Supporting Information for

A Small Molecule Receptor that Selectively Recognizes Trimethyl Lysine in a Histone Peptide with Native Protein-like Affinity

*Lindsey A. Ingerman, Matthew E. Cuellar, and Marcey L. Waters**

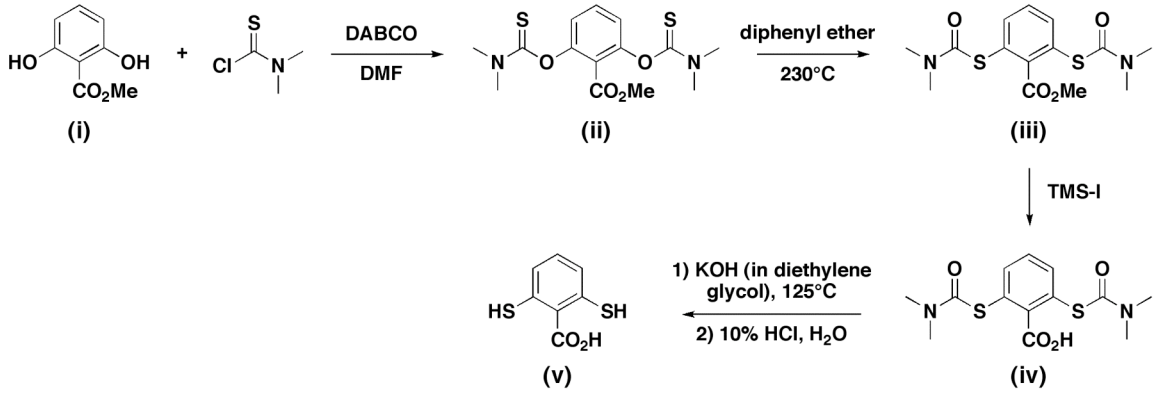
Department of Chemistry, CB 3290, University of North Carolina, Chapel Hill, NC 27599

Peptide Synthesis. All peptide synthesis was performed on a Tetras Peptide Synthesizer using Peptides International CLEAR-Amide resin. Peptides were synthesized on a 0.06 mmole scale. All amino acids with functionality were protected during synthesis. Coupling reagents were HOBt/HBTU in DMF. For the dipeptides, the N-terminus was acylated with a solution of 5% acetic anhydride and 6% 2,6-lutidine in DMF. Peptides synthesized for fluorescence anisotropy were capped with 2 equivalents of 5(6)-Carboxyfluorescein and coupled in the dark with standard coupling reagents overnight. Cleavage was performed by hand with a cocktail of 95% TFA/2.5% triisopropylsilane/2.5% H_2O for 3 hours. Peptides were purified by semipreparative reversed-phase HPLC on a C18 column at a flow rate of 4 mL/min. Peptides were purified with a linear gradient of A and B (A: 95% H_2O/5% CH_3CN with 0.1% TFA, B: 95% CH_3CN/5% H_2O with 0.1% TFA) and elution was monitored at 214 nm. Once purified, peptides were lyophilized to powder and characterized by ESI-MS.

Methylated peptides were synthesized with either 2 equivalents of Fmoc-Lys(Boc)(Me)-OH purchased from BaChem or Fmoc-Lys$(Me)_2$-OH•HCl purchase from Anaspec and coupled for 10-12 hours. The trimethyl lysine-containing peptides were synthesized by reacting the corresponding dimethylated peptides (0.6 mmol scale) prior to cleavage from the resin with MTBD (10.8 μL, 0.075 mmol) and methyl iodide (37.4 μL, 0.6 mmol) in DMF (5 mL) for 5 hours with bubbling N_2 in a peptide synthesis flask stoppered with a vented septum. After washing the resin with DMF (3x), CH_2Cl_2 (3x), and drying, the peptide was cleaved as normal.

Monomer Synthesis. Monomers **A** and **B** were synthesized following the reported procedure of Corbett, Sanders, and Otto.[1] Monomer **D** was synthesized following the reported procedure of Field and Kirrstetter.[2] We gratefully acknowledge donation of monomer **E** from Sijbren Otto. Monomers **G** was synthesized following the reported procedure of Otto and coworkers[3], as well as monomer **F** starting with 3,5-dihydroxynaphthoic acid. Monomer **H** is commercially available.

*Synthesis of Monomer **C**.* Compounds *i*, *ii*, *iii* were synthesized using standard reported procedures.



Scheme 1. Synthetic procedure for building block **C.**

2,6-Bis(dimethylcarbamoylsulfanyl)benzoic acid (iv).[4] A mixture of 2,6-Bis(dimethylcarbamoylsulfanyl)benzoic acid methyl ester *(iii)* (0.16 g, 0.47 mmol) and Me_3SiI (0.36 mL, 2.52 mmol) were refluxed at 105°C overnight. Once cool, diethyl ether (11 mL) was added and washed with 10% NaOH. The basic aqueous layers were combined and 1M HCl was added to acidify the solution. A white crystal-like solid precipitated out and was filtered (0.084 g, 0.26 mmol). ^{1}H NMR (MeOD, 400 MHz): δ = 7.59 (d, 2H), 7.42 (t, 1H), 3.07 (broad d, 12H).

2,6-Dimercaptobenzoic acid (v).[1] Under a nitrogen atmosphere 2,6-Bis(dimethylcarbamoylsulfanyl)benzoic acid *(iv)* (0.037 g, 0.011 mmol) was suspended

in a 1.75 M solution (0.38 mL) of KOH in diethylene glycol that had been purged with argon for 2 hr. The solution was heated at 105°C for 30 min. After the solution had cooled to room temperature, 2.7 mL of purged water was added followed by rapid addition of 10% HCl (0.5 mL). The product was either filtered when possible and washed with hexanes, or diluted further and purified by reverse phase HPLC (0.13 g, 0.070 mmol). ^{1}H NMR (MeOD, 400 MHz): δ = 7.19 (d, 2H), 7.07 (t, 1H).

Dynamic Combinatorial Chemistry. The relevant building blocks were individually dissolved in water, adding sufficient 1.0 M aqueous NaOH to fully deprotonate the thiols and carboxylic acids on the building blocks, using sonication when necessary. The pH of each solution was then adjusted to 8.5 using 1.0 M aqueous HCl and 1.0 M aqueous NaOH. For example, in the unbiased **A•B•D** libraries, aliquots of each monomer solution were combined in a 2 mL LC-MS vial to reach a final concentration of 2.5 mM of each monomer. In the biased $\mathbf{A_2B}$ and $\mathbf{A_2C}$ libraries, aliquots of each monomer solution were combined in a 2 mL LC-MS vial to reach a final concentration of 5 mM **A** and 2.5 mM of **B** or **C** respectively. When necessary, an aliquots of peptide guests dissolved in water were added to the reactions to reach a final concentration of 7.5 mM peptide. Any remaining volume was made up with water. The vials were capped and analyzed at various time points.

Analytical LC-MS. LC-MS was carried out on an Agilent Rapid Resolution LC-MSD system equipped with an online degasser, binary bump, autosampler, heated column compartment, and diode array detector. All separations were performed using 5 mM NH_4OAc H_2O-acetonitrile gradients at pH 5 and a Halo C18 column (4.6 × 100 mm, 2.7 micron). The MS was performed using a single quad mass spectrometer. Mass spectra

(ESI+) were acquired in ultrascan mode by using a drying temperature of 350°C, a nebulizer pressure of 45 psi, a drying gas flow of 10 L/min, and a capillary voltage of 3000 V. The reactions were monitored weekly (3 μL injections) until equilibrium was reached after about 3 weeks. The chromatography of library **A•B•D** (Figure S1) was carried out at 50°C with gradient A (Table SI). The chromatography of library $\mathbf{A_2B}$ (Figure S2) was carried out with gradient B (Table SI) using a gradient temperature, going from 50°C to 40°C, left to right. The chromatography of library $\mathbf{A_2C}$ (Figure S3) was carried out at 50°C with gradient C (Table SI). The peak areas were integrated at 254 nm and the amplification factors were calculated (A.F. = % area of $\mathbf{A_2B}$ in templated DCL / % area of $\mathbf{A_2B}$ in untemplated DCL).

Table S1. Analytical LC methods use to analyze DCC libraries **A•B•D** (Method A), $\mathbf{A_2B}$ (Method B) and $\mathbf{A_2C}$ (Method C).

Gradient A		
Time (min)	**%B**	**Flow Rate (mL/min)**
0.00	0.0	1.0
3.00	30.0	1.0
7.00	32.3	1.0
7.30	32.3	1.0
7.35	32.3	0.6
9.00	32.3	0.6
9.10	32.3	1.0
10.80	34.0	1.0
10.90	100.0	1.0
11.90	100.0	1.0
12.00	0.0	1.0
13.00	0.0	1.0

Gradient B		
Time (min)	**%B**	**Flow Rate (mL/min)**
0.00	3.0	1.0
3.00	30.0	1.0
8.25	33.0	1.0
8.30	33.0	1.5
10.30	34.0	1.5
12.00	50.0	1.5
12.10	100.0	1.5
13.50	100.0	1.5
14.00	3.0	1.0
19.00	3.0	1.0

Gradient C		
Time (min)	**%B**	**Flow Rate (mL/min)**
0.00	15.0	1.5
2.00	30.0	1.5
6.00	30.0	1.5
9.00	50.0	1.5
9.10	100.0	1.5
11.00	100.0	1.5
11.50	15.0	1.5
16.00	15.0	1.5

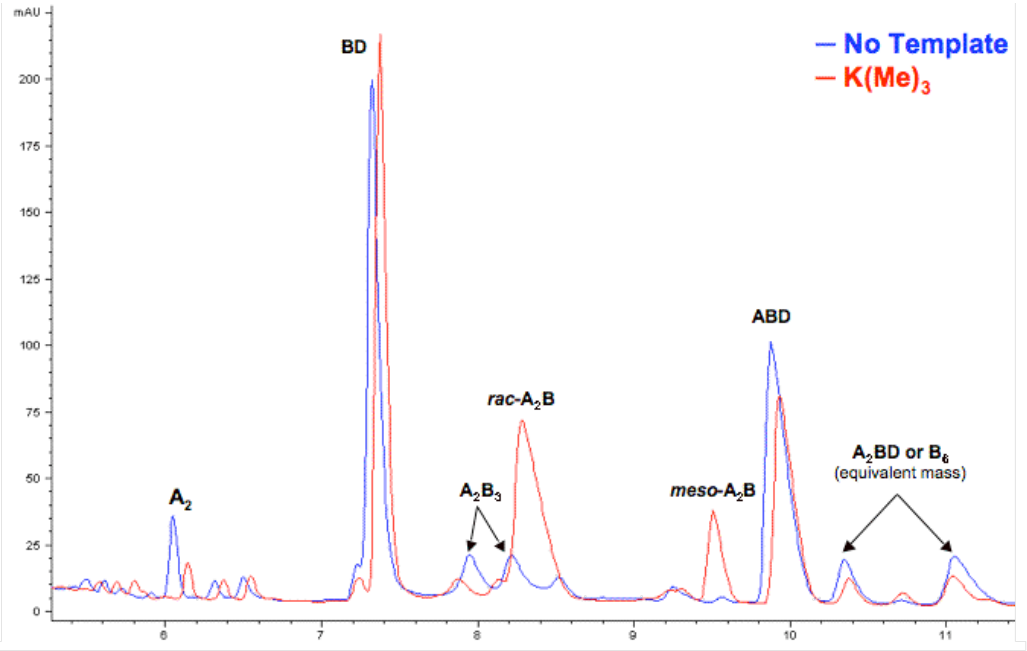


Figure S1. Part of the analytical HPLC trace at 254 nm of a DCL consisting of monomers **A**, **B**, and **D** (2.5 mM each), untemplated (blue) and in the presence of Ac-KMe_3-G-NH_2 (red).

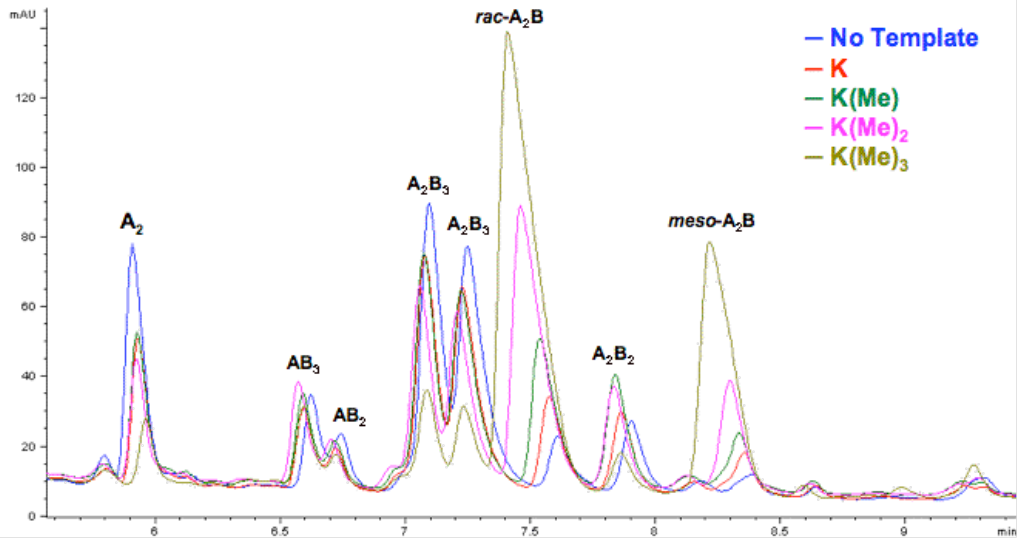


Figure S2. Part of the analytical HPLC trace at 254 nm of a biased DCL consisting of monomers **A** (5 mM) and **B** (2.5 mM), untemplated (blue) and in the presence of all methylation states, Ac-KMe$_n$-G-NH$_2$.

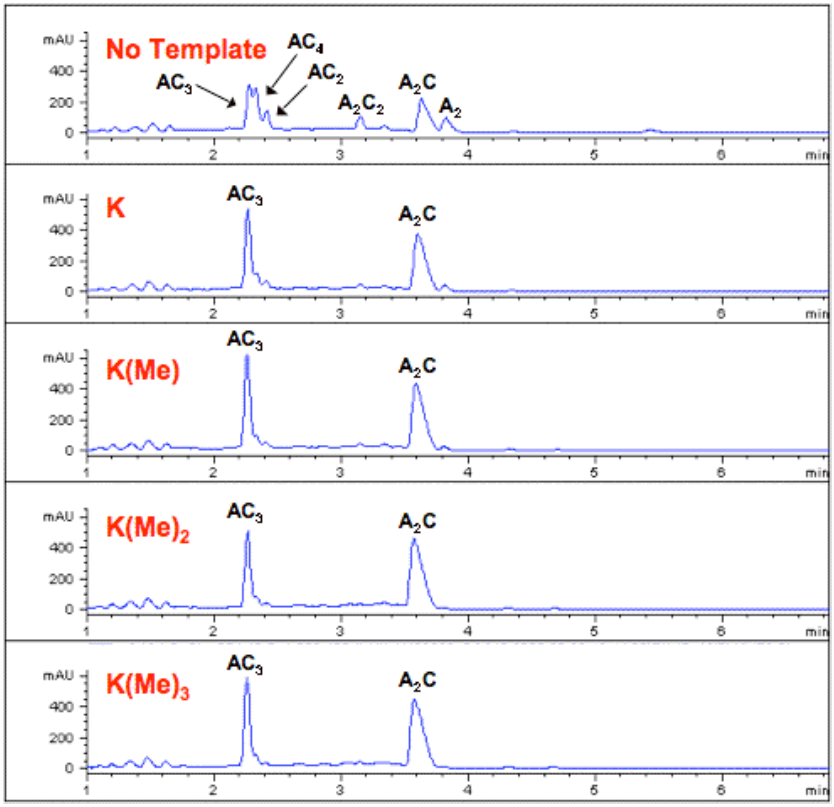


Figure S3. Part of the analytical HPLC trace at 254 nm of a biased DCL consisting of monomers **A** (5 mM) and **C** (2.5 mM), untemplated and in the presence of all methylation states, Ac-KMe$_n$-G-NH$_2$.

***Rac*-A_2B and *meso*-A_2B isolation.** The receptors were isolated by semi-preparative HPLC using 5.0 mM **A**, 2.5 mM **B** and 7.5 mM Ac-KMe_3-G-NH_2. Upon equilibration the library was filtered, and 0.5 mL injections were chromatographed using standard peptide synthesis mobile phases A and B (0-50% B 0-5 min, then held at 50% B 5-20 min) with a flow rate of 4.0 mL/min. Optimal separation was achieved using a column heater set to 40°C. The two **A_2B** peaks from 11.5 – 12.5 minutes were collected separately (Figure S2) and analyzed for purity by analytical LC-MS. A second and sometimes third purification was required to achieve pure *meso*-**A_2B** (Figure S3). Both peaks were indistinguishable by mass (Figure S4).

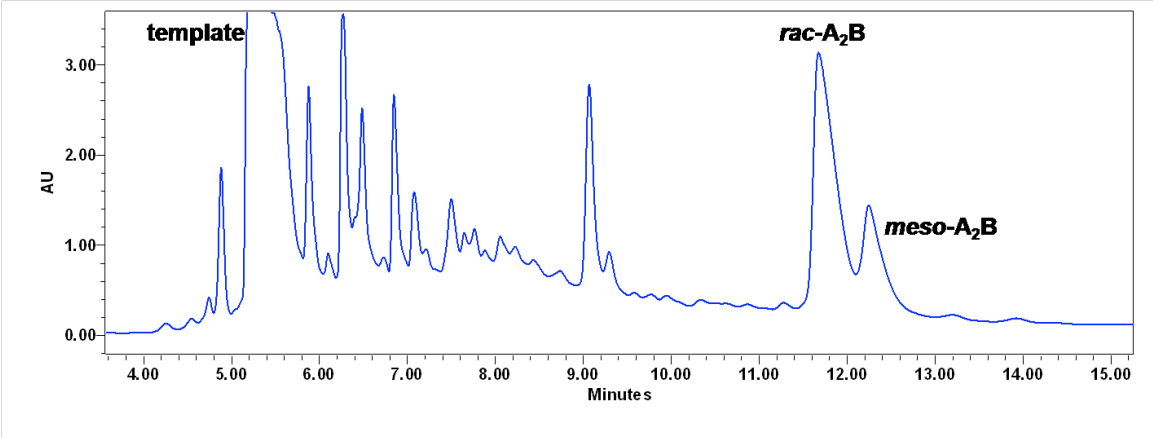


Figure S4. Semi-preparative HPLC trace of an A_2B biased library.

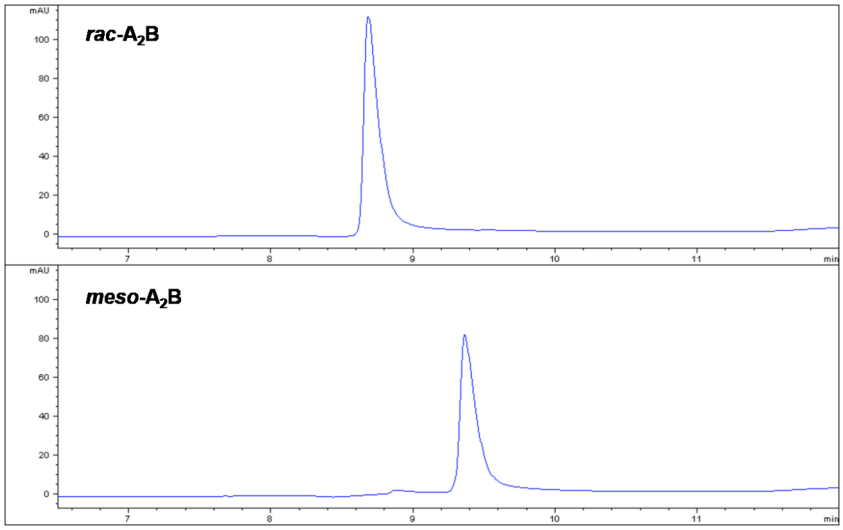


Figure S5. Analytical LC traces of purified *rac*-$\mathbf{A_2B}$ (top) and *meso*-$\mathbf{A_2B}$ (bottom).

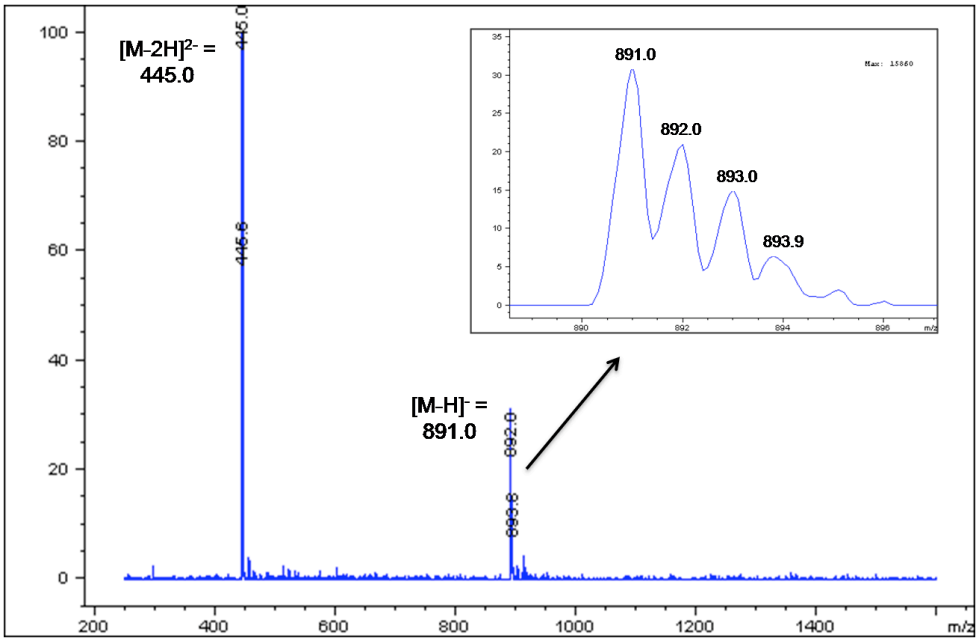


Figure S6. Mass spectra of purified *rac*-$\mathbf{A_2B}$ and *meso*-$\mathbf{A_2B}$.

NMR Measurements. Trimethyl lysine peptide NMR samples were made in 50 mM pD 9.0 sodium borate buffer (referenced to DSS) in the absence and presence of excess *rac*-**A_2B**. Samples were analyzed on a Varian Inova 600 MHz instrument at 25°C. 1D NMR spectra were collected with 64 scans using a 1-1.5 second presaturation or solvent suppression.

Fluorescence Anisotropy. Binding assays were performed with purified *rac*- or *meso*-**A_2B** and fluorescein labeled histone 3 peptides. Assays were prepared in 384-well plates (Corning) with a total volume of 50 μL per well, containing 20 μM labeled peptide and increasing concentrations of **A_2B** in buffer (10 mM phosphate pH 8.5). Plates were spun down and allowed to incubate for at least 30 minutes before analysis. Fluorescence anisotropy was measured on a PHERAstar (BMG Labtech) using FP485, 520A, and 520B filters at 27°C. The anisotropy data was plotted as a function **A_2B** concentration and each plot was fitted in KaleidaGraph to the following equation:[5]

$$r = \left(\left(\frac{(a + x + k_d) \pm \sqrt{(-a - x - k_d)^2 - 4(a \cdot x)}}{2 \cdot a} \right) \cdot (r_\infty - r_o) \right) + r_o$$

where r is fluorescence anisotropy, r_o is the initial anisotropy value, r_∞ is the maximum anisotropy value, a is the peptide concentration, x is the concentration of **A_2B**, and k_d is the dissociation constant. All measurements were taken in duplicate or triplicate.

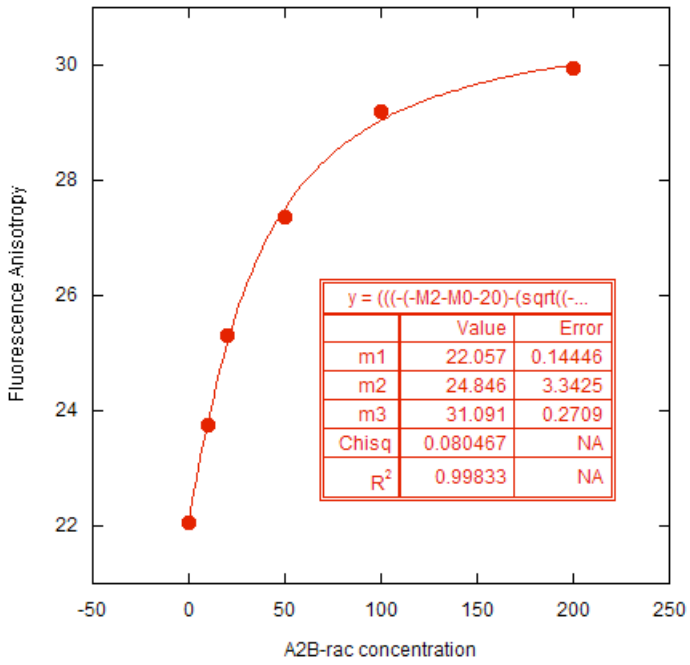


Figure S7. Fluorescence anisotropy of *rac*-$\mathbf{A_2B}$ with H3 KMe_3 (10 mM sodium phosphate buffer pH 8.5).

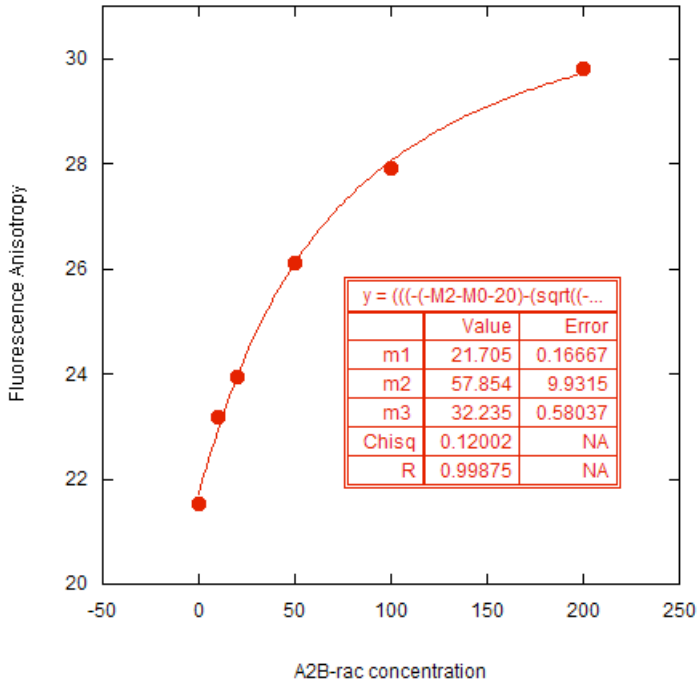


Figure S8. Fluorescence anisotropy of *rac*-$\mathbf{A_2B}$ with H3 KMe_2 (10 mM sodium phosphate buffer pH 8.5).

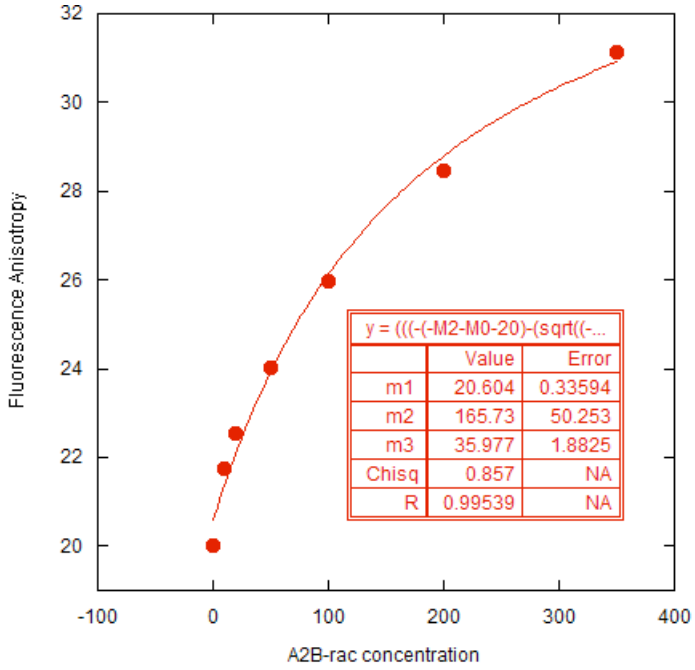


Figure S9. Fluorescence anisotropy of *rac*-$\mathbf{A_2B}$ with H3 KMe (10 mM sodium phosphate buffer pH 8.5).

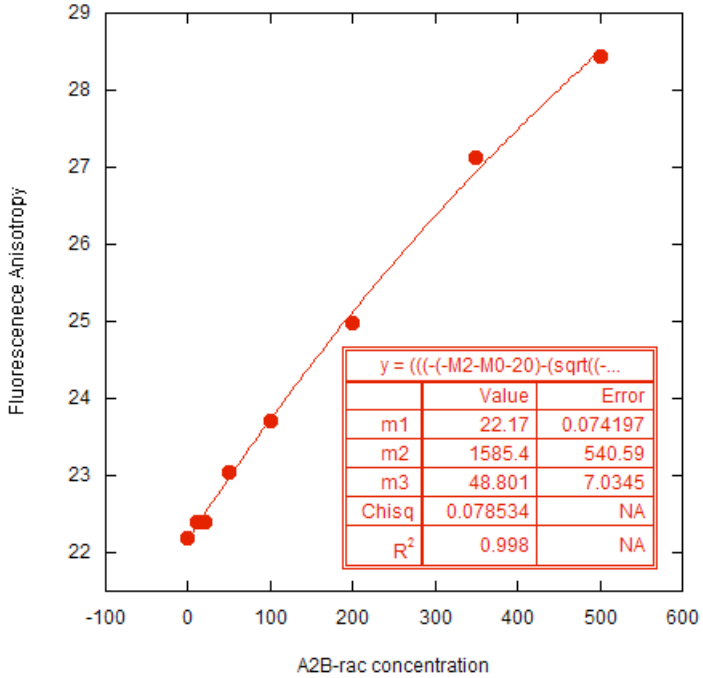


Figure S10. Fluorescence anisotropy of *rac*-$\mathbf{A_2B}$ with H3 K (10 mM sodium phosphate buffer pH 8.5).

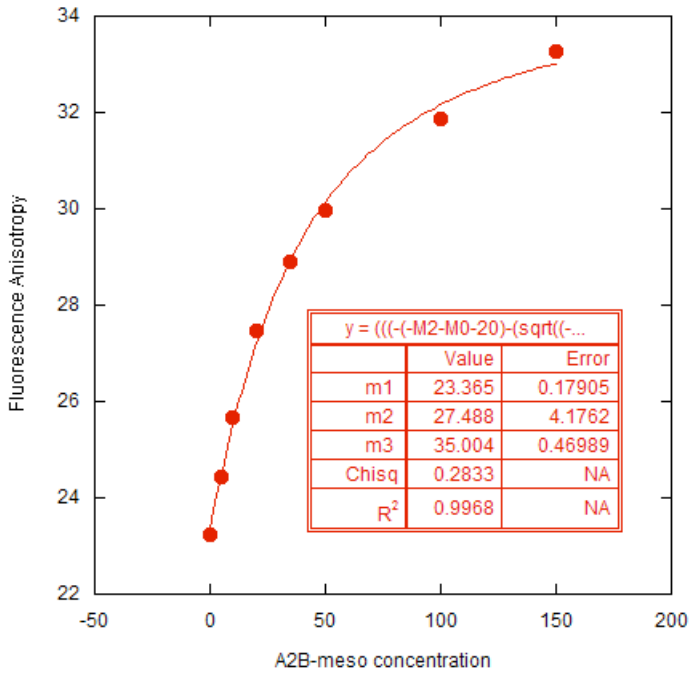


Figure S11. Fluorescence anisotropy of *meso*-$\mathbf{A_2B}$ with H3 KMe_3 (10 mM sodium phosphate buffer pH 8.5).

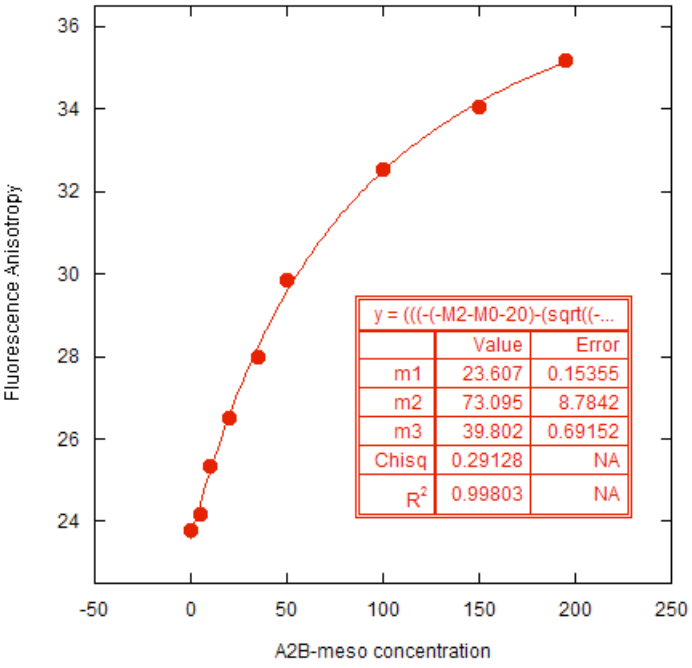


Figure S12. Fluorescence anisotropy of *meso*-$\mathbf{A_2B}$ with H3 KMe_2 (10 mM sodium phosphate buffer pH 8.5).

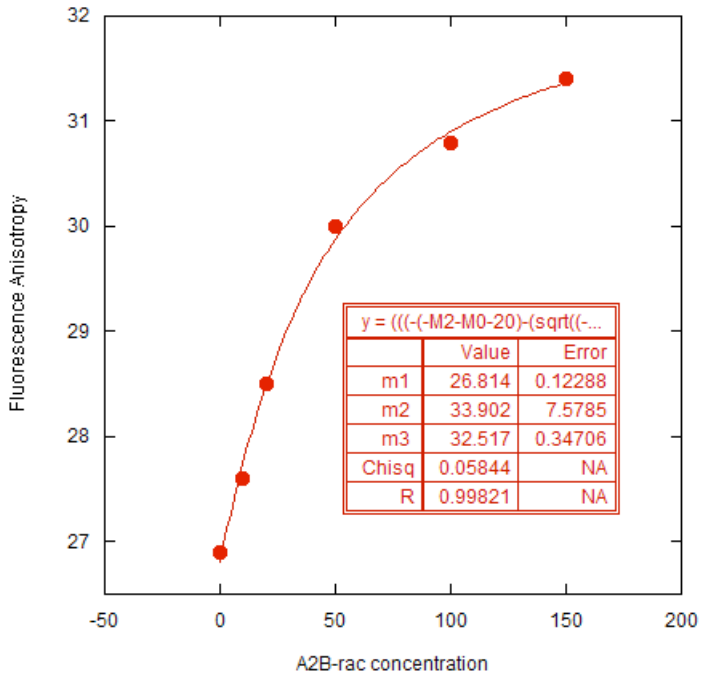


Figure S13. Fluorescence anisotropy of *rac*-**A_2B** with H3 KMe_3 where Arg8 has been mutated to Gly8 (10 mM sodium phosphate buffer pH 8.5).

Isothermal Titration Calorimetry. Isothermal titration calorimetry measurements were conducted using an ITC200 from MicroCal, LLC to verify the binding data obtained by fluorescence anisotropy. A single 0.2 μL aliquot followed by 38 aliquots of 1 μL were titrated into the calorimetric cell every 2.5 minutes. The titrant used to determine the binding to KMe_3 was a 2.0 mM solution of the H3 tail peptide in Figure S12 (in 10 mM phosphate buffer, pH 8.5), containing a Trp for concentration determination and separated from the natural sequence by 3 glycine spacers. This peptide also contained Gly8 as opposed to Arg8 which is found in the native sequence. It has been shown that this mutation has minimal affect on the overall binding affinity. The cell was filled with a 0.2 mM solution of either *rac*- or *meso*-**A_2B** (in 10 mM phosphate buffer, pH 8.5). All measurements were carried out at 26°C. It was found by ITC that *rac*-**A_2B** binds KMe_3

with a K_d of 20.0 μM, whereas *meso*-**A_2B** binds KMe_3 with a K_d of 12.8 μM (Figures S13 and S14).

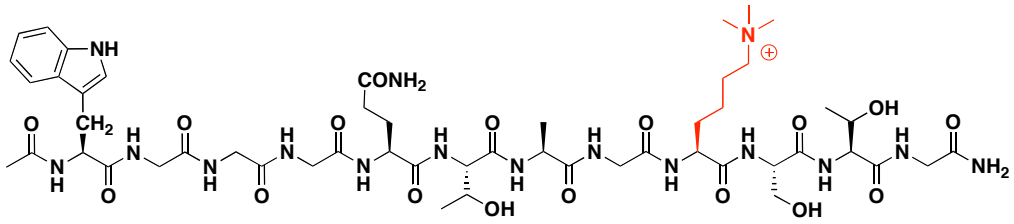


Figure S14. Histone 3 peptide (WGGG-$GKMe_3$) used as the titrant in ITC experiments to confirm **A_2B** binding to KMe_3.

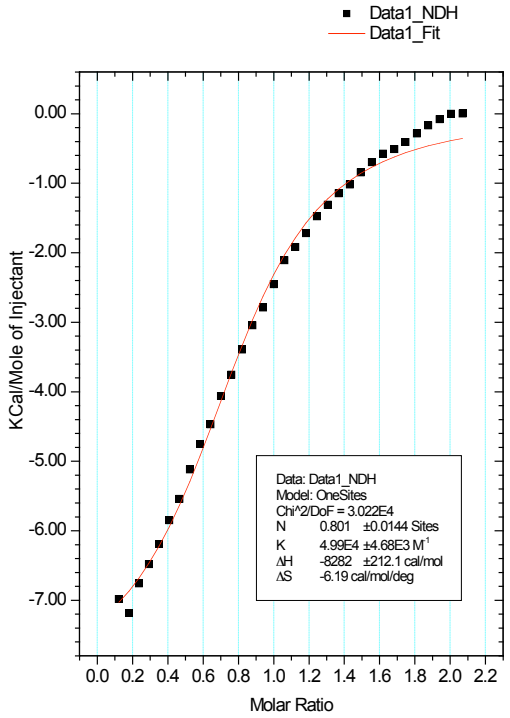


Figure S15. ITC binding curve from the titration of WGGG-$GKMe_3$ into *rac*-**A_2B**, giving a K_d of 20.0 μM.

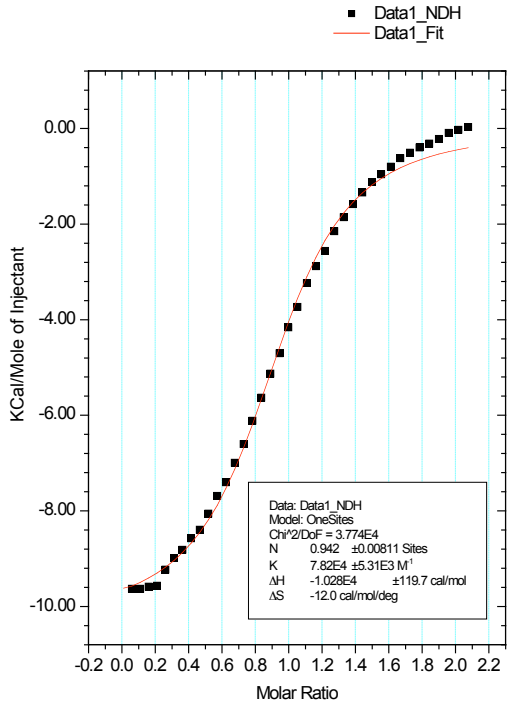


Figure S16. ITC binding curve from the titration of WGGG-GKMe$_3$ into *meso*-**A$_2$B**, giving a K_d of 12.8 μM.

[1] Corbett, P. T.; Sanders, J. K. M.; Otto, S. *Chem. Eur. J.* **2008**, *14*, 2153-2166.
[2] Staab, H. A.; Kirrstetter, R. G. H. *Liebigs Ann. Chem.* **1979**, 886.
[3] West, K. R.; Ludlow, R. F.; Corbett, P. T.; Besenius, P.; Mansfeld, F. M.; Cormack, P. A. G.; Sherrington, D. C.; Goodman, J. M.; Stuart, M. C. A.; Otto, S. *J. Am. Chem. Soc.* **2008**, *130*, 10834-10835.
[4] Suksai, C.; Gomez, S. F.; Chhabra, A.; Liu, J.; Skepper, J. N.; Tuntulani, T., Otto, S. *Langmuir* **2006**, *22*, 5994-5997.
[5] Wang, Y.; Killian, J.; Hamasaki, K.; Rando, R. R. *Biochemistry* **1996**, *35*, 12338-12346.