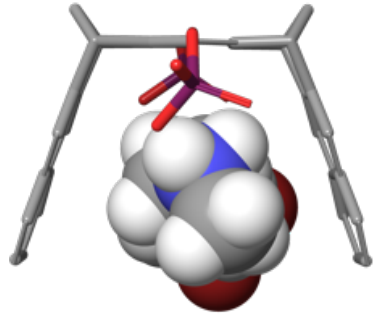
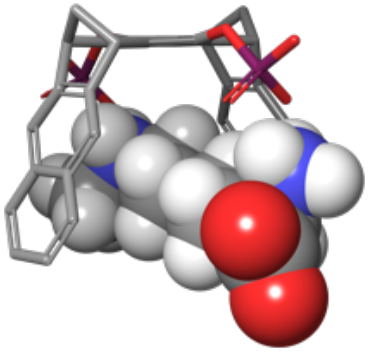
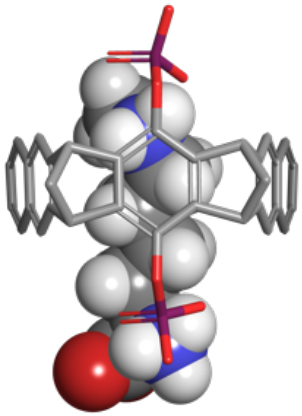
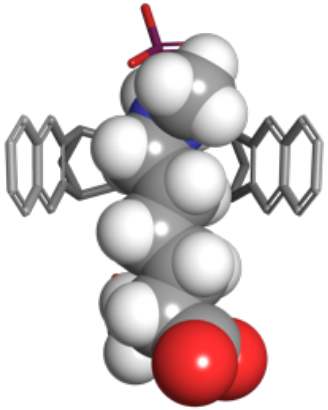
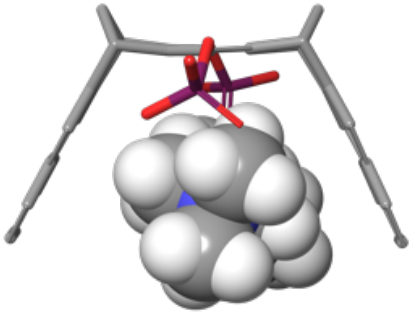
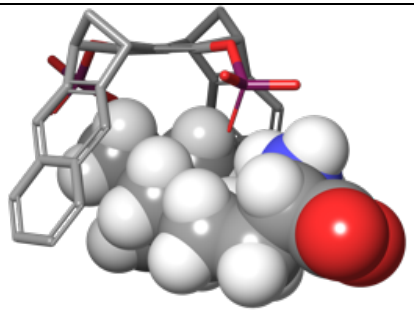
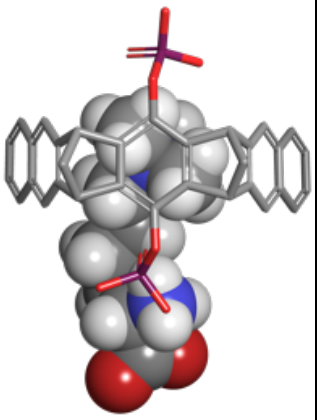
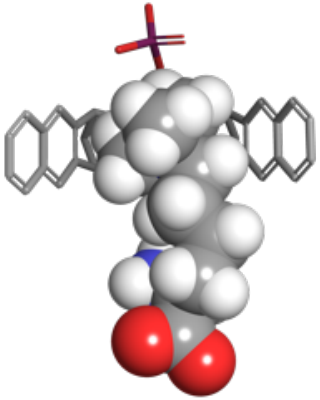
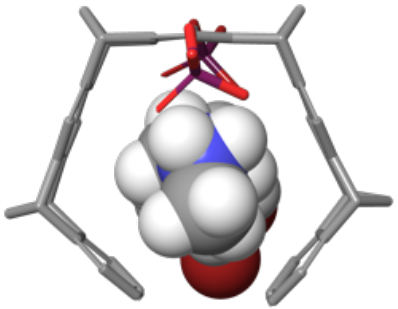
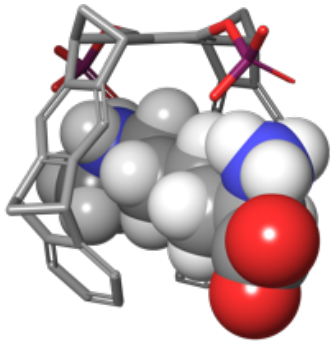
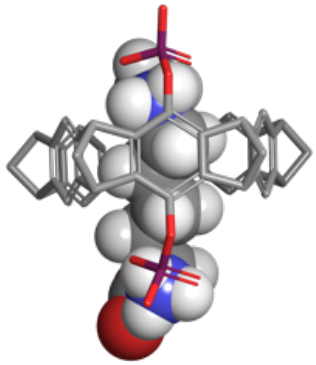
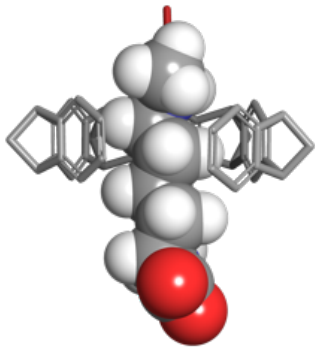
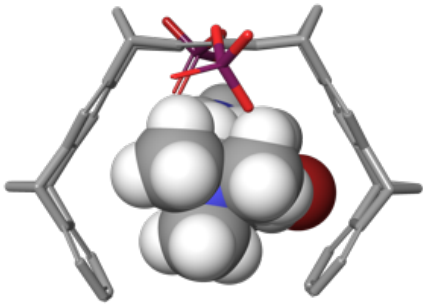
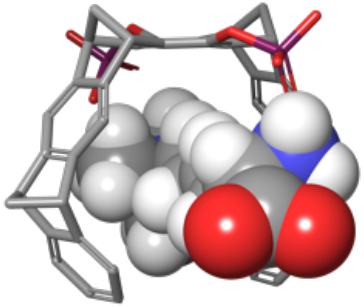
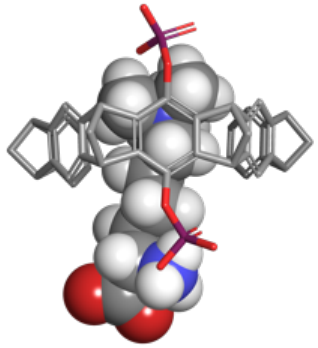
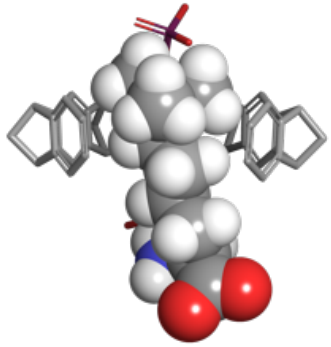
	Front	Back	Top	Bottom
MML				
TML				

Table S7: Energy-minimized complexes of host **CLR01** with MML and TML (OPLS_2005, explicit water)

	Front	Back	Top	Bottom
MML				
TML				

5. References

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

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

[3] D. Ma, G. Hettiarachchi, D. Nguyen, B. Zhang, J. B. Wittenberg, P. Y. Zavalij, V. Briken, L. Isaacs, *Nat. Chem.* **2012**, *4*, 503-510.

[4] I. Hadrovic, P. Rebmann, F. G. Klarner, G. Bitan, T. Schrader, *Front. Chem.* **2019**, *7*, 657.

[5] P. Talbiersky, F. Bastkowski, F. G. Klarner, T. Schrader, *J. Am. Chem. Soc.* **2008**, *130*, 9824-9828.

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

[7] S. A. Z. Ndendjio, L. Isaacs, *Supramol. Chem.* **2019**, *31*, 432-442.

[8] S. Dutt, C. Wilch, T. Gersthagen, P. Talbiersky, K. Bravo-Rodriguez, M. Hanni, E. Sanchez-Garcia, C. Ochsenfeld, F. G. Klarner, T. Schrader, *J. Org. Chem.* **2013**, *78*, 6721-6734.