

# Site-Specific and Regiospecific Installation of Methylarginine Analogues into Recombinant Histones and Insights into Effector Protein Binding

Daniel D. Le,<sup>†</sup> Arianna T. Cortesi,<sup>‡</sup> Samuel A. Myers,<sup>†</sup> Alma L. Burlingame,<sup>§</sup> and Danica Galonić Fujimori<sup>\*;‡;§</sup>

<sup>†</sup>Chemistry and Chemical Biology Graduate Program, <sup>‡</sup>Department of Cellular and Molecular Pharmacology, and <sup>§</sup>Department of Pharmaceutical Chemistry, University of California, San Francisco, 600 16th Street, MC2280, San Francisco, California 94158, United States

## Supporting Information

**ABSTRACT:** Arginine methylation has emerged as a widespread post-translational modification with influence over myriad cellular processes. However, the molecular mechanisms underlying such methylarginine-dependent phenomena remain unclear. To aid in this research, a facile method was developed to install methylarginine analogues on recombinant protein for use in biochemical, biophysical, and structural studies. Through chemical conjugation of novel  $\alpha,\beta$ -unsaturated amidine precursors with proteins, methylarginine mimics can be displayed with control of methylation site, extent, and regiospecificity. Analogue installation into histones using this strategy produced modified proteins that were recognized by antibodies specific to endogenous methylarginine, and these histones retained the capacity to form mononucleosomes. Moreover, a native methylarginine-specific binding domain was shown to interact with methylarginine analogue-modified substrates. This chemical conjugation method for installing methylarginine analogues provides an efficient route to produce homogeneous modified proteins for subsequent investigations of methylarginine-dependent processes.

A large and expanding group of proteins, with diverse cellular functions, are known to harbor post-translational arginine methylation.<sup>1</sup> Because the guanidino group of arginine can participate in polydentate ionic and hydrogen bonding interactions, arginine serves as a critical component of certain protein–protein<sup>2,3</sup> and protein–nucleic acid<sup>4</sup> interfaces. Arginine-mediated interactions are modulated by methylation, which exists in three major physiological forms:  $\omega$ -*N*-methylarginine (monomethylarginine, MMA),  $\omega$ -*N,N*-dimethylarginine (asymmetric dimethylarginine, ADMA), and  $\omega$ -*N,N'*-dimethylarginine (symmetric dimethylarginine, SDMA).<sup>1</sup> *In vivo*, arginine methylation is installed by protein arginine methyltransferase (PRMT) enzymes. The PRMT paralogues are classified as either Type I or Type II based on product selectivity: Type I PRMTs produce ADMA, Type II PRMTs produce SDMA, and both classes produce MMA. For example, the Type I enzyme PRMT1 methylates histone H4 at residue R3 to yield both H4R3Me1 (monomethyl) and H4R3Me2a

(dimethyl, asymmetric). Likewise, the Type II paralogue PRMT5 deposits H4R3Me1 and H4R3Me2s (dimethyl, symmetric). Ultimately, these marks interact with cellular machinery to elicit a physiological output.<sup>5,6</sup>

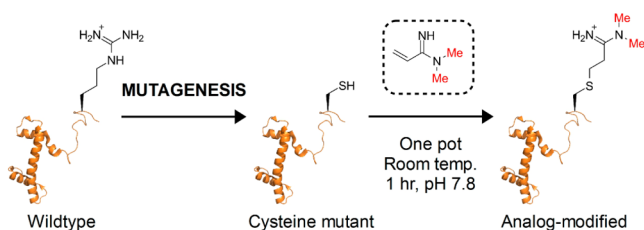
While arginine methylation is implicated in many cellular processes—such as transcriptional regulation,<sup>7</sup> DNA repair<sup>8</sup> and RNA trafficking<sup>9</sup>—there remains much to be elucidated about the underlying biochemical mechanisms. For example, methylarginine-specific binding modules are present among diverse classes of proteins;<sup>2</sup> yet, in many cases, the biological significance of methylarginine-dependent interactions is unclear. At the protein level, methylarginine-specific association can affect localization,<sup>9</sup> complex participation,<sup>10</sup> and may also influence catalytic function.<sup>6</sup> Therefore, methylated substrate proteins are necessary reagents for the study of methylarginine-dependent activity. However, current methods to access homogeneous recombinant protein with prescribed site, extent and regiospecificity of arginine methylation suffer from technical challenges that include demanding biochemical manipulations<sup>11</sup> and limited control of methylation status.<sup>12</sup> Presented here is a facile method to install methylarginine analogues on recombinant protein with defined methylation status.

Akin to the chemical conjugation strategies employed for methyl-,<sup>13</sup> acetyl-,<sup>14,15</sup> and ubiquityl-lysine analogues;<sup>16</sup> the site of methylarginine analogue installation is directed by chemo-selective reaction with a mutant cysteine residue. As such, methylarginine analogue conjugation is ideal for modifying proteins that lack cysteine residues, other than at the site of analogue installation. Methylarginine analogue precursors are based on a novel  $\alpha,\beta$ -unsaturated amidine scaffold with various degrees of *N*-methylation, which reacts with cysteinyl thiols via conjugate addition (Scheme 1). Once appended to protein, the amidine moiety serves as a guanidino group side chain mimic with shape and  $pK_a$  (acetamidine  $pK_{a(DMSO)} = 27.1$ , guanidine  $pK_{a(DMSO)} = 28.5$ )<sup>17</sup> similar to native arginine—features shared by amidine-based inhibitors for arginine-specific proteases,<sup>18</sup> nitric oxide synthases,<sup>19</sup> dimethylarginine dimethylamino-hydrolases,<sup>19</sup> arginine deiminases,<sup>20</sup> and PRMTs.<sup>21</sup> Unlike the guanidino side chain of arginine, amidine analogues possess an

Received: November 2, 2012

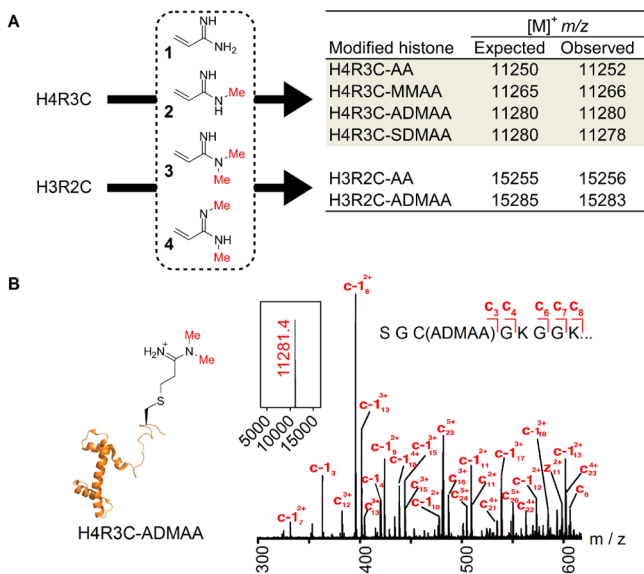
Published: February 11, 2013

### Scheme 1. Chemical Conjugation of Methylarginine Analogues with Recombinant Histone



$\epsilon$ -methylene group in place of nitrogen that introduces nonplanar geometry and precludes polar interactions at this atom. However, structural studies suggest that at least a subset of methylarginine-dependent interactions is primarily mediated via the terminal nitrogen atoms (Figure S1), which are conserved between the guanidino and amidino side chains.

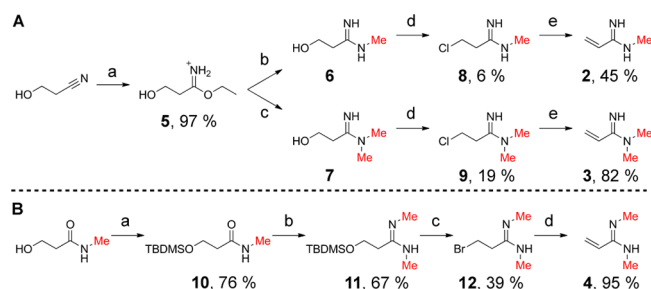
Analogue precursor compounds 1–4 (Figure 1A) were synthesized with predefined methylation extent and regio-



**Figure 1.** Site-specific and regiospecific installation of arginine analogues into histones. (A) Alkylation of histone mutants with analogue precursors 1–4. Observed  $m/z$  values for analogue-modified histones are in agreement with expected values. (B) Deconvoluted intact mass (inset) and ETD mass spectrum of H4R3C-ADMAA yielding product ions consistent with site-specific modification.

specificity at the amidine group to mimic various forms of endogenous methylarginine. Precursor 1 was prepared by addition of methylchloroaluminum amide to acrylonitrile followed by hydrolysis.<sup>22</sup> The MMA and ADMA analogue precursors (compounds 2 and 3, respectively) were produced by aminolysis of ethyl 3-hydroxypropionimidate hydrochloride (5), prepared by the Pinner reaction (Scheme 2A). Subsequently, chlorination of the hydroxyl group allowed for base-promoted  $\beta$ -elimination to yield the desired  $\alpha,\beta$ -unsaturated amidines. Access to the symmetric amidine precursor 4 was achieved via methylamine addition to an activated *N*-methylamide obtained by reaction of 3-hydroxy-*N*-methylpropanamide *tert*-butyldimethylsilyl ether (10) with triethyloxonium tetrafluoroborate (Scheme 2B). Deprotection, halogenation, and elimination yielded the target unsaturated symmetric amidine 4.

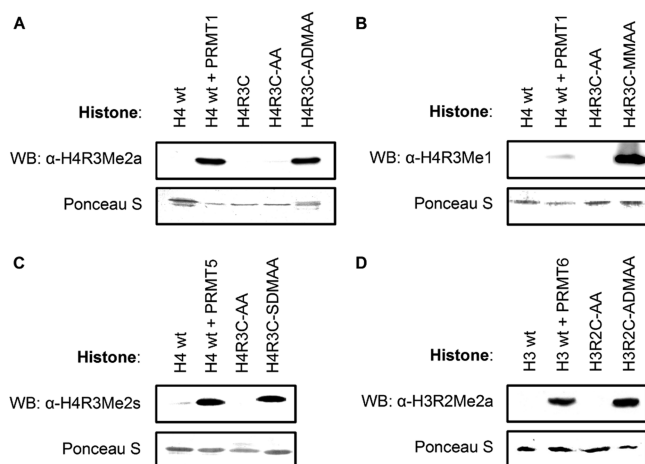
### Scheme 2. Preparation of Analogue Precursors<sup>a</sup>



<sup>a</sup>(A) Synthesis of monomethyl and asymmetric dimethyl analogue precursors: (a) HCl, EtOH, Et<sub>2</sub>O, 3 h; (b) MeNH<sub>2</sub>, MeOH, 19 h; (c) Me<sub>2</sub>NH, MeOH, 16 h; (d) PPh<sub>3</sub>, CCl<sub>4</sub>, DMF, 22–24 h; (e) triethylamine, acetonitrile, 1–5 min. (B) Synthesis of symmetric dimethyl analogue precursor (a) TBDMSO, imidazole, DMF, 24 h; (b) i. Et<sub>3</sub>OBF<sub>4</sub>, DCM, 1 h; ii. MeNH<sub>2</sub>, DCM/THF, 1 h; (c) PPh<sub>3</sub>Br<sub>2</sub>, DCM, 20 h; (d) triethylamine, acetonitrile, 1 min. All reactions were performed at room temperature.

The analogue precursors were reacted with recombinant *Xenopus* histones H3 and H4 containing cysteine substitutions: H3R2C (possesses an additional C110A mutation that is functionally akin to the wild-type protein<sup>13</sup>) and H4R3C, respectively. The reactions were performed at room temperature under denaturing and reducing conditions for ~1 h. The following analogue-modified histone H4 derivatives were produced by reaction with precursors 1–4, respectively (Figure 1A): H4R3C-AA (arginine analogue), H4R3C-MMAA (monomethyl arginine analogue), H4R3C-ADMAA (asymmetric dimethyl arginine analogue), and H4R3C-SDMAA (symmetric dimethyl arginine analogue). Likewise, histone H3R2C was reacted with 1 and 3 to yield H3R2C-AA and H3R2C-ADMAA, respectively. Formation of the analogue-modified histones was monitored by ESI-MS, and observed  $m/z$  values were in accordance with those expected (Figure 1A). Mass spectrometric analysis of alkylation reactions showed complete loss of starting protein masses concomitant with the emergence of corresponding alkylated product ions, which indicates highly efficient analogue conjugation (Figure S2). Analysis of intact H4R3C-ADMAA by tandem mass spectrometry (MS/MS) confirmed site-specific installation of the methylarginine analogue. The predominant precursor ion at  $m/z$  11281.4 corresponding to H4R3C-ADMAA (Figure 1B, inset) was subjected to electron-transfer dissociation (ETD), and the resultant product ion series was consistent with analogue installation at Cys 3 (Figure 1B).

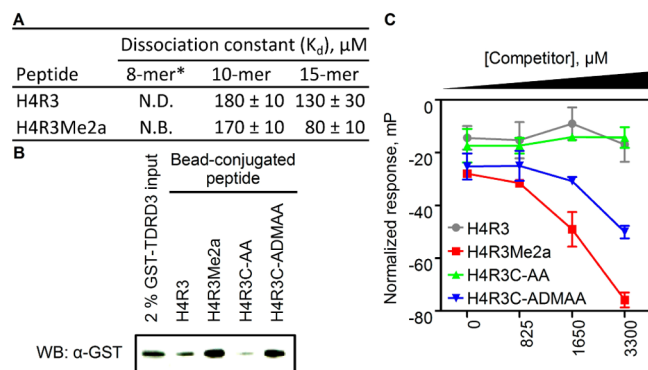
In studies of methylarginine-dependent gene expression, antibodies specific for histones harboring arginine methylation have been used to map the distribution of methylarginine along chromatin.<sup>23</sup> These antibodies were employed to assess recognition of analogue-modified histones by Western blot analysis. As expected,  $\alpha$ -H4R3Me2a antibody recognized wild-type histone H4 treated with PRMT1—which enzymatically installs native ADMA—whereas the untreated histone exhibited no signal (Figure 2A). The same antibody also recognized H4R3C-ADMAA but failed to detect both H4R3C and H4R3C-AA. Similar results were observed using  $\alpha$ -H4R3Me1 with H4R3C-MMAA (Figure 2B),  $\alpha$ -H4R3Me2s with H4R3C-SDMAA (Figure 2C), and  $\alpha$ -H3R2Me2a with H3R2C-ADMAA (Figure 2D). In addition, the analogue-modified histones displayed minimal antigen cross-reactivity among the respective antibodies (Figure S3). Together, these antibody



**Figure 2.** Analogue-modified histones were recognized by methylarginine-specific antibodies in Western blot analyses. Antibody specificities: (A) H4R3Me2a, (B) H4R3Me1, (C) H4R3Me2s, and (D) H3R2Me2a. Lower row in each panel shows protein loading control.

recognition results indicate that analogue-modified histones are reasonable mimics of native arginine and methylarginine. Next, the analogue-modified histones were used to reconstitute nucleosomes, protein:DNA complexes that form chromatin. Mononucleosomes were assembled from octamers, composed of histones H3, H2A, H2B, and either wild-type H4 or analogue-modified H4, wrapped by the 601 DNA sequence (Figure S4). Native polyacrylamide gel electrophoresis (PAGE) analysis of the WT, AA-modified, and ADMAA-modified nucleosome reconstitution experiments shows nearly equal extent of retarded DNA migration upon assembly, indicating that the process of analogue installation does not affect nucleosome reconstitution efficiency. Thus, incorporation of methylarginine analogues into nucleosome complexes provides access to homogeneous modified chromatin substrates. The ease of analogue installation makes this strategy amenable to large-scale preparation of nucleosomes.

The ability of methylarginine analogues to mimic native arginine methylation was further tested by investigating analogue association with the putative methylarginine-specific effector protein TDRD3. The Tudor domain-containing protein TDRD3 localizes to transcription start sites of many highly expressed genes and binds to H4R3Me2a; therefore, TDRD3 was postulated to be a methylarginine-specific effector protein with transcriptional co-activator function.<sup>24</sup> In contrast to these findings, an isothermal titration calorimetry (ITC) binding study indicated no detectable association between the Tudor domain of TDRD3 and an 8-mer H4R3Me2a peptide.<sup>25</sup> To gain further insight into TDRD3 binding specificity, equilibrium dissociation constants ( $K_d$ ) were determined for the Tudor domain with various peptide substrates using a fluorescence polarization anisotropy saturation binding assay. Under conditions similar to those of the aforementioned ITC study, the Tudor domain of TDRD3 demonstrated affinity for 10-mer and 15-mer peptide substrates with  $K_d$  values in the low- to mid-micromolar range (Figures 3A and S5A). The observed increase in affinity as a function of peptide length may explain the reported lack of binding with a minimal 8-mer substrate.<sup>25,26</sup> Furthermore, the 15-mer peptide displayed the most pronounced methylation-dependent binding, demonstrating a  $\sim 1.6$ -fold decrease in  $K_d$  attributable to methylation. Site-



**Figure 3.** The Tudor domain of TDRD3 binds native asymmetric dimethylarginine and the corresponding analogue-modified substrate. (A) Measured equilibrium dissociation constants from fluorescence polarization saturation binding assays. (B) Western blot analysis of GST-TDRD3 precipitated on peptide-coated beads. (C) Fluorescence polarization competition assay. N.B. = no binding. N.D. = not determined. \* = from ref 25.

directed mutagenesis (YS66A) of the aromatic cage binding pocket within TDRD3 markedly decreased peptide binding affinity; further indicating that substrate interactions are specific to the Tudor domain (Figure S6).<sup>2,25</sup> Taken together, these results show that the Tudor domain of TDRD3 is a binding module that discriminates, albeit weakly, between non-methylated and asymmetrically dimethylated arginine in histone H4.

Having confirmed TDRD3 Tudor domain specificity for native methylarginine, binding of analogue-modified peptides was tested. Histone peptides (15-mers) possessing a biotin tag were conjugated to streptavidin-coated beads, to which a solution of TDRD3 Tudor domain fusion protein (GST-TDRD3) was added. The extent of Tudor domain binding was assessed by Western blot analysis of precipitated protein. For the native H4R3Me2a peptide, adhesion of GST-TDRD3 was evident; in contrast, the non-methylated peptide (H4R3) gave a comparatively diminished signal (Figure 3B). These results were mirrored by the corresponding analogue-modified peptides: H4R3C-ADMAA and H4R3C-AA. In addition, free H4R3Me2a antagonized the association of GST-TDRD3 with bead-conjugated H4R3C-ADMAA (Figure S5B), indicating specific interaction between the Tudor domain and the analogue-modified substrate.

A fluorescence polarization competition assay was used to assess TDRD3 Tudor domain binding specificity under homogeneous solution-phase conditions. The association between His<sub>6</sub>-MBP-TDRD3 (a high-yielding soluble fusion construct of the Tudor domain) and a tetramethylrhodamine-tagged 15-mer H4R3Me2a peptide produces polarized fluorescence, which was antagonized by competitor peptides (10-mer, untagged). The displacement of fluorescent peptide due to competition causes a decrease in polarized emission inversely proportional to competitor ligand affinity. As expected, the H4R3Me2a ligand was the most potent competitor and the corresponding H4R3C-ADMAA variant was  $\sim 1.5$ -fold less effective at competitor concentrations between 1.7 and 3.3 mM (Figure 3C). Non-methylated H4R3 and H4R3C-AA competitor peptides demonstrated no affinity for TDRD3. Taken together, methylarginine analogues functionally mimic the native mark and TDRD3 exhibits only a mild preference for methylated over non-methylated arginine.



Despite the subtle substrate specificity of TDRD3, the observed methylation-dependent binding affinity differences may represent a physiological mechanism to fine-tune its proposed transcriptional co-activation function within a cellular context.

The site-specific incorporation of methylarginine analogues into recombinant protein offers a facile method to attain homogeneous substrates for biochemical, biophysical, and structural investigations into the functions of arginine methylation. As demonstrated, analogue precursors 1–4 quantitatively modify histone proteins containing a single cysteine mutation that directs site-specific conjugation. This method provides control of methylation extent (non-, mono-, and dimethylarginine) in addition to regiospecificity (symmetric and asymmetric dimethylarginine). Chemically modified histones displaying methylarginine analogues were recognized by antibodies specific for native methylation states. Additionally, such histones were capable of forming mononucleosomes, demonstrating that methylarginine analogues can provide access to modified chromatin substrates. Experiments with both endogenous asymmetric dimethylarginine and the corresponding analogue variant indicated that the methylarginine recognition module of TDRD3 preferentially binds the asymmetric dimethylated mark in histone H4. Taken together, this work demonstrates that methylarginine analogue installation into recombinant protein by chemical conjugation is an effective method to produce substrates for *in vitro* investigations of methylarginine-dependent processes.

## ■ ASSOCIATED CONTENT

### Supporting Information

Precursor synthesis and characterization, full experimental methods, and supplemental figures. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## ■ AUTHOR INFORMATION

### Corresponding Author

fujimori@cmp.ucsf.edu

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

We thank Lauren Buckley (formerly of UW Madison) for assistance with alkylation condition optimization, the laboratories of James Wells (UCSF) and Kevan Shokat (UCSF) for instrument usage, in addition to Steven Clarke (UCLA) for expression plasmids of PRMT paralogues. This work was supported by the UCSF Program for Breakthrough Biomedical Research, the Searle Scholars Program (D.G.F.), the V Scholar Award (D.G.F.), and the NSF (graduate research fellowship to D.D.L.). Mass spectrometry was provided by the Bio-Organic Biomedical Mass Spectrometry Resource at UCSF (A.L.B., Director), supported by NIH NIGMS 8P41GM103481 and the Howard Hughes Medical Institute.

## ■ REFERENCES

- (1) Bedford, M. T.; Clarke, S. G. *Mol. Cell* **2009**, *33*, 1.
- (2) Chen, C.; Nott, T. J.; Jin, J.; Pawson, T. *Nat. Rev. Mol. Cell Biol.* **2011**, *12*, 629.
- (3) Trievel, R. C.; Shilatfard, A. *Nat. Struct. Mol. Biol.* **2009**, *16*, 678.
- (4) Luscombe, N. M.; Laskowski, R. A.; Thornton, J. M. *Nucleic Acids Res.* **2001**, *29*, 2860.

- (5) Zhao, Q.; Rank, G.; Tan, Y. T.; Li, H.; Moritz, R. L.; Simpson, R. J.; Cerruti, L.; Curtis, D. J.; Patel, D. J.; Allis, C. D.; Cunningham, J. M.; Jane, S. M. *Nat. Struct. Mol. Biol.* **2009**, *16*, 304.
- (6) An, W.; Kim, J.; Roeder, R. G. *Cell* **2004**, *117*, 735.
- (7) Migliori, V.; Muller, J.; Phalke, S.; Low, D.; Bezzi, M.; Mok, W. C.; Sahu, S. K.; Gunaratne, J.; Capasso, P.; Bassi, C.; Cecatiello, V.; De Marco, A.; Blackstock, W.; Kuznetsov, V.; Amati, B.; Mapelli, M.; Guccione, E. *Nat. Struct. Mol. Biol.* **2012**, *19*, 136.
- (8) Guo, Z.; Zheng, L.; Xu, H.; Dai, H.; Zhou, M.; Pascua, M. R.; Chen, Q. M.; Shen, B. *Nat. Chem. Biol.* **2010**, *6*, 766.
- (9) McBride, A. E.; Cook, J. T.; Stemmler, E. A.; Rutledge, K. L.; McGrath, K. A.; Rubens, J. A. *J. Biol. Chem.* **2005**, *280*, 30888.
- (10) Gao, X.; Zhao, X.; Zhu, Y.; He, J.; Shao, J.; Su, C.; Zhang, Y.; Zhang, W.; Saarikettu, J.; Silvennoinen, O.; Yao, Z.; Yang, J. *J. Biol. Chem.* **2012**, *287*, 18130.
- (11) Akahoshi, A.; Suzue, Y.; Kitamatsu, M.; Sisido, M.; Ohtsuki, T. *Biochem. Biophys. Res. Commun.* **2011**, *414*, 625.
- (12) Kölbel, K.; Ihling, C.; Bellmann-Sickert, K.; Neundorff, I.; Beck-Sickingler, A. G.; Sinz, A.; Kühn, U.; Wahle, E. *J. Biol. Chem.* **2009**, *284*, 8274.
- (13) Simon, M. D.; Chu, F.; Racki, L. R.; de la Cruz, C. C.; Burlingame, A. L.; Panning, B.; Narlikar, G. J.; Shokat, K. M. *Cell* **2007**, *128*, 1003.
- (14) Huang, R.; Holbert, M. A.; Tarrant, M. K.; Curtet, S.; Colquhoun, D. R.; Dancy, B. M.; Dancy, B. C.; Hwang, Y.; Tang, Y.; Meeth, K.; Marmorstein, R.; Cole, R. N.; Khochin, S.; Cole, P. A. *J. Am. Chem. Soc.* **2010**, *132*, 9986.
- (15) Li, F.; Allahverdi, A.; Yang, R.; Lua, G. B.; Zhang, X.; Cao, Y.; Korolev, N.; Nordenskiöld, L.; Liu, C. F. *Angew. Chem., Int. Ed.* **2011**, *50*, 9611.
- (16) Valkevich, E. M.; Guenette, R. G.; Sanchez, N. A.; Chen, Y. C.; Ge, Y.; Strieter, E. R. *J. Am. Chem. Soc.* **2012**, *134*, 6916.
- (17) Bordwell, F. G.; Ji, G. Z. *J. Am. Chem. Soc.* **1991**, *113*, 8398.
- (18) Nar, H. *Trends Pharmacol. Sci.* **2012**, *33*, 279.
- (19) Wang, Y.; Monzingo, A. F.; Hu, S.; Schaller, T. H.; Robertus, J. D.; Fast, W. *Biochemistry* **2009**, *48*, 8624.
- (20) Luo, Y.; Arita, K.; Bhatia, M.; Knuckley, B.; Lee, Y. H.; Stallcup, M. R.; Sato, M.; Thompson, P. R. *Biochemistry* **2006**, *45*, 11727.
- (21) Obianyo, O.; Causey, C. P.; Jones, J. E.; Thompson, P. R. *ACS Chem. Biol.* **2011**, *6*, 1127.
- (22) Garigipati, R. S. *Tetrahedron Lett.* **1990**, *31*, 1969.
- (23) Barski, A.; Cuddapah, S.; Cui, K.; Roh, T. Y.; Schones, D. E.; Wang, Z.; Wei, G.; Chepelev, I.; Zhao, K. *Cell* **2007**, *129*, 823.
- (24) Yang, Y.; Lu, Y.; Espejo, A.; Wu, J.; Xu, W.; Liang, S.; Bedford, M. T. *Mol. Cell* **2010**, *40*, 1016.
- (25) Liu, K.; Guo, Y.; Liu, H.; Bian, C.; Lam, R.; Liu, Y.; Mackenzie, F.; Rojas, L. A.; Reinberg, D.; Bedford, M. T.; Xu, R. M.; Min, J. *PLoS One* **2012**, *7*, e30375.
- (26) Voigt, P.; Reinberg, D. *Chembiochem* **2011**, *12*, 236.