Late Stage Modification of Trimethyllysine Receptors Identified from Dynamic Combinatorial Libraries

Journal:	*Organic & Biomolecular Chemistry*
Manuscript ID	OB-ART-08-2015-001649.R1
Article Type:	Paper
Date Submitted by the Author:	09-Sep-2015

Complete List of Authors:	Pinkin, Nicholas; The University of North Carolina at Chapel Hill, Department of Chemistry Power, Amanie; The University of North Carolina at Chapel Hill, Department of Chemistry Waters, Marcey; The University of North Carolina at Chapel Hill, Department of Chemistry

Organic and Biomolecular Chemistry RSCPublishing

ARTICLE

Late Stage Modification of Receptors Identified from Dynamic Combinatorial Libraries

Cite this: DOI: 10.1039/x0xx00000x

Nicholas K. Pinkin, Amanie Power, Marcey L. Waters[a]

Received ooth January 2012,
Accepted ooth January 2012

DOI: 10.1039/x0xx00000x

www.rsc.org/

Small molecule receptors are attractive potential sensors of post-translational modifications, including methylated lysine and methylated arginine. Using dynamic combinatorial chemistry (DCC), our lab previously identified a suite of receptors that bind to Kme_3 with a range of affinities ranging from low micromolar to high nanomolar, each with a unique selectivity for Kme_3 over the lower methylation states. To enable these receptors to have broad application as Kme_3 sensors, we have developed a method for their late-stage modification, which we used to synthesize biotinylated derivatives of **A_2B**, **A_2D**, and **A_2G** in a single step. For our most attractive receptor for applications, **A_2N**, we needed to develop an alternative method for its selective functionalization, which we achieved by "activating" the carboxylic acids on the constituent monomer **A** or **N** by pre-functionalizing them with glycine (Gly). Using the resulting **Gly-A** and **Gly-N** monomers, we synthesized the novel **A_2N** variants **A_2Gly-N**, **Gly-A_2N**, and **Gly-A_2Gly-N**, which enabled the late stage biotinylation of **A_2N** wherever Gly was incorporated. Finally, we performed ITC and NMR binding experiments to study the effect that carboxylate spacing has on the affinity and selectivity of **A_2Gly-N** and **Gly-A_2N** for Kme_X guests compared to **A_2N**. These studies revealed the proximity of the carboxylates to play a complex role in the molecular recognition event, despite their positioning on the outside of the receptor.

Introduction

Since the hypothesis of the histone code by Strahl and Allis at the turn of the 21st century,[1] significant progress has been made toward understanding the complex machinery that enables the reading, writing, and erasing of histone post-translational modifications (PTMs) and the associated consequences for gene expression.[2–10] While the majority of advancements can be attributed to significant advances in antibody and mass spectrometry (MS) approaches,[11–15] we and others envisioned that synthetic receptors could offer several advantages for the study of PTMs, particularly methylated lysine (Lys) and arginine (Arg).[16–24] Although generally weaker binders than antibodies, synthetic receptors are typically simple and inexpensive to produce, they have well understood molecular structure and they offer complete batch-to-batch reliability. Already, an assortment of applications has been reported using synthetic receptors to study PTMs, some of which are only possible due to their unique properties compared to antibodies.[21–24] As the applications of synthetic receptors for studying PTMs continue to advance, these tools may enable new approaches for understanding PTMs, which may allow questions to be answered that remain difficult to address using the current tools available.

Utilizing dynamic combinatorial chemistry (DCC),[25–27] our lab has discovered a number of receptors with varied affinity and selectivity for Kme_3 over the lower Lys methylation states (Figure 1), with binding affinities ranging from the low micromolar to high nanomolar range, as well as the first reported receptor for asymmetric dimethylarginine ($aRMe_2$).[16–19]

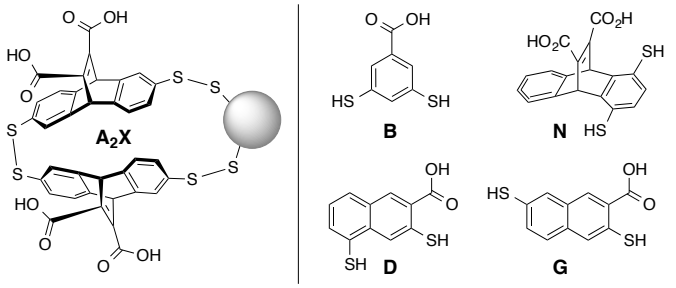


Figure 1. Receptors identified from our lab using dynamic combinatorial chemistry that bind selectively to trimethyllysine or asymmetric dimethylarginine.

To be applicable as probes for Kme_3 sensing, it is necessary to functionalize these receptors with a tag that enables readout of Kme_3 binding. This could be achieved directly by the attachment of a reporter molecule (e.g. fluorescent dye), or indirectly through the attachment of a recognition motif such as biotin, which is commonly used as an affinity tag due to its picomolar affinity to the proteins avidin and streptavidin. Herein, we report a method that enables the rapid mono-functionalization of **A_2B**, **A_2D**, and **A_2G**, which we demonstrate for the synthesis of biotinylated derivatives of each receptor (Figure 2). This method relies on the reduced reactivity of the carboxylic acids on monomer **A** using standard amide coupling reagents, enabling the selective modification of monomers **B**, **D**, and **G** in the assembled receptors.

Because the carboxylates on monomer **N** are similar to those on monomer **A**, **A_2N** cannot be directly functionalized in the same manner. Instead, we also developed a method to pre-functionalize the carboxylic acids on monomers **A** and **N**, which we utilized to distance the carboxylates from the ethanoanthracene bridge using glycine (Gly). Using the modified monomers, **Gly-A** and **Gly-N**, in combination with unmodified **A** and **N**, we show that the novel receptors **A_2Gly-N**, **Gly-A_2N**, and **Gly-A_2Gly-N** can all be accessed using dynamic combinatorial libraries (DCLs). These derivatives enable selective downstream modification of only the Gly-spaced monomers in assembled receptors, as the Gly carboxylates are reactive using the method for the mono-functionalization of **A_2B**, **A_2D**, and **A_2G**. Using this approach, we demonstrate the di-biotinylation of **A_2N** using **A_2Gly-N**. Lastly, using the Gly-functionalized derivatives of **A_2N**, we investigated the role of the proximity of the charge for guest binding using isothermal titration calorimetry (ITC) and NMR.

Results and Discussion.

Mono-Functionalization of Receptors.

The modification of **A_2B**, **A_2D**, and **A_2G** was achieved using standard amide coupling conditions, using diisopropylcarbodiimide (DIC), N-hydroxysuccinimide (NHS) and diisopropylethylamine (DIPEA). The coupling is carried out in anhydrous DMF for 48 hours, after which the solvent is removed and the modified receptors are purified by reverse phase HPLC (Figure 2). Using this method, biotinylated derivatives of each receptor were synthesized using a short polyethylene glycol (PEG) diamine linker (Biotin-PEG_2-NH_2). Using a 10-fold excess of this linker, **A_2B-Biotin**, **A_2D-Biotin**, and **A_2G-Biotin** were prepared in approximately 46%, 13%, and 17% yield in a single step, using RP-HPLC to isolate the modified receptors (see SI, yields estimated from HPLC peak areas). The poorer yields observed for **A_2D** and **A_2G** can likely be attributed to greater steric constraint of the carboxylates due to their position *ortho* to a disulfide instead of *meta*, as in **A_2B**. Although biotin could be pre-attached to monomers **B**, **D**, and **G**, and these modified monomers used to assemble the modified receptors in DCLs, late stage functionalization of the purified receptors enables potentially any marker of interest to be conjugated to the receptor in a single step. In addition to being convenient, this method enables expensive reagents such as fluorophores to be conjugated with little waste of reagent. To demonstrate this point, we coupled the commercially available Biotin-PEG_{11}-NH_2 reagent to **A_2B** using a 7.5-fold excess of reagent, which cleanly provided **A_2B-PEG_{11}-Biotin** in 60% yield (estimated from HPLC peak area) using the same coupling conditions. Due to the long length and the monodisperse nature of this PEG_{11} derivative, this reagent is expensive and cannot be synthesized and purified as easily as the PEG_2 derivative. Thus, the ability to directly attach it to the receptors in a single step allows a minimal amount of the reagent to be used.

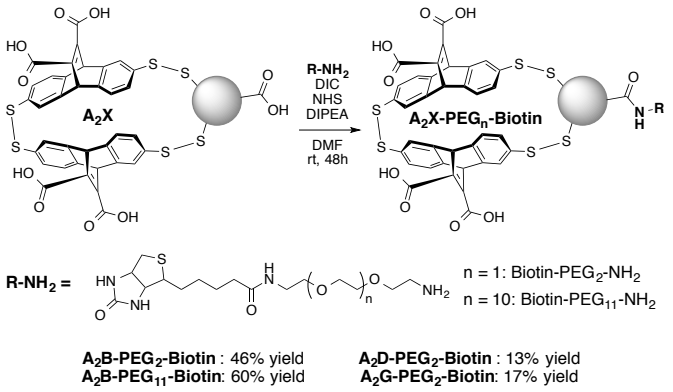


Figure 2. Approach for mono-functionalization of **A_2B**, **A_2D**, and **A_2G** with Biotin-PEG-NH_2. Conditions: R-NH_2 = 7.5 eq. (PEG_{11}) or 10 eq. (PEG_2); DIC = 5 eq. (PEG_{11}) or 7.5 eq. (PEG_2); NHS: 5 eq. (PEG_{11}) or 7.5 eq. (PEG_2); DIPEA = 10 eq. (PEG_{11}) or 12 eq. (PEG_2). Yields provided are estimated from peak area of crude HPLC traces.

Modification of A_2N.

Unlike **A_2B**, **A_2D** and **A_2F**, **A_2N** is unreactive under the same DIC coupling conditions. This is not surprising, as the carboxylates on monomers **A** and **N** are similar, and monomer **A** is not modified by DIC. Because **A_2N** binds to Kme_3 with the best combination of affinity and selectivity of any receptor we have discovered, we needed an alternative approach for the attachment of biotin that would enable **A_2N** to also have application for Kme_3 sensing.

Due to the similarity of the carboxylic acids on **A** and **N**, we focused on developing methods to modify the carboxylates of **A** and **N** prior to use in DCLs (Figure 3a). We began by protecting the thiols with triphenylmethyl (trityl) protecting groups (**Trt-A/N**), which are installed by stirring the monomer and two equivalents of triphenylmethanol in 95:5 trifluoroacetic acid (TFA)/dichloromethane (DCM). Our initial attempts at modifying the carboxylates focused on single step couplings using traditional amide coupling reagents. Like others before us,[28,29] we observed poor reactivity of the carboxylates on **A** and **N**, which has been suggested to be due to steric constraints at the bridging olefin position. Thus, we were surprised to discover that both monomers can easily be converted to their corresponding N-hydroxysuccinimidyl (NHS) esters by

coupling NHS using dicyclohexylcarbodiimide (DCC) in DCM (**Trt-A/N-OSu**). In their activated ester form, **A** and **N** can both be coupled to a variety of amines simply by stirring the two components together in DCM.[30] If the amine is used as its acid salt, an equivalent of DIPEA is needed for the reaction to occur.

With the goal of making **A_2N** reactive toward modification after isolation from a DCL, we envisioned that if we simply spaced the carboxylates from the bridgehead position, they would become more reactive under standard coupling conditions. To test this idea, we coupled glycine methyl ester to **Trt-A/N-OSu** by stirring with four equivalents of the amino acid (as its hydrochloride salt) and four equivalents of DIPEA in DCM. Once coupled, the **Trt-A/N-Gly-OMe** monomers were deprotected in two steps, starting with the removal of the trityl groups by stirring in 5:95 TFA/DCM with an excess of triisopropylsilane (TIPS) to give **Gly-A/N-OMe**. The methyl esters were then hydrolyzed using LiOH in H_2O to furnish the final **Gly-A/N** monomers.

Using a combination of the modified **Gly-A/N** monomers and unmodified monomers, **A_2Gly-N**, **Gly-A_2N**, and **Gly-A_2Gly-N** were all synthesized via preparative DCLs using butyltrimethylammonium iodide ($BuNMe_3^+$) as a guest template in 50 mM borate buffer (Figure 3b). After five days, the receptors were purified by RP-HPLC, giving combined yields of the three isomers of 40%, 35% and 53% for **A_2Gly-N**, **Gly-A_2N** and **Gly-A_2Gly-N**, respectively (estimated from HPLC peak area). This is similar to the combined yield previously observed for **A_2N** of 45%,[18] indicating that the Gly substitutions do not significantly influence the amplification of the receptor under these conditions. Interestingly, the Gly modifications had varied effects on the resolution of the *rac*-, *meso*$_1$-, and *meso*$_2$- isomers of each receptor: **A_2Gly-N** eluted similarly to **A_2N**, with the *rac*- and *meso*$_1$- isomers co-eluting first, and the *meso*$_2$- isomer eluting afterward; all isomers of **Gly-A_2N** nearly co-eluted, making it impossible to isolate a single isomer; and all three of the isomers of **Gly-A_2Gly-N** were well resolved, making it simple to isolate each one individually. As predicted, the Gly spacing enabled Biotin-PEG-NH_2 to be coupled using the same coupling conditions described in Figure 2. Using this approach, **A_2N-Biotin** was prepared in a single step from **A_2Gly-N** in 48 % yield (estimated from HPLC peak area). Importantly, this approach allowed the selective modification of monomer **N** over monomer **A**.

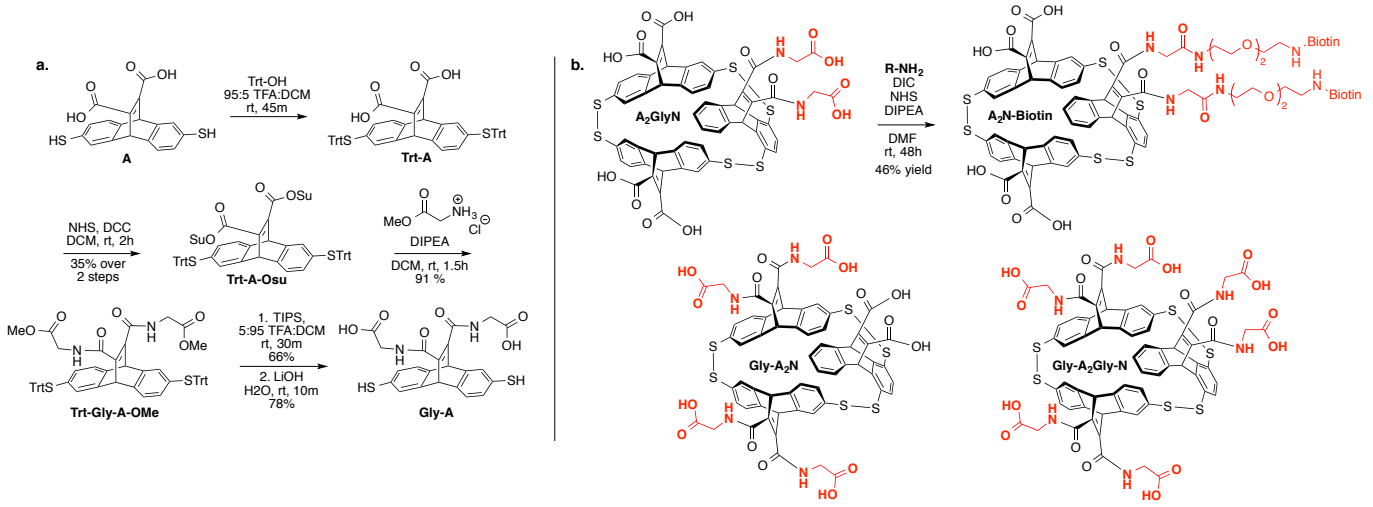


Figure 3. (a) General synthesis of glycine spaced monomers shown for **A**, but identical for monomer **N**. (b) New receptors **A_2Gly-N**, **Gly-A_2N**, and **Gly-A_2Gly-N** synthesized using the Gly-spaced monomers. Using the same coupling conditions previously optimized for **A_2B**, **A_2D**, and **A_2G**, **(Biotin-PEG_2)$_2$-A_2Gly-N** was synthesized from **A_2Gly-N** (top).

Binding Studies Using Gly-spaced Receptors.

A_2N contains a deep aromatic binding pocket that complements the larger, more hydrophobic Kme_3 over the lower Lys methylation states (Figure 4) and provides the highest affinity binding as well as the greatest selectivity of the receptors we have developed.[18] The six carboxylic acids are necessary for water solubility, but it was unclear what role, if any, their charge played in the binding of the cationic ammoniums inside the aromatic pocket. We previously observed that neighboring Arg and Lys residues could directly affect the affinity and selectivity of **A_2N** for a primary site of Kme_3 binding,[31] which we attributed to favorable electrostatic interactions between the secondary basic residues and the carboxylates.

With the Gly-spaced receptors, we aimed to address the questions regarding (1) the contribution the carboxylates toward binding inside of the pocket (primary interaction) and (2) the importance of their proximity to the binding pocket for secondary interactions with Arg outside of the pocket. This was achieved by comparing the binding interactions of the Gly-spaced derivatives of **A_2N** to model peptide guests using isothermal titration calorimetry (ITC). For comparison to **A_2N**, we used a peptide containing residues 4-12 of the histone 3 (H3) tail and an N-terminal WGGG- tag for concentration determination (Ac-WGGG-QTAR**Kme_X**STG-NH_2, X=0-3),

which we have used previously to study the affinity and selectivity of **A_2N** for Kme_3.[18] To study the contribution of neighboring charge to binding, we also studied binding to an equivalent peptide where the neighboring Arg8 is mutated to Gly (Ac-WGGG-QTAG**Kme_X**STG-NH_2, X=3). While **A_2N** and **A_2Gly-N** were used as single (*meso*$_2$-) isomers in the binding and NMR studies (*vide infra*), we could not isolate the pure *meso*$_2$- isomer of **Gly-A_2N**; instead, a mixture of isomers was used composed of predominantly the *meso*$_2$- isomer.

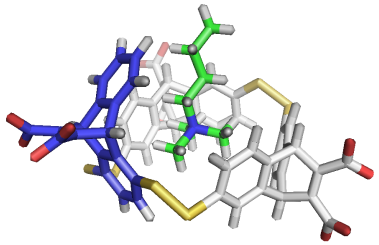


Figure 4. Gas phase minimized model of *meso*$_2$-**A_2N** binding to Kme_3. Monomer **A** is white, monomer **N** is blue, and the carboxylates are shown in red.

Comparison of Affinities and Selectivites.

With increasing spacing of the carboxylates on **A_2N**, the binding affinity for $RKme_2$ and $RKme_1$ are nearly unaffected, while the affinity for $RKme_3$, $RKme_0$, and $GKme_3$ decreased, regardless of whether the Gly residues were on **A** or **N** (Table 1). As a result, **A_2Gly-N** and **Gly-A_2N** are less selective for Kme_3 than **A_2N**. Because the higher affinity of **A_2N** for $RKme_0$ over $RKme_1$ was previously proposed to be due to the engagement of a different mode of binding only possible for $RKme_0$,[18] the loss in affinity of the Gly-spaced derivatives for this peptide suggests that the spacing changes this mode of binding.

Comparing the binding of each of the receptors to $RKme_3$, there is a decrease in affinity with the initial introduction of two Gly residues on **N** to give **A_2Gly-N** ($\Delta\Delta G$ = 1.1 kcal/mol, Table 1, compare entries 6 and 1), while the subsequent introduction of four Gly residues to give **Gly-A_2N** only caused an additional change of 0.2 kcal/mol ($\Delta\Delta G$ = 1.3 kcal/mol, compare entries 11 and 1), which is within error of the value for **A_2Gly-N**. Because the drop in affinity is not proportional to the number of Gly substitutions, this suggests that a similar mechanism may be weakening the binding of both of the modified receptors to Kme_3. A mechanism that is consistent with this observation is a conformational change that disfavors binding of the large trimethylammonium of Kme_3 inside the aromatic pocket. The binding of Kme and Kme_2 may not be affected due to their smaller size, or because they prefer to bind to a different conformation of the receptor in which the NH groups on Kme and Kme_2 are still able to hydrogen bond with water. In the case of unmodified RK peptide, we had previously proposed that it binds to the exterior of the receptor.[18] Addition of the Gly residues would be expected to influence this mode of binding as well, as is observed (entries 4 vs 9).

Importantly, the data are inconsistent with a through-space electrostatic interaction between the cationic guest and the carboxylates on the exterior of the binding pocket, which was originally proposed by Dougherty in his cyclophane host that consists of two ethenoanthracene units identical to those in monomer A.[28] Despite differences in the host geometries and the guests studied, the proximity of the modified carboxylates to the bound guest in both systems is identical. If the carboxylates interact with the bound guest through a long range electrostatic interaction, spacing with Gly would be expected to weaken this interaction equally for all cationic guests binding to the **A_2N** derivatives, but this is clearly not the case. Further studies are necessary to more fully understand the contribution of such long range interactions in molecular recognition.

As previous studies have shown that the neighboring Arg contributes to binding through interaction with the carboxylates and aromatic rings on the outside of the receptor, we investigated whether spacing of the carboxylates weakens this interaction. We find that for all three receptors, the interaction with Arg provides about 0.8 kcal/mol to the interaction energy (compare entries 1 and 5, 6 and 10, and 11 and 15). While this lack of dependence on spacing could suggest the interaction with Arg is unaffected, it may also be due to the fact that not all the carboxylates in **A_2Gly-N** and **Gly-A_2N** contain the Gly spacer.

Organic and Biomolecular Chemistry RSCPublishing

ARTICLE

Table 1. ITC Binding data for **A_2N**, **Gly-A_2N**, and **A_2Gly-N** binding to H3K9me$_X$ (Ac-WGGG-QTA[R/G]Kme$_X$STG-NH_2).

Entry	Receptor	Peptide	K_d [b] (μM)	Selectivity Factor [c]	ΔG [d] (kcal/mol)
1	A_2N[e]	RKme$_3$	0.30 ± 0.04	-	-8.91 ± 0.07
2	A_2N[e]	RKme$_2$	4.1 ± 0.5	14	-7.36 ± 0.07
3	A_2N[e]	RKme$_1$	40 ± 4	131	-6.01 ± 0.06
4	A_2N[e]	RKme$_0$	10.5 ± 0.9	35	-6.80 ± 0.05
5	A_2N[e]	GKme$_3$	1.3 ± 0.2	4.2	-8.05 ± 0.08
6	A_2Gly-N[e]	RKme$_3$	2.0 ± 0.2	-	-7.80 ± 0.07
7	A_2Gly-N[e]	RKme$_2$	5.3 ± 0.5	2.7	-7.21 ± 0.06
8	A_2Gly-N[e]	RKme$_1$	36 ± 4	18.3	-6.08 ± 0.07
9	A_2Gly-N[e]	RKme$_0$	34 ± 4	17.5	-6.10 ± 0.06
10	A_2Gly-N[e]	GKme$_3$	7 ± 1	3.8	-7.01 ± 0.08
11	Gly-A_2N[f]	RKme$_3$	2.7 ± 0.7	-	-7.6 ± 0.1
12	Gly-A_2N[f]	RKme$_2$	5.5 ± 0.7	2.0	-7.19 ± 0.08
13	Gly-A_2N[f]	RKme$_1$	40 ± 4	14.7	-6.01 ± 0.06
14	Gly-A_2N[f]	RKme$_0$	60 ± 6	21.7	-5.77 ± 0.06
15	Gly-A_2N[f]	GKme$_3$	10 ± 1	3.6	-6.84 ± 0.06

[a] Conditions: 26 °C in 10 mM borate buffer, pH 8.5. [b] Errors are from measurements taken in duplicate or triplicate and are estimated at <10%, unless otherwise noted. [c] Selectivity is calculated as the affinity for Kme$_3$ over the designated methylation state in that row. [d] Errors are propagated from errors in K_d. [e] The pure *meso*$_2$- isomer was used. [f] A mixture of isomers was used containing predominantly the *meso*$_2$- isomer.

NMR Binding Studies.

We also compared the binding properties of **A_2Gly-N**, **Gly-A_2N**, and **A_2N** to Kme$_3$ by comparing the upfield shifting induced by each receptor on the simple guest butyl trimethylammonium (BuNme$_3^+$, Table 2) under saturating conditions. For this model guest, binding to **A_2N** causes approximately the same upfield shifting of the Nme$_3$ protons (2.41 ppm) as was previously observed for the equivalent protons on the peptide guest Ac-Kme$_3$G-NH_2 (2.46 ppm).[18] This indicates that BuNme$_3^+$ is suitable for modelling binding to Kme$_3$. Using the same concentrations of **A_2N**, **A_2Gly-N** and **Gly-A_2N** (again, as a mixture of isomers), less upfield shifting and more significant broadening of the Nme$_3^+$ protons of the guest were observed with increasing Gly incorporation (Figure S35). Due to the differences in affinities of each of these receptors for Kme$_3$ (Table 1), the changes in upfield shifting may partially reflect a different proportion of bound guest in each spectrum, although this would be expected to cause proportional differences in upfield shifting for all affected protons. Instead, the differences in upfield shifting are more significant for the Nme$_3$ and γ-methylene protons compared to the more distant α- and β-methylene protons, suggesting that the guest engages in weaker cation-pi interactions with the Gly-substituted receptors, perhaps due to the inability to access the preferred conformation for optimal binding, as suggested above.

Table 2. Change in chemical shifts (Δδ) observed for BuNme$_3^+$ when bound to **A_2N**, **A_2Gly-N**, and **Gly-A_2N**. The analogous Δδ values previously observed for the peptide guest Ac-Kme$_3$G-NH_2 when bound to **A_2N** are included for reference.[18]

Receptor:		**A_2N**[a]	**A_2N**[b]	**A_2Gly-N**[b]	**Gly-A_2N**[b,c]
Guest:		Ac-Kme$_3$G-NH_2	BuNme$_3^+$	BuNme$_3^+$	BuNme$_3^+$
Δδ (ppm):	Nme$_3$	-2.46	-2.41	-2.32	-1.99
	δ	-3.45	-	-	-
	γ	-3.25	-2.93	-2.80	-2.54
	β	-2.09	-1.40	-1.34	-1.31
	α	-0.60	-0.47	-0.46	-0.49

[a] Conditions: 10 mM borate buffered D_2O (pH 8.67). [b] Conditions: 10 mM borate buffered D_2O (pH 8.5) [c] Used as a mixture of isomers containing predominantly *meso*$_2$-**Gly-A_2N**.

Conclusions

In conclusion, we have developed a straightforward method for the late stage modification of the receptors **A_2B**, **A_2D**, and **A_2G**. Using this method, we demonstrated the facile synthesis of biotinylated derivatives of each receptor containing a short PEG linker, as well as the synthesis of a biotinylated derivative of **A_2B** with a PEG$_{11}$ linker using the commercially available Biotin-PEG$_{11}$-NH_2. This simple method should enable the attachment of nearly any desired functionality to these receptors, which will enable their rapid application to new directions in the field of PTM sensing, which we are actively pursuing. Further, as DCC continues to be a valuable tool for the discovery of new receptors with unique affinities and selectivities for different PTMs, we expect that these methods

will continue to have value in modifying novel receptors that share the **A_2X** framework.

We have also developed a simple new method for the modification of monomers **A** and **N**, which have until this point proven challenging targets for functionalization.[28,29] Using this method, we showed that we could di-functionalize each monomer with Gly to yield **Gly-A** and **Gly-N**, which were used in DCLs to assemble the novel receptors **A_2Gly-N**, **Gly-A_2N**, and **Gly-A_2Gly-N**. As only the Gly-spaced monomer is reactive under the coupling conditions developed for **A_2B**, **A_2D** and **A_2G**, we were able to demonstrate the selective functionalization of **N** in a similar manner by coupling Biotin-PEG_2-NH_2 to **A_2Gly-N** to form **A_2N-Biotin**. Although not included here, this approach also enables monomer **A** to be activated for functionalization using the same conditions in any **A_2X** receptor, enabling the single step pan-functionalization of all carboxylates.

Finally, we used ITC to study the effect of spacing the carboxylates on **A** and **N** using Gly on the binding properties of **A_2Gly-N** and **Gly-A_2N** for Kme_X compared to **A_2N**. Although these carboxylates cannot interact directly with a guest bound in the aromatic pocket due to their positioning on the outside of the receptor, their modification with Gly reduced the affinity of the modified receptors for Kme_3 from high nanomolar for **A_2N** to low micromolar for **A_2Gly-N** and **Gly-A_2N**. We believe that the Gly-spacing causes a conformational change in the receptors that affects binding to the larger trimethylammonium, consistent with NMR studies that indicate less optimal guest binding with **A_2Gly-N** and **Gly-A_2N**. Overall, these studies reveal that synthetic modification of these receptors can influence binding properties, but that they still maintain affinities and selectivities in the useful range in this case. The methods developed here enable new approaches toward the late stage modification of complex macrocyclic receptors discovered from DCLs, and should facilitate the rapid development of new applications for these unique receptors.

Experimental

Detailed synthetic procedures for the modification of the receptors, as well as the preparation of Gly-A and Gly-N can be found in the SI. Biotin-PEG_2-NH_2 has been prepared previously and was synthesized according to published literature procedures, although this compound is also commercially available. Biotin-PEG_{11}-NH_2 was purchased from Quanta Biodesign. **A_2B**, **A_2D**, and **A_2G** were prepared in preparative DCLs as described previously, and were purified by reverse-phase HPLC using gradients between NH_4OAc buffered solvents (A: 10 mM NH_4OAc in H_2O; B: 10 mM NH_4OAc in 9:1 ACN:H_2O) on an Atlantis PrepT3 5 μm 10 x 150 mm C18 column.

A2Gly-N, **Gly-A_2N**, and **Gly-A_2Gly-N** were prepared on a preparative scale by dissolving the corresponding monomers (concentrations can be found in the SI) in 50 mM borate buffer (pH 8.5) with 10 mM butyltrimethylammonium iodide ($BuNme_3^+$) as a template. After 5 days, the receptors were purified by RP-HPLC using a gradient of 0-100% B in 45 minutes (A: 10 mM NH_4OAc in H_2O; B: 10 mM NH_4OAc in 9:1 ACN:H_2O) on an Atlantis PrepT3 5 μm 10 x 150 mm C18 column. After isolation, the receptors were lyophilized for 5-7 days to remove any trace NH_4OAc salts. The concentration of the receptors was determined using the same extinction coefficient determined for A2N previously, 11,665 $M^{-1}cm^{-1}$.

All NMR experiments were performed using a Bruker 400 MHz or Bruker 600 MHz instrument, as noted. Data analysis was performed using Topspin 3.1 software. VT 1D NMRs and NMR binding experiments were collected on a Bruker 600 MHz insturment in 10 mM borate buffererd D_2O (pH 8.67). Proton assignments in the binding studies were made using TOCSY analysis. High resolution mass spectrometry was performed on a Thermo LTQ-FT-ICR mass spectrometer.

Peptides were synthesized on a Tetras Peptide Synthesizer using CLEAR-Amide resin from Peptides International. Peptides were purified by RP-HPLC using Waters X-Bridge C18 columns and gradients between water (A) and acetonitrile (B) containing 0.01% trifluoroacetic acid. They were desalted by RP-HPLC using NH_4OAc buffered solvents on an Atlantis PrepT3 5 μm 10 x 100 mm C18 column, and were lyophilized for 5-7 days to remove all volatile salts

All ITC titrations were performed using a MicroCal Auto-iTC200 at 26 °C. Data analysis was performed using the built in Origin 7 software using a one site binding model. Unless otherwise noted, titrations were performed in duplicate. A 10 mM pH 8.5 sodium borate buffer was used for all experiments. All concentrations were determined using a NanoDrop2000 with a xenon flash lamp, 2048 element linear silicon CCD array detector, and 1 mm path length. ~1.1-2.4 mM solutions of peptide were titrated into ~100-180 μM solutions of receptor using 2 μL injections every 3 minutes. Heats of dilution of peptides were subtracted prior to analysis in Origin.

Acknowledgements

We gratefully acknowledge funding from the W. M. Keck foundation for this work. This material is based in part upon work supported by the National Science Foundation under Grant No. CHE-1306977 and also the National Science Foundation Graduate Research Fellowship to N.K.P. under grant no. DGE-1144081. A.P. was supported by the Biophysical Society's Short Course in Biophysics, funded by the NIH (Grant #2T36GM075791-06).

Notes and References

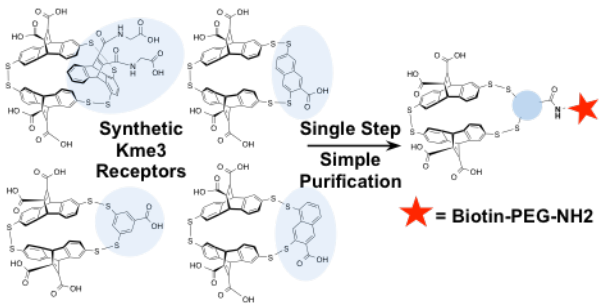
[a] Department of Chemistry, CB 3290, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599. E-mail: mlwaters@unc.edu

Electronic Supplementary Information (ESI) available: Detailed synthetic procedures; NMR characterization; Hi-Res Mass Spec; Receptor characterization; ITC binding data. See DOI: 10.1039/b000000x/

1 B. D. Strahl and C. D. Allis, *Nature*, 2000, **403**, 41–45.
2 E. J. Richards and S. C. R. Elgin, *Cell*, 2002, **108**, 489–500.
3 S. Khorasanizadeh, *Cell*, 2004, **116**, 259–272.

4 C. Martin and Y. Zhang, *Nat. Rev. Mol. Cell Biol.*, 2005, **6**, 838–849.
5 T. Kouzarides, *Cell*, 2007, **128**, 693–705.
6 S. D. Taverna, H. Li, A. J. Ruthenburg, C. D. Allis and D. J. Patel, *Nat. Struct. Mol. Biol.*, 2007, **14**, 1025–1040.
7 B. Li, M. Carey and J. L. Workman, *Cell*, 2007, **128**, 707–719.
8 P. A. Jones and S. B. Baylin, *Cell*, 2007, **128**, 683–692.
9 A. J. Bannister and T. Kouzarides, *Cell Res.*, 2011, **21**, 381–395.
10 S. B. Rothbart and B. D. Strahl, *Biochim. Biophys. Acta - Gene Regul. Mech.*, 2014, **1839**, 627–643.
11 D. E. Schones and K. Zhao, *Nat. Rev. Genet.*, 2008, **9**, 179–191.
12 Y. Zhao and O. N. Jensen, *Proteomics*, 2009, **9**, 4632–4641.
13 C. Chatterjee and T. W. Muir, *J. Biol. Chem.*, 2010, **285**, 11045–11050.
14 L.-M. P. Britton, M. Gonzales-Cope, B. M. Zee and B. A. Garcia, *Expert Rev. Proteomics*, 2011, 8, 631–643.
15 H. Huang, S. Lin, B. a. Garcia and Y. Zhao, *Chem. Rev.*, 2015, 150217145638004.
16 L. A. Ingerman, M. E. Cuellar and M. L. Waters, *Chem. Commun.*, 2010, **46**, 1839–1841.
17 L. I. James, J. E. Beaver, N. W. Rice and M. L. Waters, *J. Am. Chem. Soc.*, 2013, **135**, 6450–6455.
18 N. K. Pinkin and M. L. Waters, *Org. Biomol. Chem.*, 2014, **12**, 7059–67.
19 J. E. Beaver, B. C. Peacor, J. V. Bain, L. I. James and M. L. Waters, *Org. Biomol. Chem.*, 2015, **13**, 3220–3226.
20 C. S. Beshara, C. E. Jones, K. D. Daze, B. J. Lilgert and F. Hof, *ChemBioChem*, 2010, **11**, 63–66.
21 S. A. Minaker, K. D. Daze, M. C. F. Ma and F. Hof, *J. Am. Chem. Soc.*, 2012, 11674–11680.
22 K. D. Daze, T. Pinter, C. S. Beshara, A. Ibraheem, S. A. Minaker, M. C. F. Ma, R. J. M. Courtemanche, R. E. Campbell and F. Hof, *Chem. Sci.*, 2012, **3**, 2695–2699.
23 H. F. Allen, K. D. Daze, T. Shimbo, A. Lai, C. A. Musselman, J. K. Sims, P. A. Wade, F. Hof and T. G. Kutateladze, *Biochem. J.*, 2014, **459**, 505–512.
24 M. Florea, S. Kudithipudi, A. Rei, M. J. Gonzalez-Alvarez, A. Jeltsch and W. M. Nau, *Chem. A Eur. J.*, 2012, 3521–3528.
25 P. T. Corbett, J. Leclaire, L. Vial, K. R. West, J.-L. Wietor, J. K. M. Sanders and S. Otto, *Chem. Rev.*, 2006, **106**, 3652–3711.
26 S. Otto, R. L. E. Furlan and J. K. M. Sanders, *Science*, 2002, **297**, 590–593.
27 P. T. Corbett, J. K. M. Sanders and S. Otto, *Chemistry*, 2008, **14**, 2153–2166.
28 S. M. Ngola, P. C. Kearney, S. Mecozzi, K. Russell and D. a. Dougherty, *J. Am. Chem. Soc.*, 1999, **121**, 1192–1201.
29 R. Marquis, K. Kulikiewicz, S. Lebedkin, M. M. Kappes, C. M. S. Meunier and A. Wagner, *Chem. - A Eur. J.*, 2009, **15**, 11187–11196.
30 N. K. Pinkin, PhD Thesis, University of North Carolina, Chapel Hill, 2015.
31 N. K. Pinkin, I. Liu, J. D. Abron and M. L. Waters, *Chem. Eur. J*, 2015, Accepted.

TOC Graphic:



Approaches for the late-stage modification of receptors discovered from dynamic combinatorial libraries and the investigation of the effects of simple modifications on receptor binding and selectivity.