

Organic & Biomolecular Chemistry

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

ARTICLE

Contributions of pocket depth and electrostatic interactions to affinity and selectivity of receptors for methylated lysine in water

Cite this: DOI: 10.1039/x0xx00000x

Received 00th September 2014,
Accepted 00th January 2012

DOI: 10.1039/x0xx00000x

www.rsc.org/

J. E. Beaver[§], B. C. Peacor[§], J. V. Bain, Lindsey I. James, M. L. Waters^{*}

Dynamic combinatorial chemistry was used to generate a set of receptors for peptides containing methylated lysine (KMe_n, n = 0–3) and study the contribution of electrostatic effects and pocket depth to binding affinity and selectivity. We found that changing the location of a carboxylate resulted in an increase in preference for KMe₂, presumably based on ability to form a salt bridge with KMe₂. The number of charged groups on either the receptor or peptide guest systematically varied the binding affinities to all guests by approximately 1–1.5 kcal/mol, with little influence on selectivity. Lastly, formation of a deeper pocket led to both increased affinity and selectivity for KMe₃ over the lower methylation states. From these studies, we identified that the tightest binder was a receptor with greater net charge, with a K_d of 0.2 μM, and the receptor with the highest selectivity was the one with the deepest pocket, providing 14-fold selectivity between KMe₃ and KMe₂ and a K_d for KMe₃ of 0.3 μM. This work provides key insights into approaches to improve binding affinity and selectivity in water, while also demonstrating the versatility of dynamic combinatorial chemistry for rapidly exploring the impact of subtle changes in receptor functionality on molecular recognition in water.

Introduction

Biology is defined by supramolecular chemistry in water. However, controlling molecular recognition in water is still a considerable challenge in the field of supramolecular chemistry.^{1,2} This is because many of the reliable noncovalent interactions that provide directionality and selectivity in organic solvents, such as hydrogen bonds, are noncompetitive in water.^{3,4,5,6} The hydrophobic effect, on the other hand, while providing a significant driving force in water, provides no selectivity or directionality.^{7,8} Nature overcomes these issues by utilizing multiple noncovalent interactions in concert to achieve the high selectivities needed for proper biological function.

One biomolecular recognition event that captured our attention due to its apparent simplicity as well as its biological and medicinal importance is the recognition of methylated lysine (Lys) by aromatic pockets in so-called “reader” proteins.⁹ Reader proteins bind to various sites of methylation in histone proteins, which make up the core of nucleosomes that package DNA in the nucleus.^{10,11} These binding events further recruit other proteins that control expression of the associated DNA.¹² Dysregulation of Lys methylation (and therefore aberrant gene expression) is associated with many types of cancer.^{13,14,15} Lys can be mono-, di-, or trimethylated (KMe_n, n=1–3; Fig. 1), and the methylation state dictates the protein that binds to it.

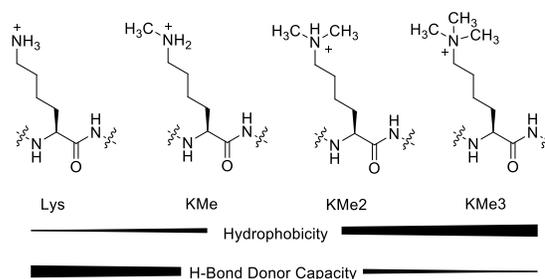


Figure 1. Methylation states of lysine in histone proteins. Increasing methylation increases hydrophobicity while decreasing hydrogen bond donor capability.

The general molecular recognition motif used by reader proteins to differentiate between the methylation states of Lys appears to be fairly simple: aromatic sidechains provide cation-π interactions with NMe groups and amides or carboxylates form hydrogen bonds or salt bridges with -NH groups (Fig. 2).^{16,17}

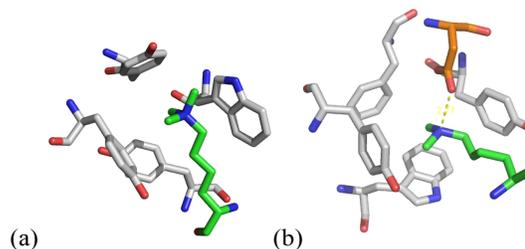


Figure 2. Exemplary binding pockets of reader proteins for methylated lysine. (a) BPTF bound to KMe3; pdb: 2F6J; (b) 53BP1 bound to KMe2; pdb: 2IG0. Green is KMe_n, gray is the aromatic pocket; orange is Asp. The hydrogen bond between KMe2 and Asp is shown in yellow.

We aimed to replicate this binding motif with synthetic receptors, using dynamic combinatorial chemistry (DCC) to identify promising candidates,¹⁸ with the dual goals of understanding molecular recognition of this class of guests and developing new tools for sensing these modified amino acids.¹⁹ To this end, we first reported a receptor, *rac*-**A₂B** (referred to as **A₂B** throughout the remaining text), which binds KMe3 in the context of a histone peptide sequence with a *K_d* of 2.6 μM in water.^{20,21} Subsequent redesign resulted in the receptor *meso*-**A₂N** (referred to as **A₂N** throughout the remaining text), which has a deeper pocket and binds KMe3 in the same sequence with a *K_d* of 0.3 μM.²¹ Herein, we describe more extensive variation of the **X** building block in dynamic combinatorial libraries (DCLs) that gave rise to a set of receptors, **A₂X**, each of which bind to methylated Lys. We find that the position of the charged group, number of charged groups, and depth of the binding pocket all influence affinity and/or selectivity in predictable ways, providing insight into approaches for achieving high affinity and high selectivity in water. This work also highlights the utility of DCC in structure/function studies and receptor optimization for molecular recognition in water. Finally, these simple receptors demonstrate some of the same principles used by reader proteins to achieve selectivity for the different methylation states of Lys.

Results and Discussion

System Design

DCC was used to conduct a structure function study of the **X**-subunit of **A₂X** by incorporating various monomers into the **X**-position. DCC using disulfide exchange is a reversible process that relies on the thermodynamic templating of favorable receptors by binding guests.^{22,23} This allows for both affinity screening and preparative synthesis of complex macrocyclic species. The **A₂X** structure was chosen for this study because the **A₂**-cleft persisted in previous receptors that bind to KMe3 identified by DCC.^{20,21,24} DCC is advantageous for structure-function studies as variation of monomer **X** allows one to investigate the influence of one component of the macrocycle without having to develop a new synthetic approach to each modified macrocycle.

Dynamic combinatorial libraries (DCLs) were biased toward the formation of **A₂X** in which monomers **A** and **X** (where **X** is a monomer shown in Fig. 3 or 5) were mixed in a 2:1 ratio in sodium borate buffer (50 mM, pH 8.5) with a total monomer concentration between 0.75 and 7.5 mM. To amplify hosts that interact primarily with the intended target, the monomers were mixed with an equimolar quantity of a tetrapeptide guest, Ac-KMe_nGGY-NH₂, where *n* is the degree of methylation, and Tyr is used as a concentration tag.[‡] In addition, one library was left untemplated to study the baseline thermodynamic state of the

library. These peptides were chosen to promote the amplification of species that specifically interact with the side chain of interest and also to remove the influence of electrostatic interactions, which would arise if the zwitterionic amino acid was used as guest. The libraries were monitored by analytical reverse phase HPLC for the amplification of any species in the presence of one guest over any other guests. Those amplified species were resynthesized (as a mixture of stereoisomers) using biased, preparative libraries and isolated for characterization and binding studies with histone H3 peptides.

Investigation of Receptors with Varied Electrostatic Interactions

Amplification of Receptors **A₂C and **A₂E**.** We previously reported the receptor **A₂B**, which binds to KMe3 (*K_d* = 2.6 ± 0.1 μM) in the context of a histone H3 peptide centered around lysine 9, with > 8-fold selectivity over unmethylated lysine.^{20,21} This binding event is driven primarily by cation-π interactions between the CH₃(δ⁺) groups and the aromatic groups in the receptor, with selectivity proposed to come from the increased cost of desolvation for each of the lower methylation states coupled with the loss of favorable CH-π interactions. Based on these observations, we wanted to explore the effect of carboxylate positioning as well as the contribution of multiple carboxylates. In isolation, electrostatic interactions are often negligible in water, but coupled with another binding event, such as the cation-π interaction between KMe3 and the aromatic cage of **A₂B**, the carboxylate may contribute more significantly.⁴ Moreover, it may contribute differently to the various methylation states of Lys, as has been seen in KMe_n reader proteins.^{9,17} To this end we used two new monomers, **C** and **E**, in exploratory DCLs to create the **A₂X** macrocycle platform, as shown in Fig. 3.

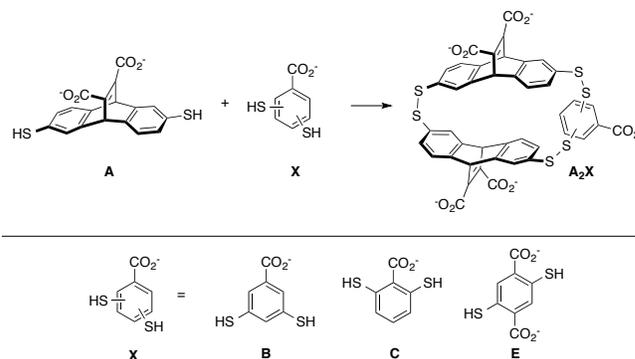


Figure 3. Monomers synthesized to explore the effect of electrostatic interactions in the **A₂X** framework

Monomer **C** was synthesized to determine the role of carboxylate position on selectivity. Specifically, positioning the carboxylate between the thiols in **C** holds it closer to the interior of the macrocycle of **A₂X**, while also rotating the carboxylate out of the plane of the benzene ring (Fig. 4). This may direct the carboxylate into the receptor binding pocket, thus mimicking the favorable hydrogen bonding and electrostatic interactions seen in several reader proteins for KMe2.^{9,17}

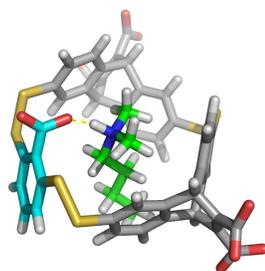


Figure 4. Molecular modelling of **A₂C** (**A** in gray and **C** in cyan) bound to butyldimethyl ammonium (green) as a model for KMe2. The position of the thiols *ortho* to the carboxylate rotates it out of the plane of the aromatic ring and into the binding pocket, allowing for a salt bridge with the NH group.

Initial DCLs with monomer **C** at 7.5 mM total monomer concentration suggested that selectivity between the methylation states is lost, based on amplification of **A₂C** by all methylation states of Lys^{20,25}. However, when the monomer and guest concentration were decreased 10-fold, **A₂C** was found to be amplified to a greater extent in the presence of KMe2 and KMe3 relative to Lys and KMe, suggesting that moving the carboxylate may alter the selectivity between KMe2 and KMe3 (Fig. S2). Thus, **A₂C** was isolated and its binding affinity and selectivity were investigated, as described below.

Monomer **E** was incorporated into the macrocycle to investigate the contribution of multiple anions toward binding affinity and selectivity. The addition of multiple anionic residues has been shown to result in cooperative binding and thus increased affinity, an event that we hypothesized monomer **E** would mimic.²⁶ Similar to **A₂C**, **A₂E** (Fig. S3) was amplified as the major species in each library that contained a Lys guest (0.75 mM total monomer), but it was amplified approximately 2-fold greater in the presence of KMe3 than each of the other methylation states of Lys. Thus, this receptor was also isolated and its binding affinity and selectivity were investigated as described below.

Characterization of Binding. Binding studies of each receptor with the histone H3 peptide Ac-WGGG-QTAR-KMe_n-STG-NH₂ (KMe_n, where *n* = 0-3) were performed by Isothermal Titration Calorimetry (ITC) in 10 mM sodium borate buffer, pH 8.5. The WGGG sequence was appended to the N-terminus of each peptide as a tag for concentration determination by UV. Control experiments performed with the same peptide lacking positively charged residues, Ac-WGGG-QTAGGSTG-NH₂, showed no interaction, though there is a possibility that the peptide sequence contributes some to the binding events with methylated lysine. The results of ITC experiments for **A₂B**, **A₂C**, and **A₂E** are shown in Table 1.

Binding To Methylated Lysine. ITC revealed that moving the carboxylate from the *meta*-position in monomer **B** to the *ortho*-position in monomer **C**, between the two thiols, had no influence on binding of **A₂C** to KMe3, KMe, or Lys relative to **A₂B**. In contrast, the binding affinity for **A₂C** to KMe2 is more favorable than for **A₂B** by a factor of 2.3, amounting to a differ-

ence of 0.5 kcal/mol, which nearly abolishes the selectivity between KMe3 and KMe2. The fact that only the binding affinity for KMe2 is improved suggests a binding motif in which the two methyl groups of KMe2 form CH(δ⁺)-π interactions with the aromatic pocket, while the N-H can form a hydrogen bond with the carboxylate of **C**, as it is perpendicular to the plane of the benzene ring, and thus directed into the pocket. Binding of KMe3 is not influenced by the position of the carboxylate because it cannot form a hydrogen bond. A likely reason that binding to KMe and Lys is unaffected is because they do not fit well in the pocket and thus do not bind with a specific geometry, so that once again the position of the carboxylate is not important.

Adding the extra carboxylate in **A₂E** improved binding to KMe3 by more than an order of magnitude ($K_d = 0.191 \pm 0.002$ μM), and provides an extra 1.3-1.5 kcal/mol binding energy for KMe, KMe2, and KMe3, compared to **A₂B**. Thus, while providing increased binding affinity, the extra carboxylate has little effect on binding selectivity between the methylated lysines. This observation is similar to that seen in Dougherty's receptor, where increasing the charge allowed for increased binding to well-solvated, polar guests through cooperative interactions, namely cation-π and electrostatic effects.²⁷ Surprisingly, however, **A₂E** increases the selectivity between KMe3 and Lys nearly two-fold. This arises from the fact that the binding affinity is only improved by 0.7 kcal/mol for **A₂E** with Lys relative to **A₂B**, while it is improved by 1.5 kcal/mol for KMe3.

Table 1. Thermodynamic binding data obtained for binding of **A₂B**, **A₂C**, and **A₂E** to Ac-WGGG-QTAR(KMe)_n-STG-NH₂ as measured by ITC.^a

Entry	Receptor ^b	Peptide	K_d^c (μM)	Selectivity ^d	ΔG^e (kcal/mol)
1 ^e	A₂B	KMe3	2.6 ± 0.1	-	-7.63 ± 0.03
2 ^e	A₂B	KMe2	6.3 ± 0.3	2.4	-7.10 ± 0.07
3 ^e	A₂B	KMe	13.9 ± 0.1	5.4	-6.64 ± 0.01
4 ^e	A₂B	Lys	22 ± 1	8.3	-6.38 ± 0.02
5 ^e	A₂B	R8GKMe3	17.1 ± 0.1	-	-6.52 ± 0.01
6	A₂C	KMe3	2.3 ± 0.1	-	-7.69 ± 0.02
7	A₂C	KMe2	2.8 ± 0.2	1.2	-7.57 ± 0.04
8	A₂C	KMe	13.8 ± 0.7	6.0	-6.63 ± 0.03
9	A₂C	Lys	22 ± 1	9.6	-6.34 ± 0.03
10	A₂C	R8GKMe3	29 ± 3	-	-6.17 ± 0.05
11 ^f	A₂E	KMe3	0.191 ± 0.002	-	-9.16 ± 0.01
12 ^f	A₂E	KMe2	0.5 ± 0.1	2.6	-8.5 ± 0.1
13 ^f	A₂E	KMe	1.6 ± 0.2	8.4	-7.92 ± 0.08
14 ^f	A₂E	Lys	6.7 ± 0.1	35	-7.05 ± 0.01
15 ^f	A₂E	R8GKMe3	2.7 ± 0.3	-	-7.59 ± 0.06

(a) All data determined by ITC, fit to one-site binding model; Conditions: 26 °C, in 10 mM sodium borate buffer, pH 8.5. (b) All receptors are mixtures of isomers except *rac*-**A₂B**. (c) Errors are from averages of three trials, unless noted otherwise. (d) Selectivity is calculated as the fold difference in affinity for KMe3 over the designated methylation state of the peptide in that row. (e) Data reported by Pinkin and Waters.²¹ (f) Average of two trials.

Influence of Neighboring Arginine (Arg) on Binding to KMe3. Previous work had shown that a neighboring charge in the peptide can influence the binding of methylated lysine to **A₂B** (Table 1, entry 5).²¹ Mutating the neighboring arginine (R8) to a neutral glycine (R8G) removes the adjacent cationic residue, revealing binding interactions more specific only to KMe3. This mutation results in around a 13-fold drop in affinity for both **A₂C** and **A₂E** with KMe3, which correlates to a 1.5 kcal/mol decrease in binding energy, indicating that arginine contributes to the binding event. **A₂B** exhibits a smaller drop in binding affinity for the mutated peptide, decreasing the binding affinity 7-fold, or 1.1 kcal/mol. This suggests that by moving the carboxylate or introducing an extra charge, the receptor has become more sensitive to neighboring charge in the peptide sequence.

Taken together, this data shows that for a series of receptors with similar binding pockets, the positioning and number of carboxylates at the benzene-derived **B** building block can provide significant improvement in binding affinity, but only subtle changes in selectivity. Furthermore, changing the charge on the peptide rather than the receptor led to a similar change in magnitude of the binding affinity, with a magnitude of 1-1.5 kcal/mol.

Investigation of Receptors with Deeper Binding Pockets

Previous efforts on the iterative redesign of **A₂B** furnished the macrocycle **A₂N**, which binds to KMe3 with high nanomolar affinity ($K_d = 0.30 \pm 0.04 \mu\text{M}$).²¹ While this affinity is comparative to **A₂E** presented here, the remarkable feature of **A₂N** is its selectivity. Whereas the receptors reported above all have between 1 and 3-fold selectivity for KMe3 over KMe2, **A₂N** exhibits 14-fold better affinity for the tri-methylated mark over KMe2 (Table 2, entry 2). This difference in affinity is consistent with the fact that receptors **A₂B**, **A₂C**, and **A₂E** have both shallow binding pockets and electrostatic interactions in close proximity to said pockets, whereas **A₂N** has a deeper aromatic binding pocket and is unable to form direct electrostatic interactions with guests bound inside. Instead, it can form more favorable cation- π and van der Waals interactions to increase its selectivity.^{28,29,30} KMe3 carries a permanent positive charge, making strong cation- π interactions between its N-CH₃ groups and the aromatic interior of the binding pocket. KMe2, KMe, and unmethylated Lys are also cationic under physiological conditions, but carry well-solvated protons that require desolvation in order to bind in an aromatic pocket.³¹ Binding of lower methylation states of Lys in a deep binding pocket requires more significant desolvation and decreases the net van der Waals interactions that contribute to binding.

We investigated three additional monomers to explore the effect of deepening the binding pocket by increasing the π -surface relative to **A₂B** (Fig. 5). Monomer **F**, a binol derivative, was synthesized to provide a flexible binding pocket and more possible cation- π contacts with the guests. However, exploratory DCLs showed no difference in amplification regardless of the methylated guest, so further studies were not pursued.

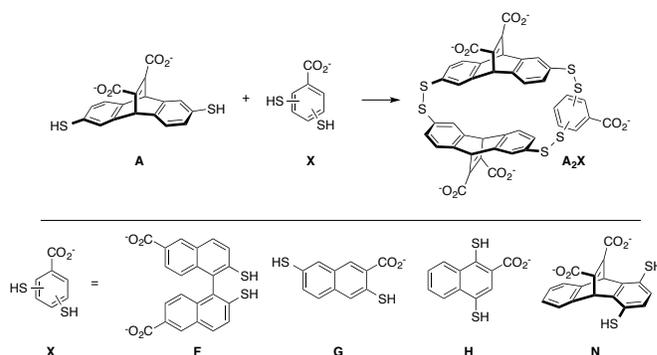


Figure 5. Monomers synthesized to explore the effect of a deeper aromatic pocket in the **A₂X** framework.

Monomer **H** was synthesized as a naphthalene-derived surrogate for monomer **N**, capable of forming a deeper hydrophobic binding pocket than **A₂B** to provide selective binding of the higher methylation states of Lys. However, DCC libraries biased toward the formation of **A₂H**, regardless of the presence of guest, produced two **A₂H** macrocycles almost exclusively after two days, suggesting that **A₂H** is the most thermodynamically favorable species in the library regardless of guest (Fig. S5). Subsequent molecular modelling showed that the unsubstituted aryl ring of the naphthalene fits ideally inside the **A₂**-cleft of **A₂H**, forming four intramolecular edge-face ArH- π interactions, and thus providing the self-templating effect observed in DCLs (Fig. 6). Binding studies by ITC confirmed that **A₂H** interacts only weakly with KMe3 and shows little to no selectivity for KMe3 over unmethylated Lys, suggesting that intermolecular interactions cannot overcome the intramolecular interactions between monomer units of this receptor.

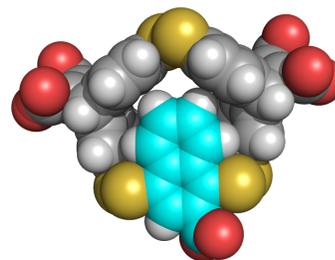


Figure 6. Molecular modelling of **A₂H** self-templating through favorable ArH- π interactions. **A** is shown in gray and **H** is shown in cyan.

Monomer **G** was incorporated into libraries to investigate the effect of a larger binding pocket on binding. Interestingly, molecular modelling of **A₂G** suggests that the macrocycle is able to adopt a twisted conformation (Fig. 7) with a deeper binding pocket than that observed for the benzene-derived monomers. DCLs formed two isomers of **A₂G** that were amplified 3.5-fold greater in the presence of KMe3 than in the presence of KMe2 and no significant amplification was observed in the presence of KMe or Lys.

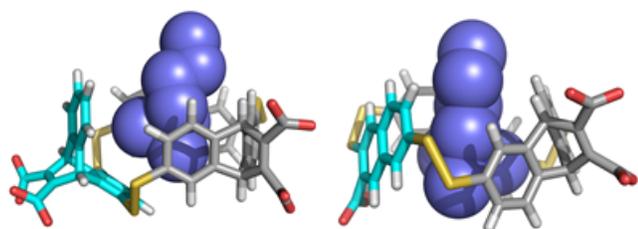


Figure 7. Molecular models of **A₂N** (left) and **A₂G** (right) bound to butyl-trimethyl ammonium (blue) as a model for KMe3. **A** is shown in gray and **N** and **G** are shown in cyan.

ITC revealed that **A₂G** binds to KMe3 ($K_d = 1.4 \pm 0.1 \mu\text{M}$) about two-fold tighter than **A₂B**. Additionally, like **A₂N**, the deepened binding pocket significantly improved selectivity for KMe3 over KMe2 by a factor of four compared to **A₂B**. Binding of KMe was similar to KMe2 (Table 2, Entries 7 and 8), which may be due to poor fit of the weaker binders inside the pocket. Lys binding is 40-fold weaker than binding to KMe3, providing the greatest selectivity yet between KMe3 and unmodified Lys. This selectivity is proposed to come from the increased surface area of the pocket and significant cost of desolvation of unmodified Lys.

Table 2. Thermodynamic binding data obtained for binding of **A₂N** and **A₂G** to Ac-WGGG-QTARK(Me)_nSTG-NH₂ as measured by ITC.^a

Entry	Receptor ^b	Peptide	K_d^c (μM)	Selectivity ^d	ΔG^e (kcal/mol)
1 ^e	A₂N	KMe3	0.30 ± 0.04	-	-8.91 ± 0.07
2 ^e	A₂N	KMe2	4.1 ± 0.5	14	-7.36 ± 0.04
3 ^e	A₂N	KMe	40 ± 4	130	-6.01 ± 0.06
4 ^e	A₂N	Lys	10.5 ± 0.9	35	-6.80 ± 0.05
5 ^e	A₂N	R8GKMe3	1.3 ± 0.2	-	-8.05 ± 0.08
6 ^f	A₂G	KMe3	1.4 ± 0.1	-	-8.00 ± 0.05
7 ^{f,g}	A₂G	KMe2	13.2 ± 2.4	10	-6.6 ± 0.1
8 ^f	A₂G	KMe	15 ± 1	11	-6.57 ± 0.04
9 ^{f,h}	A₂G	Lys	>58	>40	< -5.8
10 ^{f,g}	A₂G	R8GKMe3	5.4 ± 0.1	-	-7.19 ± 0.01

(a) All data determined by ITC, fit to one-site binding model; Conditions: 26 °C, in 10 mM sodium borate buffer, pH 8.5. (b) **A₂N** was measured as the *meso*-species, **A₂G** as a mixture of isomers. (c) Errors are from averages of three trials, unless noted otherwise. (d) Selectivity is calculated as the fold difference in affinity for KMe3 over the designated methylation state of the peptide in that row. (e) Data reported by Pinkin and Waters²¹ (f) Average of two trials. (g) Error determined by propagation from curve fitting and averages. (h) These values are approximate because the c-value for these experiments was <1.

Comparison of R8GKMe3 reveals that Arg contributes 0.8 kcal/mol of binding affinity for **A₂G** binding to KMe3. This is smaller than its contribution to binding of K9Me3 by the shallower pockets of **A₂B**, **A₂C**, and **A₂E**, but is consistent with **A₂N**. This suggests that the deep binding pockets are more selective for the targeted amino acid, while receptors with shallower

binding pockets are more promiscuous and garner affinity from electrostatic contacts in close proximity to the pocket.

Conclusions

In summary, we have probed a variety of structural effects in a conserved mercaptophane structure (**A₂X**) to explore the contribution of non-covalent interactions to molecular recognition of methylated Lys in water. DCC allowed for the rapid synthesis of several receptors that displayed one unique change to the structure, giving direct correlation between the functionality present and its effect on affinity and selectivity towards methylated lysine.

In receptors with electrostatic functionality close to the binding pocket, namely benzoic acid derivatives with one or two carboxylates, we found that increased charge served to increase the overall affinity to peptides bearing various methylation states of lysine. **A₂E** bound to KMe3 with 200 nM affinity, the best in this series, though selectivity was unchanged. Varying the charge on the peptide had a similar magnitude effect.

In receptors with a single carboxylate on **X**, moving the carboxylate to the *ortho* position between the two thiols in **A₂C** resulted in more favorable binding to KMe2, with no effect on binding to KMe3, KMe, or Lys. This result suggests a hydrogen bond interaction between KMe₂ and **C** taking place inside the pocket, analogous to the binding motifs found in many of nature's reader proteins for these marks.^{9,32}

We also explored the influence of a deeper binding pocket as found in some reader proteins.^{9,32} While reader proteins with surface groove binding pockets often display little selectivity between KMe2 and KMe3, proteins with deep binding pockets often display higher affinity and selectivity.^{9,12} By increasing the π -surface or size of the **X** monomer, we were able to create a deepened cavity for methylated lysine in **A₂G** and **A₂N**. This motif indeed results in both higher affinity and higher selectivity for trimethyl lysine relative to the parent receptor, **A₂B**. **A₂G** and **A₂N** both provide an impressive > 10-fold selectivity between KMe3 and KMe2, and a >35-fold selectivity over Lys. This increase in both affinity and selectivity likely arises through a combination of increased contacts with trimethyl lysine and higher cost of desolvation disfavoring the lower methylation states.

These receptors recapitulate the recognition mechanisms found in reader proteins, and even exhibit equal or better affinities despite the fact that reader proteins generally also bind the surrounding sequence. The synthetic receptors do not yet achieve protein-like selectivity over unmodified Lys, presumably due to the greater dependence on electrostatic interactions in the synthetic receptors.

This work demonstrates the utility of DCC for thoughtfully exploring structural motifs and their impact on molecular recognition in water. The ability to rapidly prepare receptors with slight structural changes allows subsequent design to incorporate the most advantageous structural features, fueling new receptor discovery and development with a focus on fur-

ther understanding how affinity and selectivity in water arises. Future work includes application of these receptors to sensing of methyl lysine.

Experimental

Detailed synthetic procedures for monomers **H** and **J** can be found in the ESI. Monomers **C**, **E**, and **G** were prepared following published procedures.^{20, 33, 34}

DCLs were prepared from stock solutions of relevant monomers in 50 mM sodium borate buffer, pH 8.5. Library members were combined to give final concentrations of 0.25 mM – 5 mM for each monomer, depending on the library. Guest concentration (Ac-K(Me)_nGGY-NH₂) was equal to the total monomer concentration. Libraries were prepared at a total volume of 400 μL and allowed to oxidize and equilibrate for up to three weeks. At various time intervals, 100 μL aliquots were removed, filtered with 0.22 μm PVDF syringe filter, and monitored by analytical reverse phase HPLC at 280 nm with an Atlantis T3 4.7 x 150 mm 5 μm C-18 column and buffered mobile phase to ensure library member solubility. Libraries were analyzed using optimized gradients of A and B (A: 100% H₂O, 10 mM NH₄OAc; B: 90% CH₃CN, 10% H₂O, 10 mM NH₄OAc).

Analytical LC/MS was performed on an Agilent Rapid Resolution LC-MSD system, equipped with an online degasser, binary pump, autosampler, heated column compartment, and diode array detector. All separations were performed at 45 °C in optimized gradients with mobile phases of H₂O (5 mM NH₄OAc) and CH₃CN (95% CH₃CN, 5% H₂O, 5 mM NH₄OAc) at pH 5.5 on a Zorbax Extend C18 (4.6 Å 2.1 x 50 mm, 1.8 micron). The MS was performed using a single quad mass spectrometer and all peaks were identified by negative electrospray ionization. Data analysis was performed using the software Agilent ChemStation.

ITC titrations were performed on a Microcal AutoITC200. Titrations were carried out at 298 K in buffered H₂O (10 mM sodium borate buffer, pH 8.5). The concentration of receptor was determined by measuring the UV absorbance using a NanoDrop2000 with a xenon flash lamp, 2048 element linear silicon CCD array detector, and 1 mm path length. A 1–2 mM solution of peptide was titrated into a 65–230 μM solution of receptor, using 2.0 μL increments every 3 minutes. Heats of dilution were subtracted prior to fitting using either the extrapolation method of subtraction or the direct method of subtraction. Binding curves were produced using the supplied Origin software and fit using one-site binding models.

Acknowledgements

This material is based upon work supported by the National Science Foundation under grant no. CHE-1306977, and the National Science Foundation Graduate Research Fellowship to J.E.B under grant no. DGE-1144081.

Notes and references

^a Corresponding Author: mlwaters@email.unc.edu; Department of Chemistry, CB 3290, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599

§ JEB and BCP contributed equally to this work

‡ **A**₂**E** was amplified with Ac-KMe3G-NH₂, as previously reported in reference 18.

† Electronic Supplementary Information (ESI) available. See DOI: 10.1039/b000000x/

1. E. A. Kataev and C. Mueller, *Tetrahedron*, 2014, **70**, 137-167
2. G. V. Oshovsky, D. N. Reinhoudt, and W. Verboom, *Angew. Chem. Int. Ed.*, 2007, **46**, 2366-2393.
3. E. V. Anslyn and D. A. Dougherty, *Modern Physical Organic Chemistry*, University Science Books, California, 2006, 164-165.
4. H. Schneider, *Angew. Chem. Int. Ed.* 2009, **48**, 3924-3977
5. T. Ghosh, S. Garde, and A. E. Garcia, *Biophys. J.*, 2003, **85**, 3187-3193.
6. D. N. Marti and H. R. Bosshard, *J. Mol. Biol.*, 2003, **330**, 621-637.
7. W. Blokzijl and J. B. F. N. Engberts, *Angew. Chem. Int. Ed.*, 1993, **32**, 1545-1579
8. W. M. Nau, M. Florea, and K. I. Assaf, *Isr. J. Chem.*, 2011, **51**, 559-577.
9. S. D. Taverna, H. Li, A. J. Ruthenburg, C. D. Allis, and D. J. Patel, *Nat. Struct. Mol. Biol.*, 2007, **14**, 1025-1039
10. B. D. Strahl and C. D. Allis, *Nature*, 2000, **403**, 41-45
11. A. J. Bannister and T. Kouzarides, *Cell Res.* 2011, **21**, 381-395.
12. M. Yun, J. Wu, J. L. Workman, and B. Li, *Cell Res.*, 2011, **21**, 564-578.
13. S. B. Hake, A. Xiao, and C. D. Allis, *Br. J. Cancer*, 2004, **90**, 761-769
14. G. G. Wang, C. D. Allis, and P. Chi, *Trends. Mol. Med.*, 2007, **13**, 363-372
15. W. Timp and A. P. Feinberg, *Nat. Rev. Cancer*, 2013, **13**, 497-510
16. M. A. Adams-Cioaba and J. Min, *Biochem. Cell. Biol.*, 2009, **87**, 93-105.
17. R. J. Eisert and M. L. Waters, *ChemBioChem*, 2011, **12**, 2786-2790.
18. F. B. L. Cougnon and J. K. M. Sanders, *Acc. Chem. Res.*, 2012, **45**, 2211-2221.
19. For other approaches to develop synthetic receptors for KMe₃, see: (a) K. D. Daze, M. C. F. Ma, F. Pineux and F. Hof, *Org. Lett.*, 2012, **14**, 1512–1515; (b) K. D. Daze and F. Hof, *Acc. Chem. Res.*, 2013, **46**, 937–945; (c) M. A. Gamal-Eldin and D. H. Macartney, *Org. Biomol. Chem.*, 2013, **11**, 488–495.
20. L. A. Ingerman, M. E. Cuellar, and M. L. Waters, *Chem. Commun.*, 2010, **46**, 1839-1841.
21. N. K. Pinkin and M. L. Waters, *Org. Biomol. Chem.*, 2014, **12**, 7059-7067.
22. S. Otto, R. L. E. Furlan, and J. K. M. Sanders, *Science*, 2002, **297**, 590-593.
23. S. Hamieh, V. Saggiomo, P. Nowak, E. Mattia, R. F. Ludlow, and S. Otto, *Angew. Chem. Int. Ed.*, 2013, **52**, 12368-12372
24. L. I. James, J. E. Beaver, N. W. Rice, and M. L. Waters, *J. Am. Chem. Soc.*, 2013, **135**, 6450-6455.
25. L. A. Ingerman, PhD Thesis, The University of North Carolina at Chapel Hill, 2010.

Journal Name

26. B. Ciani, M. Jourdan, and M. S. Searle, *J. Am. Chem. Soc.*, 2003, **125**, 9038-9047.
27. S. M. Ngola, P. C. Kearney, S. Mecozzi, K. Russell, and D. A. Dougherty, *J. Am. Chem. Soc.*, 1999, **121**, 1192-1201.
28. J. C. Ma and D. A. Dougherty, *Chem. Rev.*, 1997, **97**, 1303-1324.
29. R. M. Hughes, K. R. Wiggins, S. Khorasanizadeh, and M. L. Waters, *Proc. Natl. Acad. Sci.*, 2007, **104**, 11184-11188.
30. R. M. Hughes and M. L. Waters, *J. Am. Chem. Soc.*, 2006, **128**, 13586-13591.
31. D. H. Leung, R. G. Bergman, and K. N. Raymond, *J. Am. Chem. Soc.*, 2008, **130**, 2798-2805.
32. D. J. Patel and Z. Wang, *Annu. Rev. Biochem.*, 2013, **82**, 81-118.
33. L. Vial, R. F. Ludlow, J. Leclaire, R. Perez-Fernandez, and S. Otto, *J. Am. Chem. Soc.* 2006, **128**, 10253-10257
34. K. R. West, R. F. Ludlow, P. T. Corbett, P. Besenius, F. M. Mansfeld, P. A. G. Cormack, D. C. Sherrington, J. M. Goodman, M. C. A. Stuart, and S. Otto. *J. Am. Chem. Soc.*, 2008, **130**, 10834-10835.