

A small molecule receptor that selectively recognizes trimethyl lysine in a histone peptide with native protein-like affinity†

Lindsey A. Ingerman,^a Matthew E. Cuellar^b and Marcey L. Waters^{*a}

Received (in Austin, TX, USA) 7th January 2010, Accepted 27th January 2010

First published as an Advance Article on the web 12th February 2010

DOI: 10.1039/c000255k

A small molecule receptor that mimics the HP1 chromodomain's affinity for trimethyl lysine has been identified from a dynamic combinatorial library. Discrimination for trimethyl lysine over the lower methylation states parallels that of the native protein receptor. These studies demonstrate the feasibility of small molecule receptors as potential sensors for protein post-translational modifications.

Understanding the role of post-translationally modified histone proteins in gene expression is crucial to the study of both development and disease, and cancer in particular.¹ Many post-translational modifications (PTMs) have been shown to function by recruiting non-histone proteins to chromatin, dictating the higher-order chromatin structure in which DNA is packaged, and resulting in gene expression or gene silencing depending on the type and location of the PTM.² One such covalent modification is the site-specific methylation of Lys, which can be mono-, di-, or tri-methylated. Depending on the position of the Lys residue and the degree of methylation, different proteins are recruited, resulting in variable transcriptional outcomes.³ A significant protein–protein interaction induced by Lys methylation is the binding of histone 3 (H3) K9Me₃ to the HP1 chromodomain (Fig. 1a), which results in gene silencing. This binding event is the result of the recognition of the trimethylammonium in an aromatic cage *via* cation– π interactions, as well as an extended surface

groove that forms additional sequence specific interactions with the adjacent residues of the histone tail.⁴

Typically antibodies are used to identify PTMs, but these have some inherent limitations; for example, antibodies cannot be used to identify unknown PTMs due to their sequence specificity.⁵ Recently, DNA aptamers have been used to develop protein affinity reagents for acylated lysine PTMs, emphasizing the need for new molecular tools for the recognition of PTMs.⁶ The development of small molecules which mimic chromodomains is a field which has remained largely unexplored.⁷ Synthetic receptors have some significant potential advantages, including better reproducibility, lower cost, lower molecular weight, and the possibility of being used within cells. We report here the use of dynamic combinatorial chemistry (DCC) to identify synthetic receptors for trimethyl lysine (KMe₃) that bind with comparable affinities and selectivities as the native HP1 chromodomain.⁴

DCC allows the formation of a library of hosts under equilibrium conditions.⁸ In the presence of a molecular target such as methylated lysine, favorable host–guest binding interactions drive the synthesis and amplification of the best receptor(s) at the expense of other oligomers. By screening the same library against all Lys methylation states, this method provides a rapid approach to screen for selectivity as well. A range of building blocks (Fig. 2) were investigated. These monomers combine aromatic groups to facilitate cation– π interactions, carboxylates for water solubility, and thiols which can undergo disulfide exchange to establish an equilibrium. Disulfide exchange is compatible with most protein functional groups in aqueous solution at close to neutral pH.⁹ **A** and **B** have been shown previously to form receptors that interact favorably with alkylated ammonium groups.^{10†}

Dynamic combinatorial libraries (DCLs) were prepared by mixing equimolar amounts of 3–4 building blocks in water at pH 8.5. Upon reaching equilibrium, the resulting DCLs were analyzed by LC–MS. The introduction of the dipeptide Ac-KMe₃-G-NH₂ as a template resulted in significant changes

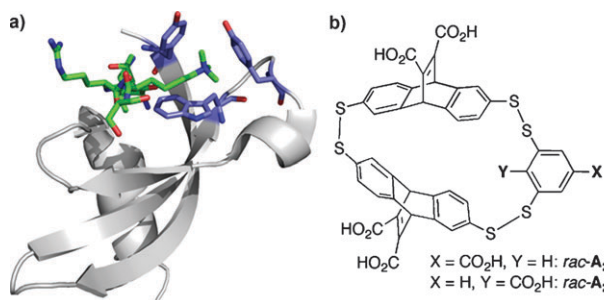


Fig. 1 (a) Structure of the HP1 chromodomain bound to the H3 K9Me₃ peptide (green) with the aromatic pocket shown in blue (pdb: 1KNE). (b) Structure of *rac*-**A2B** and *rac*-**A2C**.

^a Department of Chemistry, CB 3290, University of North Carolina, Chapel Hill, NC 27599, USA. E-mail: mlwaters@unc.edu; Fax: +1 919 962 2388; Tel: +1 919 843 6522

^b Department of Medicinal Chemistry, University of Minnesota, 308 Harvard St. SE, Minneapolis, MN 55455, USA

† Electronic supplementary information (ESI) available: Experimental procedures, LC, MS, fluorescence anisotropy, and ITC data. See DOI: 10.1039/c000255k

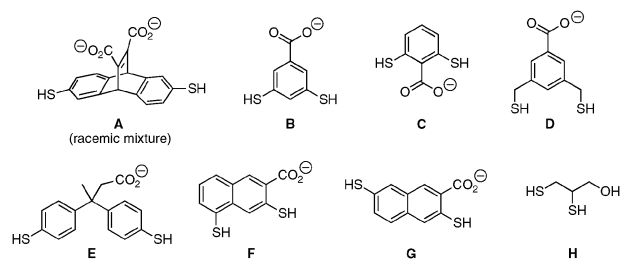


Fig. 2 Monomers used in DCLs.

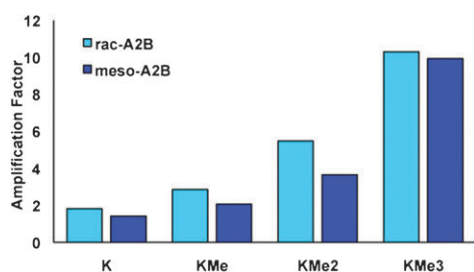


Fig. 3 Extent of amplification of *rac*-**A₂B** and *meso*-**A₂B** with Ac-KMe_n-G-NH₂ guests relative to the untemplated library.

in the composition of several libraries. For example, amplification of receptors *rac*-**A₂B** and *meso*-**A₂B** from one library,[§] as well as **A₂C** and **AC₃** from another, was observed as identified by ESI-MS (Fig. 1b). Interestingly, amplification of only a single isomer of the isomeric receptor **A₂C** was observed. When a library of **A** and **B** was screened in the presence of an 8-residue histone tail sequence containing KMe₃, a similar degree of amplification of *rac*- and *meso*-**A₂B** was observed, indicating that other nearby sidechains do not significantly influence binding.

Using biased libraries in which building blocks **A** and **B** as well as **A** and **C** were each mixed in a 2 : 1 ratio (7.5 mM total), the selectivity of the receptors for different methylation states of lysine was investigated, using Ac-KMe_n-G-NH₂ (*n* = 0–3) as the guests. The amplification of both **A₂B** diastereomers was dependent on the extent of methylation, with approximately 10-fold amplification with KMe₃ and less than 2-fold amplification with Lys (Fig. 3), suggesting significant selectivity. In contrast, no selectivity was observed for receptors **A₂C** and **AC₃**, as similar amplification was observed for all methylation states as well as unmethylated Lys (see ESI†). This lack of selectivity may be the result of additional electrostatic and/or hydrogen bonding interactions with each guest due to the position of the carboxylic acid in **C** as compared to **B**. Because of this lack of selectivity, **A₂C** and **AC₃** were not investigated further.

Both **A₂B** isomers were isolated for further studies using semi-preparative HPLC. Fluorescence anisotropy was used to measure the dissociation constant of *rac*- and *meso*-**A₂B** to a peptide consisting of residues 5–11 of the histone 3 protein (H3 K9Me_n), with each of the methylation states at Lys9, appended with an N-terminal carboxyfluorescein (Fam) for fluorescence detection.¶ The H3 K9Me₃ peptide was found to bind both *rac*- and *meso*-**A₂B** with binding affinities of about

25–30 μM, which is equivalent to its binding affinity to the HP1 chromodomain (Table 1).⁴ Moreover, H3 K9Me₃ binds to both isomers of **A₂B** with greater than 2-fold selectivity over its dimethylated counterpart, exhibiting slightly better selectivity than the native protein.⁴ The mono- and unmethylated peptides exhibited considerably weaker affinity for *rac*-**A₂B**, with no appreciable binding to the unmethylated histone tail. Thus, the trend in *both affinity and selectivity* of **A₂B** for the histone tail peptide parallels that of the native protein.⁴

H3 K9Me_n: Fam-QTAR-KMe_n-STG-NH₂, *n* = 0–3

A mutant H3 tail with Gly in place of Arg8 exhibited comparable affinity to that of the native sequence (Table 1), indicating that the adjacent basic Arg does not contribute significantly to binding. Additionally, Lys9 was mutated to Gly, which resulted in total loss of binding to *rac*-**A₂B**. NMR studies of the KMe₃ dipeptide in the presence of *rac*-**A₂B** exhibit upfield shifting of the lysine methyl groups by 0.83 ppm at 25 °C, indicating that the binding event is indeed occurring at the site of modification.** These findings confirm that the key binding interaction is between the trimethylammonium group and the aromatic pocket, as in the histone–chromodomain protein–protein interaction.

There are likely several factors that result in the observed selectivity of **A₂B**. First, the binding pocket may be too large for Lys and KMe. In addition, Lys, KMe, and KMe₂ can all form hydrogen bonds with water, and binding to the pocket may require some degree of desolvation, which would be unfavorable. Lastly, model systems have shown that cation–π interactions are enhanced with increasing methylation of the ammonium group.¹¹ Thus, KMe₃ is expected to have stronger cation–π interactions with **A₂B** than with the lower methylation states.

In conclusion, we report the identification of a synthetic receptor for KMe₃ that exhibits both comparable affinity and selectivity to the native HP1 chromodomain. The mass of **A₂B** is less than 900 Da, as compared to approximately 6300 Da for the HP1 chromodomain and about 150 kDa for a typical antibody. The comparable binding affinity to the native protein is impressive, particularly given that the synthetic receptor appears to bind only to the KMe₃ sidechain, whereas the chromodomain binds to the surrounding sequence as well. The observed selectivity for higher methylation states with **A₂B** can be attributed to differences in the magnitude of the cation–π interactions, as well as differences in size and desolvation penalty. The difference in selectivity observed

Table 1 Dissociation constants for H3 histone tail peptides with varying methylation states at K9 as determined by fluorescence anisotropy^a

Peptide	<i>rac</i> - A₂B <i>K_d</i> ^b /μM	<i>meso</i> - A₂B <i>K_d</i> ^b /μM	HP1 chromodomain <i>K_d</i> ^c /μM
H3 K9Me ₃	25 ± 3	28 ± 4	10 ^e (21 ± 2 at 25 °C) ^f
H3 K9Me ₂	58 ± 10	73 ± 9	15 ^e (39 ± 7 at 25 °C) ^f
H3 K9Me	166 ± 50	N.A. ^d	96 ^e
H3 K9	> 1200	N.A. ^d	> 1000 ^e
H3 K9Me ₃ R8G	34 ± 8	N.A. ^d	—

^a Conditions: 27 °C, 10 mM phosphate buffer, pH 8.5. ^b Errors are from the fit. ^c Values for *rac*- and *meso*-**A₂B** binding to H3 K9Me₃ with no fluorophore are 20 μM and 13 μM, respectively, as measured by ITC (see ESI†). ^d Not available due to limited material (*meso*-**A₂B**), although binding is expected to be comparable to that of *rac*-**A₂B**. ^e Conditions: 15 °C in pH 7.5 phosphate buffer. Values are taken from ref. 4c. ^f Conditions: 25 °C in 50 mM phosphate buffer, pH 8.0, 25 mM NaCl (Eisert and Waters, unpublished results).

when **B** is replaced with **C** is surprising, and demonstrates that subtle structural changes can have a significant effect on binding. Such sequence-independent receptors have the potential to be widely used in the investigation of unknown protein Lys methylation sites,¹² unlike antibodies which are sequence selective. This work suggests that small molecule receptors identified *via* DCC have a promising future as affinity reagents for PTMs, as they can easily be synthetically modified with a fluorophore, cell-penetrating peptide, or other tag. We are now investigating the application of these receptors to sensing of trimethyl lysine in proteins.

We thank the Army Research Office and Defense Threat Reduction Agency (DTRA) for support (W911NF06-1-0169). We also thank Prof. Michel R. Gagné and James W. Jorgenson (UNC) for helpful discussions. We also gratefully acknowledge Sijbren Otto for donation of monomer **E**.

Notes and references

‡ Monomers **A** and **B** are based on a host originally reported by Dougherty. See ref. 4d.

§ *Rac*- and *meso*-**A₂B** have been identified as receptors for other cationic guests. See ref. 10.

¶ Binding affinities for H3 K9Me₃ lacking the fluorophore were confirmed by ITC (see ESI†).

|| Interestingly, *rac*-**A₂B** binds to H3 K9Me₃ approximately 6-fold more tightly than to other trimethylated ammonium ions or NMe₄⁺ reported previously. See ref. 10.

** No NMR studies were performed with *meso*-**A₂B** because of limited availability of this compound due to significant co-elution with *rac*-**A₂B**.

- 1 (a) S. R. Bhaumik, E. Smith and A. Shilatifard, *Nat. Struct. Mol. Biol.*, 2007, **14**, 1008–1016; (b) G. G. Wang, C. D. Allis and P. Chi, *Trends Mol. Med.*, 2007, **13**, 363–372.

- 2 (a) T. Jenuwein and C. D. Allis, *Science*, 2001, **293**, 1074–1080; (b) T. Kouzarides, *Cell*, 2007, **128**, 693–705.
- 3 C. Martin and Y. Zhang, *Nat. Rev. Mol. Cell Biol.*, 2005, **6**, 838–849.
- 4 (a) S. A. Jacobs and S. Khorasanizadeh, *Science*, 2002, **295**, 2080–2083; (b) P. R. Nielsen, D. Nietlispach, H. R. Mott, J. Callaghan, A. Bannister, T. Kouzarides, A. G. Murzin, N. V. Murzina and E. D. Laue, *Nature*, 2002, **416**, 103–107; (c) R. M. Hughes, K. R. Wiggins, S. Khorasanizadeh and M. L. Waters, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 11184–11188; (d) For a review of cation– π interactions, see: J. C. Ma and D. A. Dougherty, *Chem. Rev.*, 1997, **97**, 1303–1327.
- 5 (a) N. Blow, *Nature*, 2007, **447**, 741–744; (b) C. C. Kazanecki, A. J. Kowalski, T. Ding, S. R. Rittling and D. T. Denhardt, *J. Cell. Biochem.*, 2007, **102**, 925–935.
- 6 B. A. R. Williams, L. Lin, S. M. Lindsay and J. C. Chaput, *J. Am. Chem. Soc.*, 2009, **106**, 6330–6331.
- 7 C. S. Beshara, C. E. Jones, K. D. Daze, B. J. Lilgert and F. Hof, *ChemBioChem*, 2009, **11**, 63–66.
- 8 For reviews, see: (a) 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; (b) J.-M. Lehn, *Chem.–Eur. J.*, 1999, **5**, 2455–2463; (c) S. Ladame, *Org. Biomol. Chem.*, 2008, **6**, 219–226; (d) R. F. Ludlow and S. Otto, *Chem. Soc. Rev.*, 2008, **37**, 101–108; (e) S. Otto, *Curr. Opin. Drug Discovery Dev.*, 2003, **6**, 509–520.
- 9 (a) S. Otto, R. L. E. Furlan and J. K. M. Sanders, *J. Am. Chem. Soc.*, 2000, **122**, 12063–12064; (b) O. Ramström and J.-M. Lehn, *ChemBioChem*, 2000, **1**, 41–48.
- 10 (a) S. Otto, R. L. E. Furlan and J. K. M. Sanders, *Science*, 2002, **297**, 590–593; (b) P. T. Corbett, J. K. M. Sanders and S. Otto, *Chem.–Eur. J.*, 2008, **14**, 2153–2166.
- 11 (a) R. M. Hughes, M. L. Benshoff and M. L. Waters, *Chem.–Eur. J.*, 2007, **13**, 5753–5764; (b) L. M. Salonen, C. Bucher, D. W. Banner, W. Haap, J.-L. Mary, J. Benz, O. Kuster, P. I. Seiler, W. B. Schweizer and F. Diederich, *Angew. Chem., Int. Ed.*, 2009, **48**, 811–814; (c) A. L. Whiting, N. M. Neufeld and F. Hof, *Tetrahedron Lett.*, 2009, **50**, 7035–7037.
- 12 J. Huang and S. L. Berger, *Curr. Opin. Genet. Dev.*, 2008, **18**, 152–158.