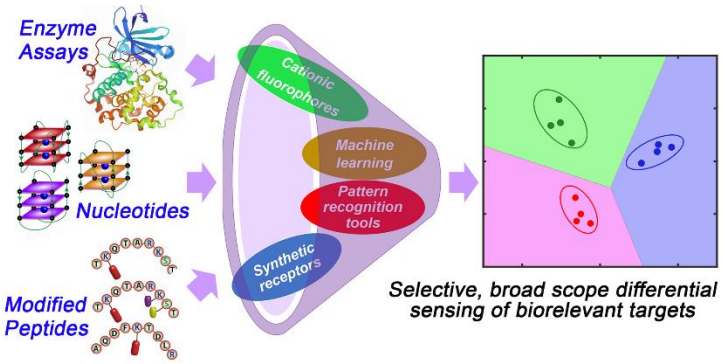
Combining Excellent Selectivity with Broad Target Scope: Biosensing with Arrayed Deep Cavitand Hosts

Wenwan Zhong[1,2] and Richard J. Hooley[1,3]

[1]Department of Chemistry; [2]Environmental Toxicology Graduate Program; [3]Department of Biochemistry; University of California–Riverside, Riverside, CA 92521, U.S.A.

CONSPECTUS



Simple macrocyclic water-soluble hosts such as cucurbiturils, cyclophanes and calixarenes have long been used for biosensing *via* indicator displacement assays. Using multiple hosts and dyes in an arrayed format allows pattern recognition-based "chemical nose" sensing, which confers exquisite selectivity, even rivaling the abilities of biological recognition tools such as antibodies. However, a challenge in indicator displacement-based biosensing with macrocyclic hosts is that selectivity and scope are often inversely correlated: strong selectivity for a specific target can limit wide application, and broad scope sensing can suffer from a lack of selectivity between similar targets. This problem can be addressed by using water-soluble, self-folding deep cavitands as hosts. These flexible bowl-shaped receptors can be easily functionalized with different motifs at the upper and lower rim, and the large cavities can bind many different fluorescent dyes, causing either fluorescence enhancement or quenching upon binding. Cavity-based affinity is strongest for NMe_3^+ groups such as trimethyllysine, and we have exploited this for site-selective recognition of post-translational lysine methylations in oligopeptides. The host recognizes the NMe_3^+ group, and by applying differently functionalized hosts in an arrayed format, discrimination between identical modifications at different positions on the oligopeptide is possible. Multiple recognition elements can be exploited for selectivity, including a defined, yet "breathable" cavity, and variable upper rim functions oriented towards the target. While the performance of the host:guest sensing system is impressive for lysine methylations, the most important advance is the use of multiple different sensing mechanisms that can target a broad range of different biorelevant species. The amphiphilic deep cavitands can both bind fluorescent dyes and interact with charged biomolecules. These non-cavity-based interactions, when paired with additives such as heavy metal ions, modulate fluorescence response in an indirect manner, and these different mechanisms allow selective recognition of serine phosphorylation, lysine acetylation and arginine citrullination. Other targets include heavy metals, drugs of abuse and protein isoforms. Furthermore, the hosts can be applied in supramolecular tandem assays of enzyme function: the broad scope allows analysis of such different enzymes as chromatin writers/erasers, kinases and phosphatases, all from a single host scaffold. Finally, the indirect sensing concept allows application in sensing different oligonucleotide secondary structures, including G-quadruplexes, hairpins, triplexes, and i-motifs. Discrimination between DNA strands with highly similar structures such as G-quadruplex strands with bulges and vacancies can be achieved. Instead of relying on a single highly specific fluorescent probe, the synthetic hosts tune the fluorophore-DNA interaction, introducing multiple recognition equilibria that modulate the fluorescence signal. By applying machine learning algorithms, a classification model can be established that can accurately predict the folding state of unknown sequences. Overall, the unique recognition profile of self-folded deep cavitands provides a powerful, yet simple sensing platform, one that can be easily tuned for a wide scope of biorelevant targets, in complex biological media, without sacrificing selectivity in the recognition.

KEY REFERENCES

- Liu, Y.; Perez, L.; Mettry, M.; Gill, A.D.; Byers, S. R.; Easley, C.J.; Bardeen, C.J.; Zhong, W.; Hooley, R.J. Site Selective Reading of Epigenetic Markers by a Dual-Mode Synthetic Receptor Array. *Chem. Sci.* **2017**, *8*, 3960 – 3970.[1] *This paper describes the use of arrayed cavitand:dye complexes to sense post-translational modifications in oligopeptides, and an analysis of the dual-mode recognition properties of the hosts.*
- Liu, Y.; Lee, J.; Perez, L.; Gill, A.D.; Hooley, R.J.; Zhong, W. Selective Sensing of Phosphorylated Peptides and Monitoring Kinase and Phosphatase Activity with a Supramolecular Tandem Assay. *J. Am. Chem. Soc.* **2018**, *140*, 13869–13877.[2] *By changing the nature of the dye, the mechanism of sensing of the arrayed host:dye complexes could be altered, selectively targeting serine phosphorylation rather than lysine methylation. This also includes applications in monitoring kinase and phosphatase activity.*
- Liu, Y.; Perez, L.; Gill, A.D.; Mettry, M.; Li, L.; Wang, Y.; Hooley, R.J.; Zhong, W. Site-Selective Sensing of Histone Methylation Enzyme Activity via an Arrayed Supramolecular Tandem Assay. *J. Am. Chem. Soc.* **2017**, *139*, 10964–10967.[3] *In this paper, the sensor arrays were applied to monitor enzymatic activity of methyltransferases and demethylases. The arrayed host:guest complexes were functional in the presence of the enzymes and their cofactors, and site-selective enzyme activity monitoring was possible.*

- Chen, J.; Hickey, B.L.; Wang, L.; Lee, J.; Gill, A.D.; Favero, A.; Pinalli, R.; Dalcanale, E.; Hooley, R.J.; Zhong, W. Selective Discrimination and Classification of G-Quadruplex Structures with a Host:Guest Sensing Array. *Nat. Chem.* **2021**, *13*, 488–495.[4] *The "indirect" sensing mechanism of arrayed host:dye complexes was exploited to detect structural changes in DNA G-quadruplexes, including discrimination and classification of structures that varied only in backbone topology.*

INTRODUCTION

The field of biosensing has been greatly expanded in recent years by the application of macrocyclic synthetic receptors.[5] To function as an effective host sensor for biological species, the synthetic receptor must contain a defined cavity into which the target species can bind, and also allow pairing with a suitable optical reporter molecule. This dye can be displaced upon recognition of the target, providing an optical output that converts simple recognition into "sensing", using indicator displacement assays (Figure 1a).[6] Small molecule hosts (as opposed to surface-attached sensors[5]) are also greatly improved if they are water-soluble. Achieving all these requirements is not trivial, and the "best" (i.e. most popular) receptors for this process are macrocycles such as cucurbit[n]urils (CB),[7] cyclophanes,[8] sulfonato-calixarenes (CX)[9] and shallow phosphonate cavitands[10] (Figure 1c), as well as other aromatic macrocycles such as pillararenes.[11] Each receptor type has its advantages and disadvantages, and some exquisite applications have been shown that highlight the creativity and technical prowess of the scientists involved.[12] Substrate selectivity can be enhanced by applying the hosts as differential sensors.[13] By using multiple hosts and fluorophores in an array, a fluorescence fingerprint can be generated that, after statistical data processing,[14] provides a simple output that allows differentiation and classification of highly similar substrates (illustrated in Figure 1b). The differential sensing strategy allows synthetic receptor-based sensors to approach levels of selectivity in biosensing that are usually only associated with antibodies or other platforms that exploit biomacromolecules, with none of the synthetic complexity or expense associated with biological sensors.[15]

Challenges still remain, however. The major limitation with using these privileged hosts is that they tend to favor specific targets: each host binds one specific type of molecular structure very strongly. For example (and with apologies for this simplistic generalization), the "best substrates" for CBs are N,N'-alkyldiammonium ions,[16] CX and phosphonate cavitands favor methylammonium ions ($R\text{-}NMe_3^+$)[17] and $R\text{-}NH_2Me^+$,[18,19] respectively), and cyclophanes favor soft cations such as methylated ammonium[20] or guanidinium ions.[21] Molecular tweezers are also effective, and target unmodified lysines in proteins.[22,23] All these species are essential molecular markers, and can be found as post-translational modifications (PTMs) in proteins,[24] drugs of abuse[25] and neurotransmitters,[26] among others. They are quite similar in structure to each other, however, mostly falling into the category of "soft nitrogen-containing cations", and as such, many of the most impressive applications involve a narrow subset of targets. Other vital targets such as small anions,[27] lipids,[28] sugars,[29] phosphates[30] and nucleotides[31] require other types of sensors for detection, with varying degrees of success.

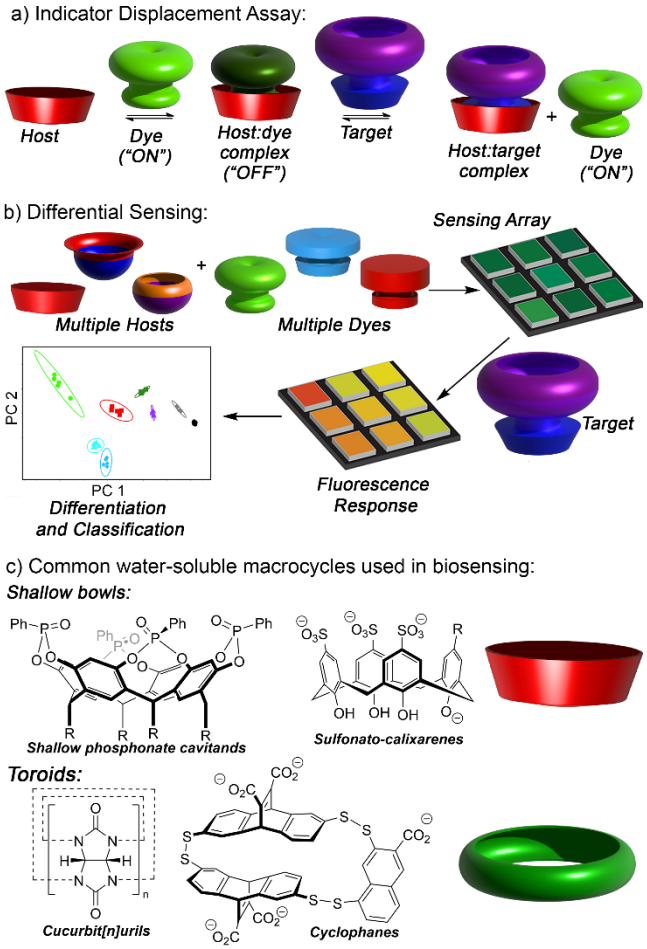


Figure 1. Synthetic Receptors as Biosensors. a) Illustration of the indicator displacement assay concept; b) the extension to differential sensing; c) examples of other, well-known macrocycles commonly used for molecular recognition and sensing in water.

Obviously, these challenges are endemic for biosensing with small molecule receptors: small, inflexible cavities cannot rival antibodies or RNA for *both* substrate binding affinity and scope, as their defined binding site is limited in the size, shape and charge of groups that can be recognized. As a result, there is an unfilled niche in this area – can more flexible, complex synthetic receptors be used that exploit multiple different recognition and sensing mechanisms for different types of targets, and are not restricted to sensing *via* 1:1 host:guest binding and indicator displacement assays? In this Account, we describe our work on biosensing with water soluble, flexible deep cavitands, and hopefully illustrate the power of these hosts to combine both strong selectivity and broad scope in sensing a wide array of complex biological targets.

MOLECULAR RECOGNITION WITH WATER-SOLUBLE DEEP CAVITANDS

Self-folding deep cavitands, pioneered by Rebek,[32-34] can act as host molecules for a variety of targets by presenting a large, defined cavity with a single entry portal to suitably sized guests. These larger cavities introduce a challenge for solubility (especially in water) and conformational stability: species with large open cavities tend to fill those spaces, either by aggregation,[35] dimerization[36] or simply adopting an "undesired" conformation.[37] One solution is to synthesize large, rigid deep cavitands, which has been explored to great success

by Gibb.[38] The other technique is to use flexible walls and rely on "self-folding" for conformational stability. This can be achieved by introducing self-complementary hydrogen bonding groups at the upper rim of the hosts. Depending on the groups used, the self-folding technique can even confer conformational stability on the hosts in water. This, of course, is essential for biorecognition, which at minimum must be performed in aqueous solution, and hopefully in complex biological media. Water-solubility is usually conferred on the cavitands by incorporating charged functional groups, either at the upper or lower rim. We have generally focused on the benzimidazole deep cavitand scaffold (Figure 2a), which uses water molecules as the hydrogen-bonding agents to ensure proper folding. It should be noted here that Rebek and Yu have investigated similar types of cavitand scaffolds as small molecule hosts and reactors.[39]

Our most successful cavitand for biorecognition purposes is the anionic cavitand **TCC** (Figure 2a).[40] When folded, the cavity is highly electron rich, with 8 aromatic panels surrounding an internal space of ~180 Å^3. This is inspired by the acetylcholinesterase active site, which consists of an "aromatic box" of tryptophan and tyrosine sidechains.[41] As such, **TCC** is an excellent host for R-NMe$_3^+$ ions (Figure 2c), notably choline and acetylcholine[40,42] and can also bind hydrocarbons of the correct size and shape.[43] Recognition of R-NMe$_3^+$ ions is generally both enthalpically and entropically favorable. Cation-π and CH-π interactions between host and guest provide an enthalpic driving force, and as R-NMe$_3^+$ ions are relatively "greasy", entropically favorable desolvation of the surrounding water molecules also occurs upon binding. Importantly, as the cavitand is open-ended, the selectivity extends to a wide range of "R" groups. Examples include surfactants,[44] proteins,[45] polymers,[46] and modified fluorophores.[47]

There is one more important trait shown by **TCC**: as the charged carboxylates reside only on the upper rim of the host, it is lipophilic and quite prone to self-aggregation. This has allowed application as a host and sensor in membrane bilayers,[44-46, 48] but also has a large (and complex) effect on the recognition and sensing of charged and hydrophobic species in water, which is one of the keys to its broad target scope.

SELECTIVE RECOGNITION AND SENSING WITH ARRAYED CAVITAND•FLUOROPHORE COMPLEXES

While the molecular recognition aspects of deep cavitands have been extensively analyzed by NMR, coupling this recognition to an optical output is more challenging. Indicator/fluorescence displacement assays[6] have been widely used with calixarenes (especially tetrasulfonatocalix[4]arene, CX4,[49] or tailored CX4 derivatives[50]), cucurbit[n]urils,[7] cyclophanes[20] and pillararenes.[11] Unfortunately, common fluorophores such as lucigenin (LCG) do not bind in deep cavitands, so we have focused on two different types of fluorophores: combinations of fluorescein or rhodamine isothiocyanates with cholamine (**FCAT** and **RCAT**, respectively), and styrylpyridinium dyes such as **DSMI**, **PSMI**, **MSMI** and **DQMI** (Figure 2b). Each of these dyes shows differing affinity and fluorescence response with different cavitands, which is ideal for broad target scope and differential sensing. Our initial focus was to establish the scope of the recognition/sensing system for peptides modified with methylated lysines: as cavitands such as **TCC**, **CHC** and **NHC** bind R-NMe$_3^+$ ions, they are perfectly suited for selective recognition of trimethyllysine groups. Using arrayed deep cavitands has a key advantage over simple hosts – while the cavities in different deep cavitands are essentially identical, the upper rim functional groups are oriented directly at *the peptide backbone*. This introduces secondary interactions with neighboring sidechains which, in addition to recognizing the Kme$_3$ group, can be used to further increase target selectivity.

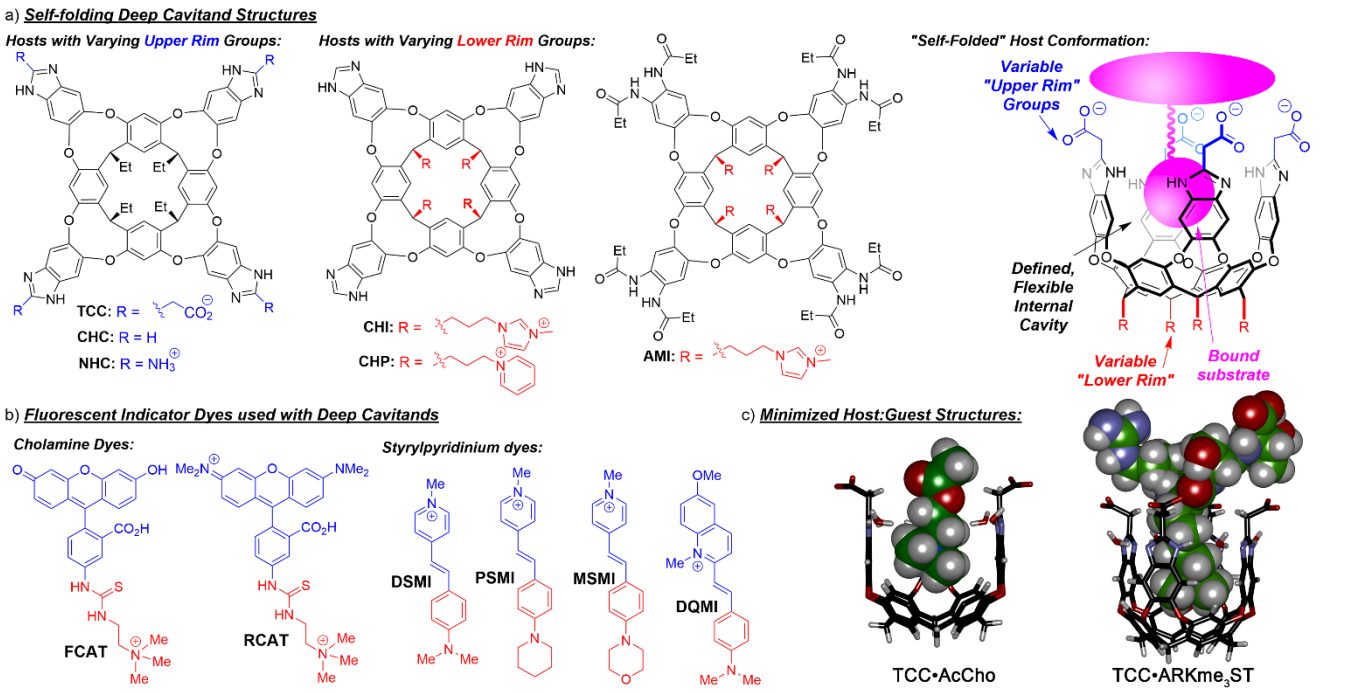


Figure 2. Structures of sensor components. a) Deep cavitand structures used for sensing biological targets, and the folded conformation in solution; b) suitable fluorophores; c) minimized structures of host:guest complexes between **TCC** and R-NMe$_3^+$-containing guests.

The simple **TCC•FCAT** host:dye pair is a powerful sensor for Kme_3 groups in oligopeptides.[51] The mechanism is slightly more complex than simple displacement: upon binding, **TCC** causes 80% quenching of **FCAT** *via* a triggered aggregation mechanism (Figure 3a). The host:guest binding interaction between **TCC** and **FCAT** triggers the formation of a micellar aggregate of the amphiphilic cavitand, and **FCAT** is self-quenched upon aggregation. Upon addition of the modified histone H3 peptide $H3K9me_3(1\text{-}21)$, the Kme_3 residue outcompetes the dye, displacing it and turning on the sensor. Neither the parent unmethylated peptide (H3K9) or the H3K9me variant caused appreciable recovery of the fluorescence signal: this simple cavitand•dye sensor is *state-selective* for different lysine methylations. It should be noted that this selectivity, while an exciting proof-of-principle for us, is also seen with other known hosts such as calixarenes[52] and cyclophanes,[53] which show stronger affinities for Kme_3[54] and greater levels of $Kme_3/Kme_2/Kme$ selectivity[55] than **TCC**. However, the most important advantage of the deep cavitand system is its adaptability: state-selectivity can be achieved with a single host:guest complex, but for more challenging targets and a finer level of selectivity, multiple arrayed host:guest complexes are needed. The constituents of these arrays can be quite varied, including different cavitand hosts, dyes, or other environmental factors.

The initial array focused on varying cavitand structure and pH: three hosts (negative **TCC**, neutral **CHC** and cationic **NHC**, Figure 2, 3c) were combined with **RCAT** at different pH values.[1] The thioureas in **RCAT** can interact synergistically with the cavitand upper rims: the C=S group can favorably interact with acids, whereas the urea NH groups can interact with bases. Biasing the pH of the system changes the balance of these interactions, adding a subtle layer of recognition selectivity between the hosts and dyes (Figure 3b), as well as the hosts and the oligopeptide targets. This effect was shown by discriminating a series of cholamine derivatives,[1] and then applied to more complex methylated lysines. Fourteen peptides were tested, based on the histone H3 sequence, with a variety of different modifications. These included varying the methylation state at K9 ($K9me_0$–$K9me_3$), changing the position of the trimethylation on the backbone ($K4/9/27/36/79me_3$), combining other PTMs such as phosphorylations and acylations with Kme_3 (e.g. $K4me_3K9AcS10p$, $K79me_3T80p$), and varying the peptide length (Figure 3).

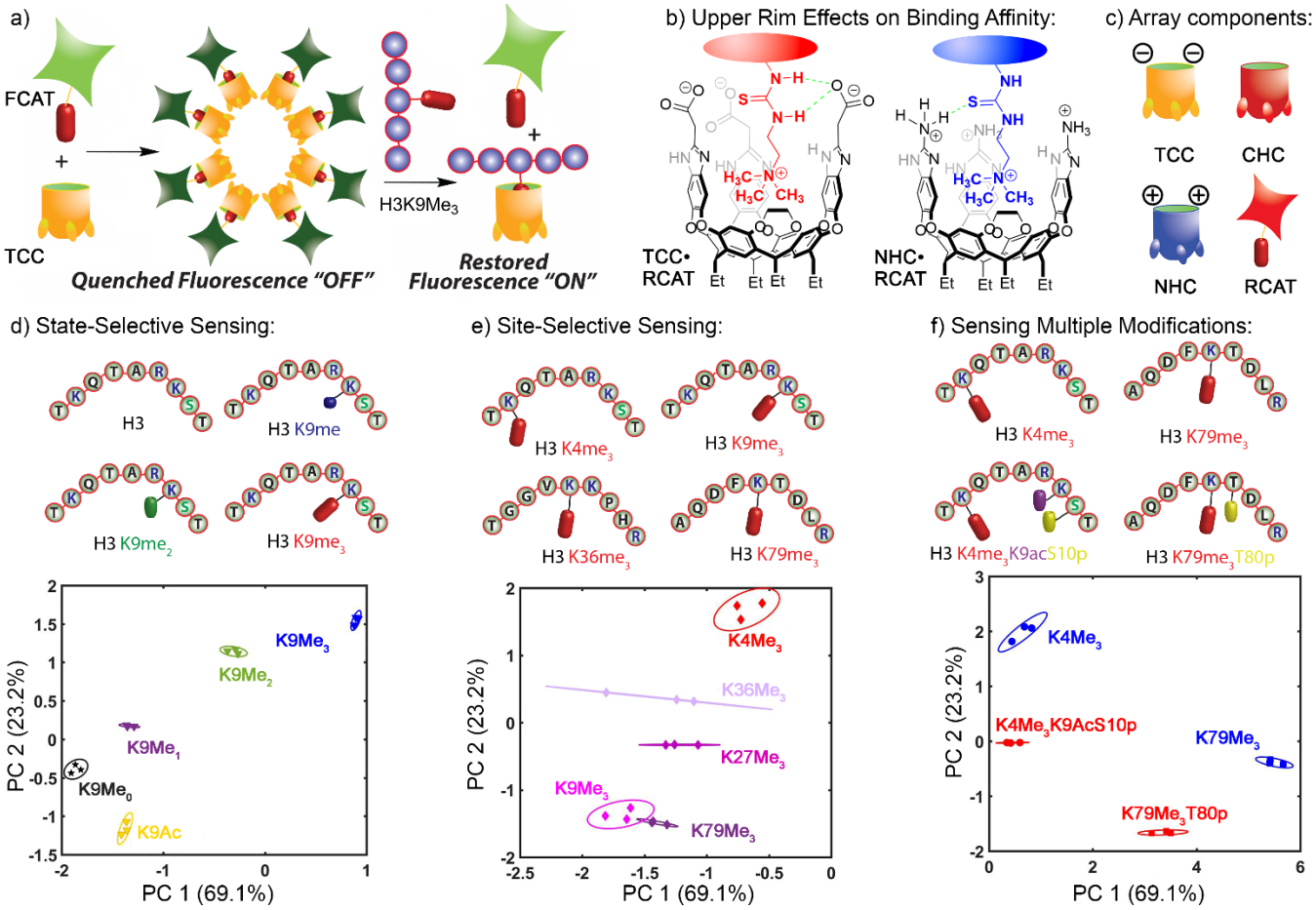


Figure 3. Differential sensing of lysine trimethylations in oligopeptides.[1] a) Illustration of the aggregation-dependent mechanism of indicator displacement shown by the TCC cavitand; b) upper rim groups provide a second layer of binding selectivity; c) array components used for differential sensing. Subsets of the PCA scores plot from the observed fluorescence data indicating discrimination of d) methylation state; e) methylation site; f) multiple methylations. Partially reproduced from ref 1, copyright 2017, Royal Society of Chemistry.

A six-factor array was used (the three cavitands **TCC**, **CHC** and **NHC**, **RCAT** dye, in either pH 7.4 or 9.0 buffer). Fluorescence changes were monitored upon addition of peptide to each sensor well in the array, and the outputs analyzed by Principal Component Analysis (PCA, Figure 3d-f). Differentiating methylation state is simple, but the much greater challenge is to differentiate between identical modifications at different *positions* on the backbone. This can be achieved with the 6-component array: peptide strands of the same length, with Kme_3 at different positions (K4/9/27/36/79) could almost all be differentiated with 95% confidence (Figure 3e). The differentiation here is subtle, as the cavitands all bind the Kme_3 group, but the relative response is affected by interactions between the upper rim groups of the cavitand and the adjacent residues on the peptide backbone. The exact interactions that bias the affinity in each state are not obvious, but two more sets of experiments shed some light on the situation. If modifications are added to residues adjacent to the Kme_3 group (e.g. $K79me_3T80p$, Figure 3f), significant changes in response are observed, presumably due to the large change in charge gained from converting a threonine OH to a phosphate group. The changes need not be adjacent, however – $K4me_3K9AcS10p$ can be distinguished from $K4me_3$, which indicates that the larger peptide structure is affecting the recognition. In addition, the sensor can discriminate peptides where the Kme_3 group resides at the same position, but the length of the peptide changes.[1] As opposed to smaller, rigid hosts such as CX4 and CB7 that are highly specific for small molecule targets and functional groups on peptides, the more lipophilic cavitands are affected by the superstructure of long peptides. The triggered aggregation process seen by the **TCC•FCAT** complex is a good illustration of this: the cavitands are not only affected by what binds *inside* the cavity, but *external* interactions also affect the fluorescence response. This has drawbacks – these hosts are notably poor at sensing individual amino acids, and teasing out the mechanisms of sensing becomes challenging, but the benefits in enhanced selectivity and scope are large.

INDIRECT SENSING OF NON-METHYLATION MODIFICATIONS

The next step was to utilize these unexpected, non-cavity-based interactions (i.e. *indirect sensing*) and extend the target scope to modifications other than lysine methylations. The anionic **TCC** can interact with cationic peptides in solution, *via* charge-based interactions rather than cavity-based recognition. This effect is not dominant when binding Kme_3 groups, as the cavity binding outcompetes charge interactions. However, when **1** is combined with styrylpyridinium dyes such as **DSMI** (Figure 2b), the **TCC•DSMI** complex can be disrupted by a cationic peptide, which changes the emission of the dye. This allows host:guest arrays to be applied in targeting *non-methylation* PTMs, using the cavity to bind a suitable dye, and the upper rim carboxylates to target cationic peptides.

The most obvious application is one that gives a large change in charge to maximize selectivity, namely phosphorylation of alcoholic sidechains (Ser, Tyr, Thr).[2] Figure 4b shows the effect of a series of cationic peptides and their phosphorylated counterparts on the fluorescence of the host:dye complexes, illustrating the selectivity of **TCC•DSMI**. Whereas the cationic peptide interacts with the host and competitively displaces **DSMI** (which is protruding from the top of the host), this interaction is reduced with a phosphorylated peptide. The sensor is not targeting the phosphate group, but the *absence* of cationic peptide interactions. Importantly, **DSMI** binds more weakly in the **TCC** than the **FCAT** or **RCAT** dyes,[1,2] and so can be displaced upon addition of the cationic peptide. The more strongly binding dyes can only be displaced by $R\text{-}NMe_3^+$ groups, so are minimally affected by unmodified peptides. This allows tailored selectivity for different modifications, depending on the dye used.

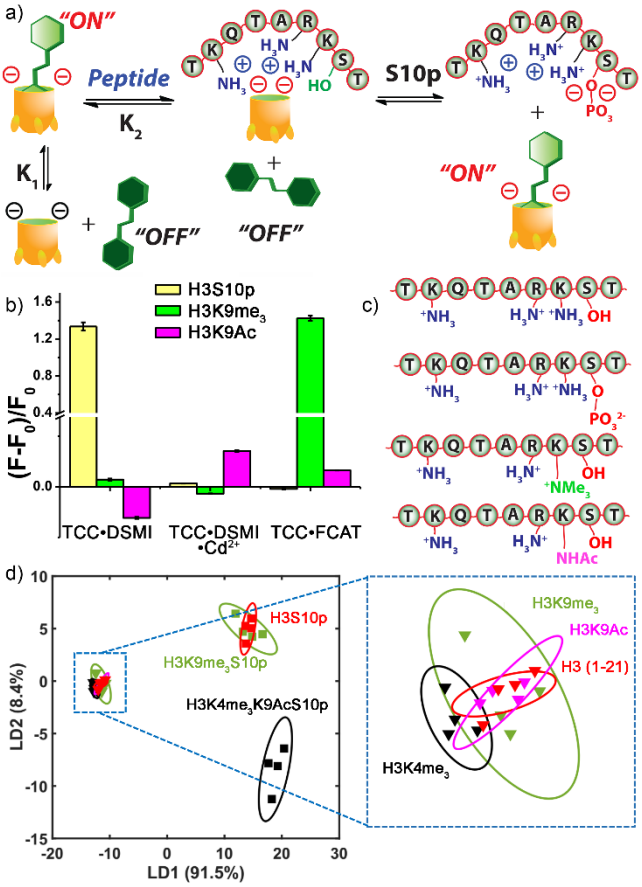


Figure 4. Differential sensing of peptide phosphorylations.[2] a) Indirect sensing mechanism; b) Fluorescence responses of modified peptides to different host•dye complexes; c) representative peptide sequences; d) PCA scores plots from the observed fluorescence data indicating discrimination between phosphorylated peptides and those with other modifications.

To achieve the greatest possible differentiation, arrayed sensing would be ideal, but as only the anionic **TCC** can interact with the cationic peptides, other, non-cavitand elements must be incorporated. Fortunately, **TCC** shows a surprising affinity for heavy metal ions in solution.[56] Even though the only functional groups present in the host are carboxylates, micromolar affinity for large metal ions is seen, mainly due to the fact that the carboxylate groups in **TCC** (when folded) are pre-oriented to chelate large metals. When **TCC** is paired with a dye (either **RCAT** or **DSMI**) and metal ions, this introduces multiple different fluorescence responses. The metals can either bind to the **TCC•dye** complex, displace the dye from the cavitand, or both. Also, the fluorescence response of the **TCC•dye•M** complexes is variable: heavy metals quench the dye, whereas filled shell metals (e.g. Zn^{2+}) do not. This has interesting applications in environmental analysis,[56] but for the purposes of this review, the most important result is that simply adding 10-50 μM metal

ion easily expands the array, without needing to synthesize new hosts or dyes.

A 4-component phosphate sensing array, consisting of **TCC**, **DSMI**, and $Cu^{2+}/Co^{2+}/Ni^{2+}$/no metal was exposed to a series of modified peptides as before (some of which are shown in Figure 4c), and the fluorescence responses were analyzed by Linear Discriminant Analysis (LDA). The **TCC•DSMI•M^{2+}** sensor array can discriminate phosphorylated peptides from their unphosphorylated counterparts, even when those peptides contain Kme_3 modifications. For example, the three phosphorylated peptides H3S10p, H3K4me_3S10p and H3K4me_3K9AcS10p are fully separated from H3, H3K4me_3, H3K9me_3 and H3K9Ac (Figure 4d). The array is also capable of selectively discriminating phosphorylation at tyrosine or threonine from serine in different peptide strands. Most interestingly, the sensing of peptide phosphorylations can be performed in phosphate buffer! As the target of the sensor is the unmodified peptide, not the phosphate group, the sensing of peptide phosphorylations is tolerant to phosphate in the media. Some degradation of performance is seen in phosphate buffer when compared to Tris, but the fact that phosphate sensing is functional in phosphate buffer at all is remarkable.

This concept can be extended to other types of non-methylation PTM, although this becomes significantly more challenging as the charge differences lessen. Lysine acylations can be discriminated from unmodified peptides, as the positive charge in the NH_3^+ group is removed upon modification. Similarity, arginine citrullinations (Ci) can be targeted:[57] the change in sidechain in this case involves converting the guanidinium group in Arg to a urea, i.e. C=NH_2^+ becomes C=O. After some optimization, a 4-factor array (**TCC•FCAT**, with no metal, La^{3+}, Ni^{2+} or Zn^{2+}) was shown to distinguish H3R2$_{Ci}$ and H3R2$_{Ci}$R8$_{Ci}$ from H3(1-21), so in this case, the sensor shows modification state selectivity. However, positional selectivity is more challenging – discriminating between H3R2$_{Ci}$ and H3R8$_{Ci}$ requires two dyes, two metals and analysis *via* 3D PCA for full separation. Still, this is a challenging target for sensing, so even this moderate selectivity is impressive.

SUPRAMOLECULAR TANDEM ASSAYS

Small molecule hosts such as CB[7] and CX4 have been widely used in supramolecular tandem assays of enzyme function,[58] including cholinesterases,[59] proteases,[60] oxidases,[61] lysine methyltransferase,[62] and kinases,[63] among others.[64,65] These systems usually monitor indicator displacement by either the reactant or product of the enzymatic reaction over time. This is a simple, effective monitoring method that relies on the affinity of small molecules for the host. However, this can be challenging to apply for large, complex targets and cannot detect positionally-selective enzyme reactions such as those of chromatin writers and erasers: as the deep cavitand sensors can provide positionally selective sensing, the obvious next step was to apply this to enzyme reaction monitoring.[3,51]

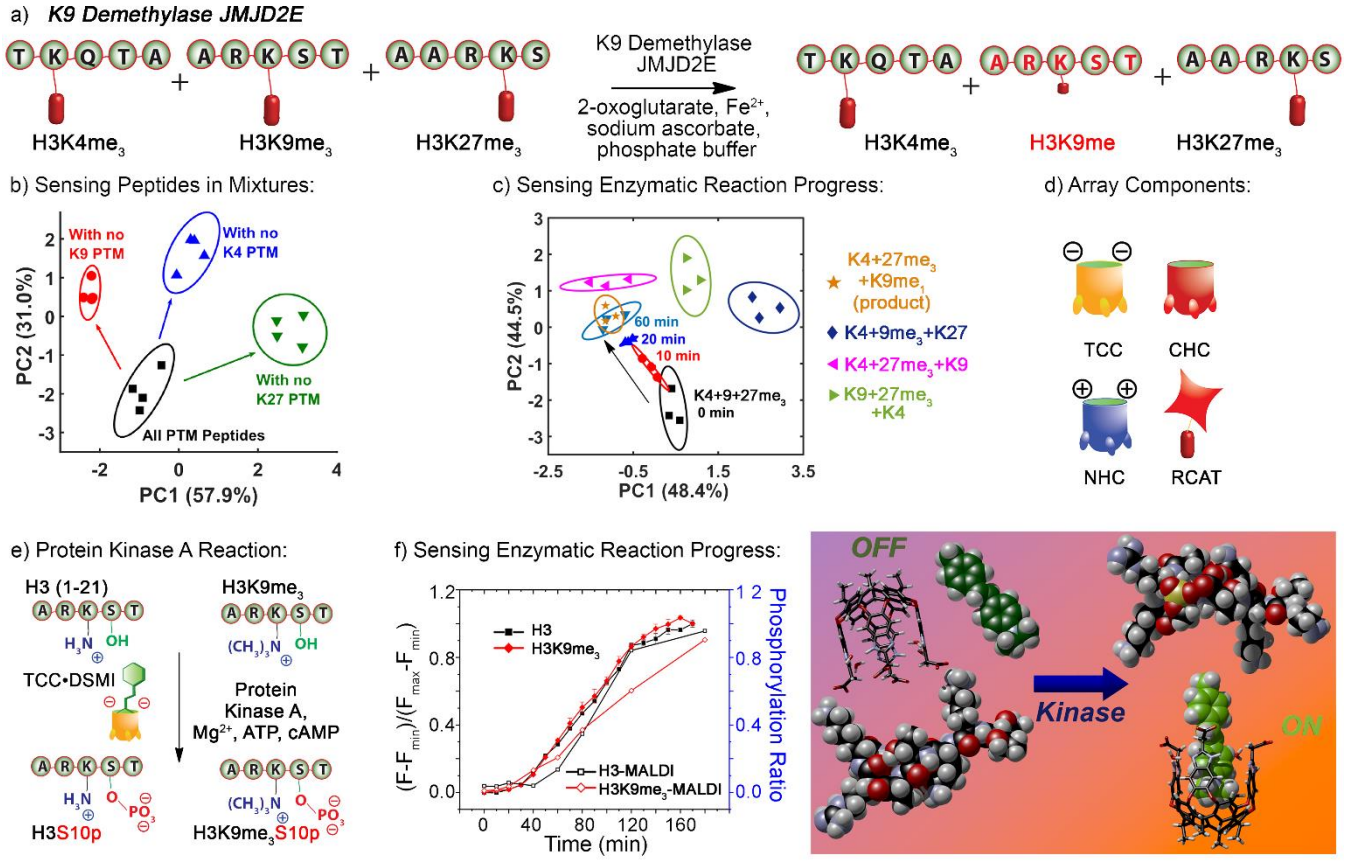


Figure 5. Supramolecular Tandem Assays. a) K9 demethylase process;[51,3] b) PCA scores plots from the observed fluorescence data indicating discrimination between lysine methylations in peptide mixtures;[3] c) PCA monitoring the JMJD2E demethylation reaction; d) array components used for the JMJD2E assay; e) Protein kinase A dephosphorylation process; f) cavitand-based monitoring of the protein kinase A reaction.[2]

The key to all supramolecular tandem assays is to ensure that the target binding process is not affected by the required conditions for the enzymatic reaction. Whereas simple recognition of Kme_3 peptides can be performed in a controlled aqueous environment, monitoring the function of lysine demethylase JMJD2E requires that the sensor be tolerant to the enzyme itself, as well as cofactors such as Fe^{2+}, sodium ascorbate and 2-oxoglutarate. Fortunately, the **TCC•FCAT** sensing pair is not affected these conditions. While Fe^{2+} has some effect on the fluorescence response, it is small in enough in this case to not affect peptide detection. Upon adding **TCC•FCAT** to a solution of $H3K9me_3$ in the presence of JMJD2E and required cofactors, the modified peptide displaces **FCAT** from the host, "turning on" the fluorophore.[51] Over time, as the enzyme removes the methyl groups from $H3K9me_3$, the fluorophore is reclaimed by the host, and the drop in fluorescence over time allows monitoring of enzyme activity, including determining the K_i of a suitable inhibitor.[51] This shows that the **TCC•FCAT** sensing pair can monitor lysine demethylase activity, but this is not unique, as Nau has shown this type of assay using CB[7].[62] However, the cavitand-based arrays are capable of much more than simply monitoring activity: by using multiple cavitands and dyes, *site-selective* monitoring of enzyme activity is possible in complex mixtures of modified peptides.[3] The arrayed tandem assay follows the same principle as the simple one-component FDA process: Kme_3 groups displace fluorophores from the cavitand(s), and enzymatic removal of those methyl groups allows recombination of the cavitand•dye complex, lowering the observed fluorescence. In this case, a three-component methylation-sensitive array was used, with **TCC, CHC, NHC** cavitands and **RCAT** dye at pH 7.4 (to ensure proper enzyme function). For monitoring the site-selective JMJD2E demethylation process, three modified peptides were combined and subjected to enzymatic demethylation (Figure 5a): $K4me_3$, $K9me_3$ and $K27me_3$. Only $K9me_3$ is demethylated by JMJD2E, and the product is the monomethylated peptide K9me. The reactions were added to each array well, and the sensor responses at various time points were subjected to PCA, together with a series of controls corresponding to all possible outcomes of the reaction (i.e. demethylation of the other peptides as well as the "correct" target).

As the reaction proceeds (Figure 5a), the reaction mixture moves toward the position where K9me + $K4me_3$ + $K27me_3$ is located, and further away from the other two controls, indicating that the enzyme acts on only the K9 methylation site, but not K4 nor K27 (Figure 5b, c). The analysis is not only capable of distinguishing incorrect reaction products, but also the state of the product: fully demethylated K9 is distinguishable from K9me. This methylation-sensitive array is also capable of site-selective monitoring of methylation writer enzymes.

As well as sensing the "obvious" targets of lysine methylation enzymes with cavitand•dye arrays, the indirect sensing concept can be used to monitor kinase (Figure 5e) and phosphatase activity. The challenge here is that kinases require the presence both ATP and cAMP for reaction, which is quite difficult for sensors that detect the phosphate group! Fortunately, as the **TCC•DSMI** sensor shows no affinity for small phosphates and its fluorescence is mediated by the nature of the cationic peptide, it is capable of monitoring kinase activity. Figure 5f shows one example of the sensing performance: phosphorylation of peptides H3(1-21) and $H3Kme_3$ by protein kinase A was monitored by both the **TCC•DSMI** sensor and by MALDI-MS analysis of aliquots drawn from solution. The fluorescence changes follow the phosphorylation process well, and this sensor is not affected by the presence of a Kme_3 group, despite using the same cavitand host. Other examples of enzymes that can be analyzed include Aurora Kinase A and alkaline phosphatase.[2]

INDIRECT SENSING MECHANISMS FOR STRUCTURAL ANALYSIS OF DNA 3D STRUCTURE

The indirect sensing mechanism for phosphate sensing illustrates that cavitands are capable of sensing biorelevant targets that do not bind in the cavity, as there are multiple mechanisms of interaction and fluorescence response that can be exploited. The logical extension of this concept is to use the hosts as fluorescence modulators rather than the recognition element, i.e. use *dyes* as the primary recognition motif, and add the hosts as competitive sensing elements. This allows sensing of unexpected targets, namely non-canonically folded nucleic acid strands.[4,66] The anionic cavitand **TCC** has no affinity for DNA or RNA at all, which is quite obvious. However, the styrylpyridinium dyes are ideally shaped for interaction with DNA: they are rigid, cationic and slightly kinked, mirroring the general structure of known DNA ligands.[67] As there are myriad small molecule dyes that can simply "detect" DNA, we focused on a different goal: can arrayed host:guest complexes sense small differences in DNA structure that cannot be simply detected by sequence analysis? Nucleotide strands can form complex 3D structures other than the paradigmatic double helix: examples of noncanonical folds include G-quadruplexes (G4s), Hoogsteen triplexes, and hairpins and i-motifs, among others. Despite some notable recent advances,[68,69] it remains challenging to analyze non-canonical structures from sequences alone, and often requires modified nucleotides,[70-72] X-ray crystallography and/or multidimensional NMR spectroscopy for complete structural analysis.

The initial target was a series of DNA G4 strands. G4s are generally classified into three different structural topologies: parallel, antiparallel or hybrid (Figure 6a). The structural differences between the topologies are small, varying in the conformation of the guanine glycosidic angles (*syn* or *anti*) and the relative orientation of the strands. The styrylpyridinium dyes bind G4 DNA strands with micromolar affinity,[4,66] and show enhanced emission while bound. They also show strong affinity for cavitand hosts in water. As such, we combined a series of hosts, **TCC, CHI**, **CHP**, **AMI**, (Figure 2a) and a shallow phosphonate cavitand provided by the Dalcanale group[10] with **DSMI** and **PSMI** dyes and monitored the fluorescence changes upon addition of a series of DNA strands. This 10-component array was highly responsive to changes in nucleotide structure. The potential of the array was shown by focusing on different G4s: 23 different sequences (6 parallel, eight antiparallel and nine hybrid, all with lengths of 22-26 bases) could all be fully discriminated based on their folding type. This result was further confirmed by Canonical Discriminant Analysis (Figure 6b), which shows correct classification of 114 out of 115 samples (including repeats) into the three topologies.

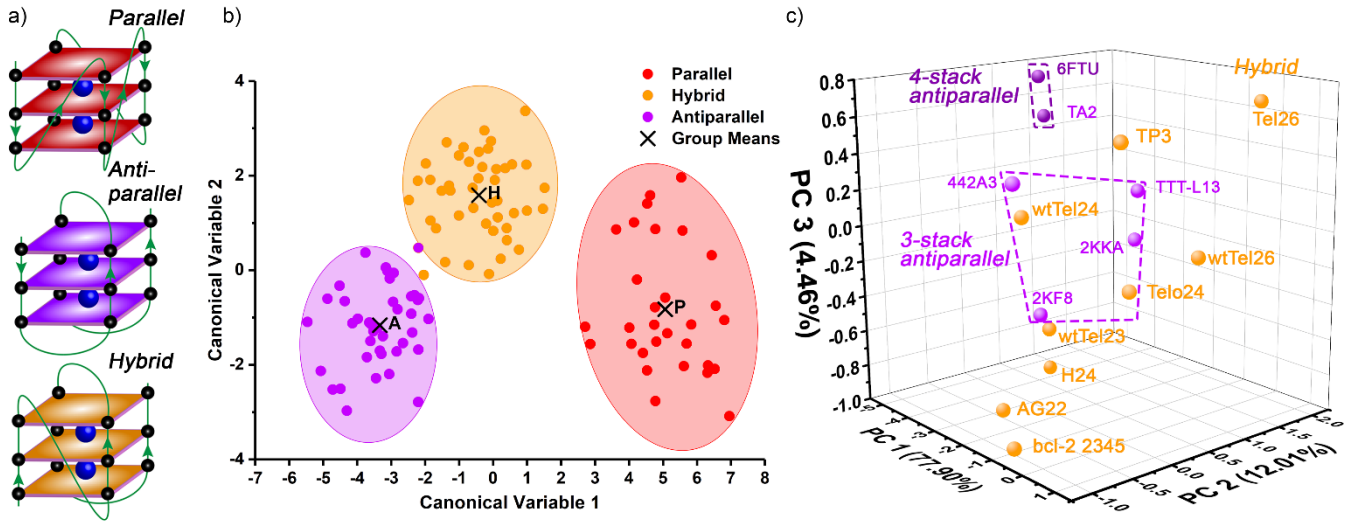


Figure 6. Nucleotide Sensing. a) Different G4 topologies; b) CDA scores plot of 115 G4 samples, grouped by topology; c) 3D PCA scores plot of 15 antiparallel and hybrid G4 strands with 10-component cavitand:dye sensing array.[4]

While the classification ability of the array is excellent, the discriminatory power is also impressive. The sensor can easily discriminate the parallel G4 structures from the hybrid/antiparallel topologies. Figure 6c focuses on the differentiation between hybrid and antiparallel strands which both contain one or two of the four G strands in an orientation opposite to the others, and display *anti* and *syn* guanines. They are even sometimes considered as a single topology group,[73] making their differentiation by simple methods highly challenging. The antiparallel and hybrid G4s that co-locate close to each other are all of similar lengths (22-26 nts) and all display three G-quartet stacks. The antiparallel structures that vary in the number of G-quartet stacks are fully separated from the 3-stack G-quartets (either antiparallel or hybrid). Some impressive discrimination effects can also be seen for the 3-stack G4s (Figure 6c). Although *2KKA* has high sequence similarity with *2KF8*, *AG22*, *wtTel23* and *wtTel24* (varying only by the absence of one T or one A on the 3' or 5' end), it is well separated from the other strands. The limit of detection (LOD) ranges from 3 - 36 nM, and the material usage is low: the minimum total DNA consumption is < 12 pmol. The sensor also performs well in the presence of interfering compounds like lysozyme, lactose, and serum, and can sense topology switches triggered by substituting K^+ with Na^+ in the buffer.

The excellent selectivity shown by the initial array in discriminating G4 structures allowed application to other folds.[66] The key to the detection and the differential selectivity is the synergistic combination of *both* host and dye interacting with DNA folded structures. Figure 7a shows a representation of the sensing mechanism, which illustrates the complexity of the system. The dyes all bind to the DNA and show differential emission enhancement when bound. The host•dye complexes can also form ternary complexes with folded DNA. Cationic cavitands can interact with structured DNA while binding the dyes, and this complex interaction adds a level of selectivity that is not possible with the dyes alone. The hosts are essential for selective sensing, even though the dyes are the main reporters.

An improved, modified sensor array (hosts **CHI**, **CHP**, **AMI** and dyes **DSMI**, **PSMI**, **DQMI**, **MSMI**)[66] was applied to a series of 18 other non-classical DNA folded structures, including hairpins, Hoogsteen triplexes, and i-motifs, as well as imperfect G4s with bulges or vacancies. These targets ranged from entirely different folding motif structures, such as triplexes, to those with very small differences, such as bulges or vacancies in G4 structure. Interrogating 18 DNAs with 16 array elements established a multidimensional data set with 1440 data points. This large and multidimensional data set illustrates the power and the limitations of the array-based analysis: it is not simple to quantify the classification effect, nor easy to pick out the most effective components from such a large and complex data set. This is where machine learning approaches can be exploited - to select the most optimal array elements, evaluate the classification effect, and even *predict* unknown DNA folding.

The fluorescence output from the array was analyzed by the support vector machine (SVM) algorithm,[74,75] followed by a PCA step to minimize data dimensions, and visualized in a 2D SVM Decision Region Boundary (DRB) plot (Figure 7b, 7c). Training the SVM-RFE algorithm with the data from the known DNA pool can lead to a classification model that permits the use of the fluorescence responses from an "unknown" DNA to predict its folding motif. To test this hypothesis, we chose 4 DNA strands with known folding motif as "unknown" targets: *c-myc 2345* and *EAD4* (known to form a parallel G4 structure), *APE 1-4 track* (an i-motif), and *Telo24* (a hybrid G4). The resultant PC values were then projected to the original DRB plot by the classifiers resulting from running Scaled 2D PCA-coupled SVM-RFE on the training dataset. The location of the unknown DNA (shown as solid blocks in Figure 7b) shows successful prediction: all of the 10 repeats for each of the "unknown" targets located within the boundary for their corresponding folding. This is remarkable, because the DNA pool contained structures with high similarity, including parallel vs. hybrid G4s, and parallel G4s with bulges or vacancies that differ by only one base.

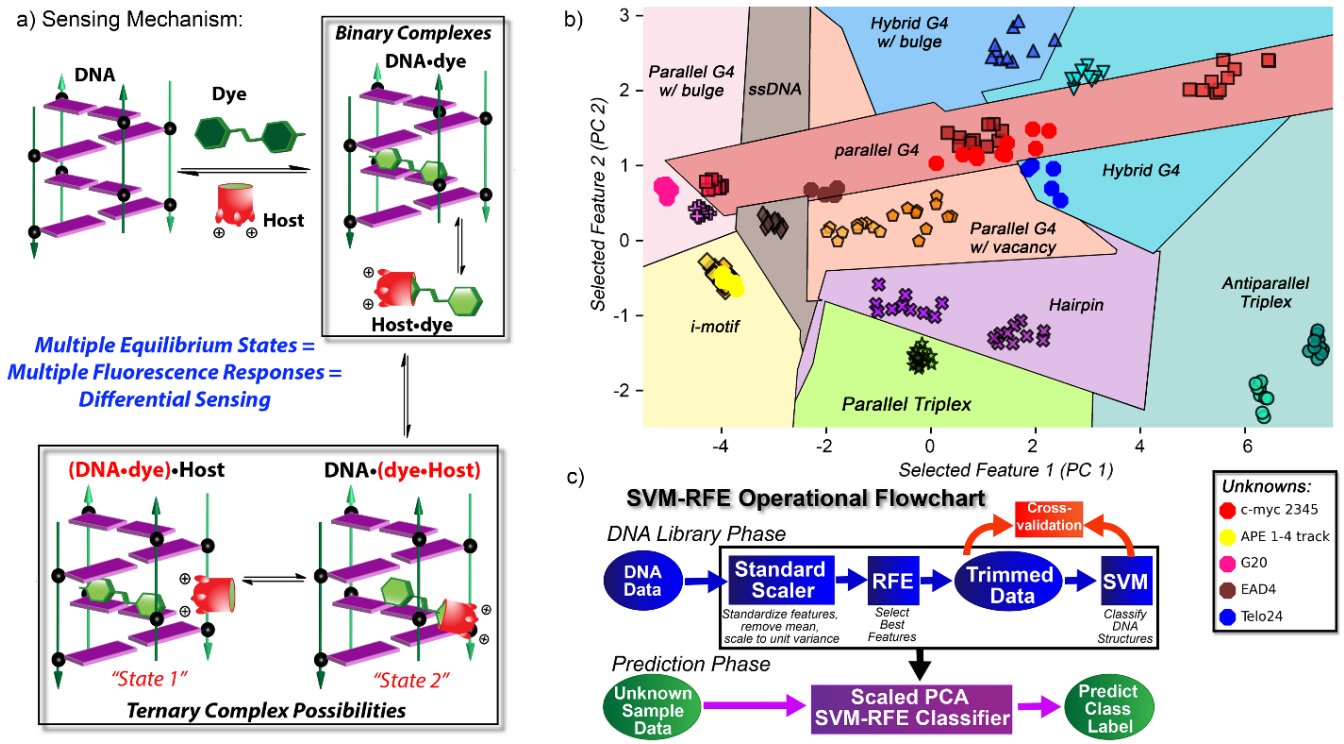


Figure 7. Machine Learning Optimized Sensing. a) Schematic of the sensing mechanism of the synthetic receptor array using an i-motif as example; b) SVM-RFE Decision Region Boundary (DRB) plot showing classification of different DNA folding motifs; c) Operational flowchart of the SVM-RFE process.

OUTLOOK

The outlook for biosensing with synthetic receptors is very exciting. The last four decades (or so) have seen a steady progression in the "difficulty" of application: the initial focus was on molecular recognition in organic solvents, then progressed to function in water and biological media. The complexity of targets has increased along with it, from the pioneering work on cation recognition, to anions, small neutral molecules and then truly biologically important species such as amino acids, peptides and nucleotides. Water-soluble deep cavitands are an enticing structural motif for biosensing, as they are high affinity hosts for important biological targets, can be easily tuned, and exploit multiple different recognition mechanisms for differential sensing of diverse targets in the presence of different targets. As well as the targets described here, we have used cavitand:dye combinations to detect other relevant molecules, including protein isoforms,[76] biothiols,[77] and hydrophobic species such as steroids,[78] cannabinoids[79] or insect pheromones.[80]

There are a couple of important "next steps" for synthetic receptor-based biosensing. Differential sensing is a powerful tool, and can take advantage of Big Data to use machine learning and/or AI approaches to truly achieve the selectivity and sensitivity of antibody-based sensing. An important task is to selectively recognize and confirm target identity in a high complex mixture of similar components, rather than detecting individual standards in pure solutions. This will require a combination of synthetic organic chemistry and statistical analysis, to provide a large suite of host and dye combinations, and the computing power to pick out essential elements for optimized sensing. Also, as has been elegantly described by Hof,[50] an important task is to move away from pure water as a medium and to focus on biorelevant conditions for recognition, including high salt concentration, the presence of lipid aggregates (membranes and liposomes) and other competitive species, as well as function in biomedia such as urine, saliva, serum, cells and living organisms. Some elegant work has been performed to illustrate the difficulties in molecular recognition in these types of media, but also the opportunities.[81,82] Obviously, as the application medium becomes more challenging, choice of receptor tends to revert to those with the most well understood mechanisms of action, such as calixarenes and cucurbiturils. Flexible deep cavitands are far more variable and complex in their behavior: this has allowed us to study a wide range of targets and show powerful selectivity, but we have only just started to apply them in complex media: we have shown that selectivity persists in serum,[4] urine[78,79] and saliva,[79] and applied them in human cells,[47] but, happily, there is still much to be done.

AUTHOR INFORMATION

Corresponding Authors

Wenwan Zhong – Department of Chemistry and the Environmental Toxicology Graduate Program; University of California-Riverside, Riverside, CA, 92521, U.S.A. ORCID: orcid.org/0000-0002-3317-3464; Email: wenwan.zhong@ucr.edu

Richard J. Hooley – *Department of Chemistry; Department of Biochemistry; University of California-Riverside, Riverside, CA, 92521, U.S.A.* ORCID: orcid.org/0000-0003-0033-8653; Email: richard.hooley@ucr.edu

Notes

The authors declare no competing financial interest.

Biographies

Wenwan Zhong is Professor of Chemistry at University of California – Riverside (UCR), specializing in separation and sensing. She received her B.S. degree in Applied Chemistry from University of Science and Technology of China, and her Ph.D. in Analytical Chemistry at Iowa State University under the direction of Prof. Edward S. Yeung. After finishing her postdoctoral training at the Los Alamos National Laboratory, she established her independent research group at UCR in 2006. The Zhong group focuses on development of analytical methods to study biomarkers and understand their functions, utilizing molecular tools like functional nucleic acids, synthetic or protein receptors, and nanomaterials.

Richard J. Hooley obtained his undergraduate degrees in Natural Sciences from Emmanuel College, Cambridge University, and his Ph.D. in the laboratory of Prof. Martin F. Semmelhack at Princeton. Following a postdoctoral position in the lab of Prof. Julius Rebek Jr. at The Scripps Research Institute in La Jolla, CA he began his independent career in the Department of Chemistry at the University of California – Riverside in 2008 and was promoted to full Professor in 2018. The Hooley group works on various aspects of organic, inorganic and supramolecular chemistry, with a focus on the synthesis and applications of biomimetic self-assembled cages and hosts, and their functions in supramolecular catalysis, molecular recognition and biosensing.

ACKNOWLEDGMENTS

The authors would like to thank the National Science Foundation (CHE-1707347) for support.

REFERENCES

(1) Liu, Y.; Perez, L.; Mettry, M.; Gill, A.D.; Byers, S. R.; Easley, C.J.; Bardeen, C.J.; Zhong, W.; Hooley, R.J. Site Selective Reading of Epigenetic Markers by a Dual-Mode Synthetic Receptor Array. *Chem. Sci.* **2017**, *8*, 3960 – 3970.

(2) Liu, Y.; Lee, J.; Perez, L.; Gill, A.D.; Hooley, R.J.; Zhong, W. Selective Sensing of Phosphorylated Peptides and Monitoring Kinase and Phosphatase Activity with a Supramolecular Tandem Assay. *J. Am. Chem. Soc.* **2018**, *140*, 13869–13877.

(3) Liu, Y.; Perez, L.; Gill, A.D.; Mettry, M.; Li, L.; Wang, Y.; Hooley, R.J.; Zhong, W. Site-Selective Sensing of Histone Methylation Enzyme Activity via an Arrayed Supramolecular Tandem Assay. *J. Am. Chem. Soc.* **2017**, *139*, 10964–10967.

(4) Chen, J.; Hickey, B.L.; Wang, L.; Lee, J.; Gill, A.D.; Favero, A.; Pinalli, R.; Dalcanale, E.; Hooley, R.J.; Zhong, W. Selective Discrimination and Classification of G-Quadruplex Structures with a Host:Guest Sensing Array. *Nat. Chem.* **2021**, *13*, 488–495.

(5) Pinalli, R.; Pedrini, A.; Dalcanale, E. Biochemical sensing with macrocyclic receptors. *Chem. Soc. Rev.* **2018**, *47*, 7006–7026.

(6) Sedgwick, A. C.; Brewster II, J. T.; Wu, T.; Feng, X.; Bull, S. D.; Qian, X.; Sessler, J. L.; James, T. D.; Anslyn, E. V.; Sun, X. Indicator displacement assays (IDAs): the past, present and future. *Chem. Soc. Rev.* **2021**, *50*, 9-38.

(7) Barrow, S. J.; Kasera, S.; Rowland, M. J.; del Barrio, J.; Scherman, O. A., Cucurbituril-Based Molecular Recognition. *Chem. Rev.* **2015**, *115*, 12320-12406.

(8) Beaver, J.; Waters, M.L. Molecular Recognition of Lys and Arg Methylation *ACS Chem. Biol.* **2016**, *11*, 643–653.

(9) Daze, K. D.; Hof, F., The Cation–π Interaction at Protein–Protein Interaction Interfaces: Developing and Learning from Synthetic Mimics of Proteins That Bind Methylated Lysines. *Acc. Chem. Res.* **2013**, *46*, 937-945.

(10) Pinalli, R.; Dalcanale, E.; Supramolecular Sensing with Phosphonate Cavitands. *Acc. Chem. Res.* **2013**, *46*, 399-411.

(11) Ogoshi, T.; Yamagishi, T.; Nakamoto, Y. Pillar-Shaped Macrocyclic Hosts Pillar[*n*]arenes: New Key Players for Supramolecular Chemistry. *Chem. Rev.* **2016**, *116*, 7937-8002.

(12) van Dun, S.; Ottmann, C.; Milroy, L.-G.; Brunsveld, L. Supramolecular Chemistry Targeting Proteins. *J. Am. Chem. Soc.* **2017**, *139*, 13960-13968.

(13) You, L.; Zha, D.; Anslyn, E. V. Recent advances in supramolecular analytical chemistry using optical sensing. *Chem. Rev.* **2015**, *115*, 7840–7892.

(14) Stewart, S.; Ivy, M. A.; Anslyn, E. V. The use of principal component analysis and discriminant analysis in differential sensing routines. *Chem. Soc. Rev.* **2014**, *43*, 70–84.

(15) Bradbury, A.; Plückthun, A. Standardize antibodies used in research. *Nature* **2015**, *518*, 27-29.

(16) Cao, L.; Šekutor, M.; Zavalij, P. Y.; Mlinarić-Majerski, K.; Glaser, R.; Isaacs, L. Cucurbit[7]uril·Guest Pair with an Attomolar Dissociation Constant. *Angew. Chem. Int. Ed.* **2014**, *53*, 988-993.

(17) Daze, K. D.; Pinter, T.; Beshara, C. S.; Ibraheem, A.; Minaker, S. A.; Ma, M. C. F.; Courtemanche, R. J. M.; Campbell, R. E.; Hof, F., Supramolecular hosts that recognize methyllysines and disrupt the interaction between a modified histone tail and its epigenetic reader protein. *Chem. Sci.* **2012**, *3*, 2695-2699.

(18) Pinalli, R.; Brancatelli, G.; Pedrini, A.; Menozzi, D.; Hernández, D.; Ballester, P.; Geremia, S.; Dalcanale, E. The origin of selectivity in the complexation of *N*-methyl amino acids by tetraphosphonate cavitands. *J. Am. Chem. Soc.* **2016**, *138*, 8569–8580.

(19) Bontempi, N.; Biavardi, E.; Bordiga, D.; Candiani, G.; Alessandri, I.; Bergese, P.; Dalcanale, E. Probing lysine mono-methylation in histone H3 tail peptides with an abiotic receptor coupled to a non-plasmonic resonator. *Nanoscale* **2017**, *9*, 8639–8646.

(20) Peacor, B. C.; Ramsay, C. M.; Waters, M. L., Fluorogenic sensor platform for the histone code using receptors from dynamic combinatorial libraries. *Chem. Sci.* **2017**, *8*, 1422-1428.

(21) James, L. I.; Beaver, J. E.; Rice, N. W.; Waters, M. L., A Synthetic Receptor for Asymmetric Dimethyl Arginine. *J. Am. Chem. Soc.* **2013**, *135*, 6450-6455.

(22) Acharya, S.; Safaie, B. M.; Wongkongkathep, P.; Ivanova, M. I.; Attar, A.; Klärner, F.-G.; Schrader, T.; Loo, J. A.; Bitan, G.; Lapidus, L. J. Molecular basis for preventing α-synuclein aggregation by a molecular tweezer. *J. Biol. Chem.* **2014**, *289*, 10727–10737

(23) Sinha, S.; Lopes, D. H. J.; Du, Z.; Pang, E.S.; Shanmugam, A.; Lomakin, A.; Talbiersky, B.; Tennstaedt, A.; McDaniel, K.; Bakshi, R.; Kuo, P.-Y.; Ehrmann, M.; Benedek, G. B.; Loo, J. A.; Klärner, F.-G.; Schrader, T.; Wang, C.; Bitan, G. Lysine-Specific Molecular Tweezers Are Broad-Spectrum Inhibitors of Assembly and Toxicity of Amyloid Proteins. *J. Am. Chem. Soc.* **2011**, *133*, 16958–16969.

(24) McGovern, R. E.; Snarr, B. D.; Lyons, J. A.; McFarlane, J.; Whiting, A. L.; Paci, I.; Hof, F.; Crowley, P. B. Structural study of a small molecule receptor bound to dimethyllysine in lysozyme. *Chem. Sci.* **2015**, *6*, 442-449.

(25) Beatty, M. A.; Selinger, A. J.; Li, Y.; Hof, F. Parallel synthesis and screening of supramolecular chemosensors that achieve fluorescent turn-on detection of drugs in saliva. *J. Am. Chem. Soc.* **2019**, *141*, 16763–16771.

(26) Kearney, P. C.; Mizoue, L. S.; Kumpf, R. A.; Forman, J. E.; McCurdy, A.; Dougherty, D. A. Molecular recognition in aqueous media. New binding studies provide further insights into the cation-π interaction and related phenomena. *J. Am. Chem. Soc.* **1993**, *115*, 9907-9919.

(27) Gale, P. A.; Howe, E. N. W.; Wu, X. Anion Receptor Chemistry. *Chem* **2016**, *1*, 351-422.

(28) Herschede, S.R.; Gneid, H.; Dent, T.; Jaeger E.B.; Lawson, L.; Busschaert, N. Bactericidal urea crown ethers target phosphatidylethanolamine membrane lipids. *Org. Biomol. Chem.* **2021**, *19*, 3838-3843.

(29) Edwards, N. Y.; Sager, T.W.; McDevitt, J. T.; Anslyn, E. V. Boronic Acid Based Peptidic Receptors for Pattern-Based Saccharide Sensing in Neutral Aqueous Media, an Application in Real-Life Samples. *J. Am. Chem. Soc.* **2007**, *129*, 13575-13583.

(30) Hargrove, A. E.; Nieto, S.; Zhang, T.; Sessler, J. L.; Anslyn, E. V., Artificial Receptors for the Recognition of Phosphorylated Molecules. *Chem. Rev.* **2011**, *111*, 6603-6782.

(31) Florea, M.; Nau, W. M. Implementation of anion-receptor macrocycles in supramolecular tandem assays for enzymes involving nucleotides as substrates, products, and cofactors. *Org. Biomol. Chem.* **2010**, *8*, 1033-1039.

(32) Hooley, R.J.; Rebek, J., Jr. Chemistry and Catalysis in Functional Cavitands. *Chem. Biol.,* **2009**, *16*, 255–264.

(33) Rudkevich, D.M.; Hilmersson, G.; Rebek, J., Jr. Self-folding cavitands. *J. Am. Chem. Soc.* **1998**, *120*, 12216–12225.

(34) Biros, S. M.; Rebek, J., Jr. Structure and binding properties of water-soluble cavitands and capsules. *Chem. Soc. Rev.* **2007**, *36*, 93-104.

(35) Hooley, R.J.; Gavette, J.V.; Mettry M.; Ajami, D.; Rebek, J., Jr. Unusual Orientation And Reactivity of Alkyl Halides in Water-Soluble Cavitands. *Chem. Sci.* **2014**, *5*, 4382–4387.

(36) Cram, D. J.; Choi, H. J.; Bryant, J. A.; Knobler, C. B. Host-guest complexation. 62. Solvophobic and entropic driving forces for forming velcraplexes, which are 4-fold, lock-key dimers in organic media. *J. Am. Chem. Soc.* **1992**, *114*, 7748-7765.

(37) Moran, J. R.; Ericson, J. L.; Dalcanale, E.; Bryant, J. A.; Knobler, C. B.; Cram, D. J. Vases and kites as cavitands. *J. Am. Chem. Soc.* **1991**, *113*, 5707-5714.

(38) Jordan, J. H.; Gibb, B. C. Molecular containers assembled through the hydrophobic effect. *Chem. Soc. Rev.* **2015**, *44*, 547–585.

(39) Mosca, S.; Yu, Y.; Rebek, J., Jr. Preparative scale and convenient synthesis of a water-soluble, deep cavitand. *Nat. Protoc.* **2016**, *11*, 1371–1387.

(40) Biros, S.M.; Ullrich, E.C.; Hof, F.; Trembleau, L.; Rebek, J., Jr. Kinetically Stable Complexes in Water: The Role of Hydration and Hydrophobicity. *J. Am. Chem. Soc.* **2004**, *126*, 2870.

(41) Ma, J. C.; Dougherty, D. A. the cation-π interaction *Chem. Rev.* **1997**, *97*, 1303-1324.

(42) Hof, F.; Trembleau, L.; Ullrich, E. C.; Rebek, J., Jr. Acetylcholine Recognition by a Deep, Biomimetic Pocket. *Angew. Chem. Int. Ed.* **2003**, *42*, 3150-3153.

(43) Hooley, R.J.; Van Anda, H. J.; Rebek, J., Jr. Extraction of Hydrophobic Species into a Water-Soluble Synthetic Receptor. *J. Am. Chem. Soc.,* **2007**, *129*, 13464–13473.

(44) Perez, L.; Mettry, M.; Caulkins, B.G.; Mueller, L.J.; Hooley, R.J. Lipid Bilayer Environments Control Exchange Kinetics of Deep Cavitand Hosts and Enhance Disfavored Guest Conformations. *Chem. Sci.* **2018**, *9*, 1836 – 1845.

(45) Ghang, Y.-J.; Lloyd, J.J.; Moehlig, M.P.; Arguelles, J.K.; Mettry, M.; Zhang, X.; Julian, R.R.; Cheng, Q.; Hooley, R.J. Labeled Protein Recognition at a Membrane Bilayer Interface by Embedded Synthetic Receptors. *Langmuir* **2014**, *30*, 10161–10166.

(46) Liu, Y.; Young, M.C.; Moshe, O.; Cheng, Q.; Hooley, R.J. A Membrane-Bound Synthetic Receptor Promotes Growth of a Polymeric Coating at the Bilayer-Water Interface. *Angew. Chem. Int. Ed.* **2012**, *51*, 7748–7751.

(47) Ghang, Y.-J.; Schramm, M.P.; Zhang, F.; Acey, R.A.; David, C.N.; Wilson, E.H.; Wang, Y.; Cheng, Q.; Hooley, R.J. Selective Cavitand-Mediated Endocytosis of Targeted Imaging Agents into Live Cells. *J. Am. Chem. Soc.* **2013**, *135*, 7090–7093.

(48) Liu, Y.; Liao, P.; Cheng, Q.; Hooley, R.J. Protein and Small Molecule Recognition Properties of Deep Cavitands in a Supported Lipid Membrane Determined by Calcination-Enhanced SPR Spectroscopy. *J. Am. Chem. Soc.* **2010**, *132*, 10383–10390.

(49) Guo, D. S.; Liu, Y. Supramolecular Chemistry of *p*-Sulfonatocalix[*n*]arenes and Its Biological Applications. *Acc. Chem. Res.* **2014**, 47, 1925-1934.

(50) Hof, F. Host–guest chemistry that directly targets lysine methylation: synthetic host molecules as alternatives to bio-reagents. *Chem. Commun.* **2016**, *52*, 10093-10108.

(51) Liu, Y.; Perez, L.; Mettry, M.; Easley, C.J.; Hooley, R.J.; Zhong, W. Self-Aggregating Deep Cavitand Acts as a Fluorescence Displacement Sensor for Lysine Methylation. *J. Am. Chem. Soc.* **2016**, *138*, 10746–10749.

(52) Beatty, M. A.; Borges-González, J.; Sinclair, N. J.; Pye, A. T.; Hof, F. Analyte-Driven Disassembly and Turn-On Fluorescent Sensing in Competitive Biological Media *J. Am. Chem. Soc.* **2018**, *140*, 3500-3504.

(53) Pinkin, N. K.; Waters, M. L. Development and mechanistic studies of an optimized receptor for trimethyllysine using iterative redesign by dynamic combinatorial chemistry. *Org. Biomol. Chem.* **2014**, *12*, 7059–7067.

(54) Ingerman, L. A.; Cuellar, M. E.; Waters, M. L. A small molecule receptor that selectively recognizes trimethyl lysine in a histone peptide with native protein-like affinity. *Chem. Commun.* **2010**, *46*, 1839-1841.

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

(56) Liu, Y.; Mettry, M.; Gill, A.D.; Perez, L.; Zhong, W.; Hooley, R.J. Selective Heavy Element Sensing with a Simple Host:Guest Fluorescent Array. *Anal. Chem.* **2017**, *89*, 11113–11121.

(57) Gill, A.D.; Hickey, B.L.; Wang, S.; Xue, M.; Zhong, W.; Hooley, R.J. Sensing of Citrulline Modifications in Histone Peptides by Deep Cavitand Hosts. *Chem. Commun.* **2019**, *55*, 13259–13262.

(58) Dsouza, R. N.; Hennig, A.; Nau, W. M. Supramolecular Tandem Enzyme Assays. *Chem. - Eur. J.* **2012**, *18*, 3444–3459.

(59) Guo, D.-S.; Yang, J.; Liu, Y. Specifically Monitoring Butyrylcholinesterase by Supramolecular Tandem Assay. *Chem. - Eur. J.* **2013**, *19*, 8755–8759.

(60) Ghale, G.; Ramalingam, V.; Urbach, A. R.; Nau, W. M. Determining Protease Substrate Selectivity and Inhibition by Label-Free Supramolecular Tandem Enzyme Assays. *J. Am. Chem. Soc.* **2011**, *133*, 7528–7535.

(61) Nau, W. M.; Ghale, G.; Hennig, A.; Bakirci, H.; Bailey, D. M. Substrate-Selective Supramolecular Tandem Assays: Monitoring Enzyme Inhibition of Arginase and Diamine Oxidase by Fluorescent Dye Displacement from Calixarene and Cucurbituril Macrocycles. *J. Am. Chem. Soc.* **2009**, *131*, 11558–11570.

(62) Florea, M.; Kudithipudi, S.; Rei, A.; González-Álvarez, M. J.; Jeltsch, A.; Nau, W. M. A Fluorescence-Based Supramolecular Tandem Assay for Monitoring Lysine Methyltransferase Activity in Homogeneous Solution. *Chem. - Eur. J.* **2012**, *18*, 3521–3528.

(63) Sahoo, H.; Hennig, A.; Florea, M.; Roth, D.; Enderle, T.; Nau, W. M. Single-Label Kinase and Phosphatase Assays for Tyrosine Phosphorylation Using Nanosecond Time-Resolved Fluorescence Detection. *J. Am. Chem. Soc.* **2007**, *129*, 15927–15934.

(64) Gober, I. N.; Waters, M. L. Supramolecular Affinity Labeling of Histone Peptides Containing Trimethyllysine and Its Application to Histone Deacetylase Assays. *J. Am. Chem. Soc.* **2016**, *138*, 9452–9459.

(65) Hennig, A.; Bakirci, H.; Nau, W. M. Label-free continuous enzyme assays with macrocycle-fluorescent dye complexes. *Nat. Methods* **2007**, *4*, 629–632.

(66) Chen, J.; Gill, A.D.; Hickey, B.L.; Gao, Z.; Cui, X.; Hooley, R.J.; Zhong, W. Machine Learning Aids Classification and Discrimination of Non-canonical DNA Folding Motifs by an Arrayed Host:guest Sensing System. *J. Am. Chem. Soc.* **2021**, *143*, 12791–12799.

(67) del Villar-Guerra, R.; Gray, R. D.; Trent, J. O.; Chaires, J. B. A rapid fluorescent indicator displacement assay and principal component/cluster data analysis for determination of ligand–nucleic acid structural selectivity. *Nucleic Acids Res.* **2018**, *46*, e41.

(68) del Villar-Guerra, R.; Trent, J. O.; Chaires, J. B. G-Quadruplex secondary structure obtained from Circular Dichroism spectroscopy. *Angew. Chem. Int. Ed.* **2018**, *57*, 7171–7175.

(69) Zuffo, M.; Xie. X.; Granzhan, A. Strength in numbers: development of a fluorescence sensor array for secondary structures of DNA. *Chem.-Eur. J.* **2019**, *25*, 1821–1818.

(70) Eubanks, C.S.; Forte, J.E.; Kapral, G.J.; Hargrove, A.E. Small molecule-based pattern recognition to classify RNA structure. *J. Am. Chem. Soc.* **2017**,*139*, 409–416.

(71) Eubanks, C.S.; Hargrove, A.E. RNA structural differentiation: opportunities with pattern recognition. *Biochemistry* **2019**, *58*, 199–213.

(72) Eubanks, C.S. Zhao, B.; Patwardhan, N.N.; Thompson, R.D.; Zhang, Q.; Hargrove, A.E. Visualizing RNA conformational changes via pattern recognition of RNA by small molecules. *J. Am. Chem. Soc.* **2019**, *141*, 5692–5698.

(73) Kwok, C. K.; Merrick, C. J. G-Quadruplexes: prediction, characterization, and biological application. *Trends Biotechnol.* **2017**, *35*, 997–1013.

(74) Ivanciuc, O. Applications of support vector machines in chemistry. *Rev. Comput. Chem.* **2007**, 291–400.

(75) Cortes, C.; Vapnik, V. Support-vector networks. *Machine Learning* **1995**, *20*, 273–297.

(76) Liu, Y.; Gill, A.D.; Duan, Y.; Perez, L.; Hooley, R.J.; Zhong, W. A Supramolecular Sensor Array for Selective Immunoglobulin Deficiency Analysis. *Chem. Commun.* **2019**, *55*, 11563–11566.

(77) Liu, Y.; Duan, Y.; Gill, A.D.; Perez, L.; Jiang, Q.; Hooley, R.J.; Zhong, W. Metal-Assisted Selective Recognition of Biothiols by a Synthetic Receptor Array. *Chem. Commun.* **2018**, *54*, 13147 – 13150.

(78) Gill, A.D.; Perez, L.; Salinas, I.N.Q.; Byers, S.R.; Liu, Y.; Hickey, B.L.; Zhong, W.; Hooley, R.J. Selective Array-based Sensing of Anabolic Steroids in Aqueous Solution by Host:Guest Reporter Complexes. *Chem-Eur. J.* **2019**, *25*, 1740 –1745.
(79) Gill, A.D.; Hickey, B.L.; Zhong, W.; Hooley, R.J. Selective Sensing of THC and Related Metabolites in Biofluids by Host:Guest Arrays. *Chem. Commun.* **2020**, *56*, 4352–4355.
(80) Hickey, B.L.; Chen, J.; Zou, Y.; Gill, A.D.; Zhong, W.; Millar, J.G.; Hooley, R.J. Enantioselective Sensing of Insect Pheromones in Water. *Chem. Commun.* **2021**, *57*, 13341–13344.
(81) Norouzy, A.; Azizi, Z.; Nau, W. M. Indicator Displacement Assays Inside Live Cells. *Angew. Chem. Int. Ed.* **2015**, *54*, 792-795.
(82) Pan, Y.-P.; Hu, X.-Y.; Guo, D.-S. Biomedical Applications of Calixarenes: State of the Art and Perspectives. *Angew. Chem. Int. Ed.* **2021**, *60*, 2768–2794.