

Article

# Comparison of Molecular Recognition of Trimethyllysine and Trimethylthialysine by Epigenetic Reader Proteins

Jordi C. J. Hintzen <sup>1,†</sup>, Jordi Poater <sup>2,†</sup>, Kiran Kumar <sup>3,†</sup>, Abbas H. K. Al Temimi <sup>4,†</sup>, Bas J. G. E. Pieters <sup>4</sup>, Robert S. Paton <sup>3,\*</sup>, F. Matthias Bickelhaupt <sup>4,5,\*</sup> and Jasmin Mecinović <sup>1,4,\*</sup>

<sup>1</sup> Department of Physics, Chemistry and Pharmacy, University of Southern Denmark, Campusvej 55, 5230 Odense, Denmark

<sup>2</sup> ICREA and Departament de Química Inorgànica i Orgànica & IQTCUB, Universitat de Barcelona, Martí I Franquès 1–11, 08028 Barcelona, Spain

<sup>3</sup> Chemistry Research Laboratory, University of Oxford, 12 Mansfield Road, Oxford OX1 3TA, UK

<sup>4</sup> Institute for Molecules and Materials, Radboud University, Heyendaalseweg 135, 6522 AJ Nijmegen, The Netherlands

<sup>5</sup> Department of Theoretical Chemistry and Amsterdam Center for Multiscale Modeling, Vrije Universiteit Amsterdam, De Boelelaan 1083, 1081HV Amsterdam, The Netherlands

\* Correspondence: robert.paton@colostate.edu (R.S.P.); f.m.bickelhaupt@vu.nl (F.M.B.); mecinovic@sdu.dk (J.M.); Tel.: +45-6550-3603 (J.M.)

† These authors contributed equally to this work.

Academic Editors: Ivo Piantanida and Adegboyega K. Oyelere

Received: 26 February 2020; Accepted: 16 April 2020; Published: 21 April 2020



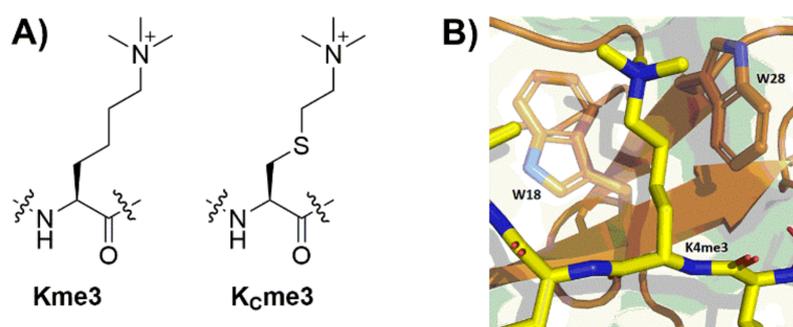
**Abstract:** Gaining a fundamental insight into the biomolecular recognition of posttranslationally modified histones by epigenetic reader proteins is of crucial importance to understanding the regulation of the activity of human genes. Here, we seek to establish whether trimethylthialysine, a simple trimethyllysine analogue generated through cysteine alkylation, is a good trimethyllysine mimic for studies on molecular recognition by reader proteins. Histone peptides bearing trimethylthialysine and trimethyllysine were examined for binding with five human reader proteins employing a combination of thermodynamic analyses, molecular dynamics simulations and quantum chemical analyses. Collectively, our experimental and computational findings reveal that trimethylthialysine and trimethyllysine exhibit very similar binding characteristics for the association with human reader proteins, thereby justifying the use of trimethylthialysine for studies aimed at dissecting the origin of biomolecular recognition in epigenetic processes that play important roles in human health and disease.

**Keywords:** epigenetics; histone; lysine methylation; molecular recognition; noncovalent interactions

## 1. Introduction

Biomolecular recognition of posttranslationally modified histone proteins is centrally important to regulation of the activity of human genes [1]. One of the most important and widespread histone modifications is lysine methylation, which is found on core histones and histone tails [2,3]. Histone lysine methyltransferases (KMTs) catalyze the transfer of methyl group from S-adenosylmethionine (SAM) to lysine  $\epsilon$ -amino group, leading to three different methylation states (i.e., monomethyllysine Kme, dimethyllysine Kme2 and trimethyllysine Kme3, Figure 1A), which can be removed by histone lysine demethylases (KDMs) [4,5]. Methylated lysine residues play different roles in epigenetic processes, as these marks are specifically recognized by structurally diverse classes of epigenetic reader proteins.

The highest mark, trimethyllysine, is recognized by the aromatic cage-containing readers, including tandem tudor domains (TTD), chromodomains (CD) and the plant homeodomain (PHD) zinc finger proteins (Figure 1B) [6]. To gain a better understanding of the exact role of lysine methylation in epigenetics, it is important to develop novel chemical tools for studying the molecular mechanisms that govern the molecular recognition of methylated lysines by reader proteins. An installation of chemically modified methylated lysine analogues into histone proteins [7] and histone peptides [8–11] has been a valuable method to study how changes in structure affect the association with reader proteins. In addition, a variety of methods have been developed to incorporate unnatural amino acids in both histone proteins and peptides, notably by auxotrophic expression systems [12] or employing the amber stop codon (TAG) [13]. Major drawbacks of both methods include severe limitations of amino acid variants that can be incorporated into histones. Furthermore, when using auxotrophic strains, the protein expression yield is decreased dramatically, and the process of developing an amber codon pair is time consuming and laborious [14,15].



**Figure 1.** (A) Structures of of trimethyllysine (**Kme3**) and trimethylthialysine (**K<sub>C</sub>me3**); (B) view of the KDM5A<sub>PHD3</sub> (orange) structure complexed with histone H3K4me3 (yellow) (PDB ID: 2KGI).

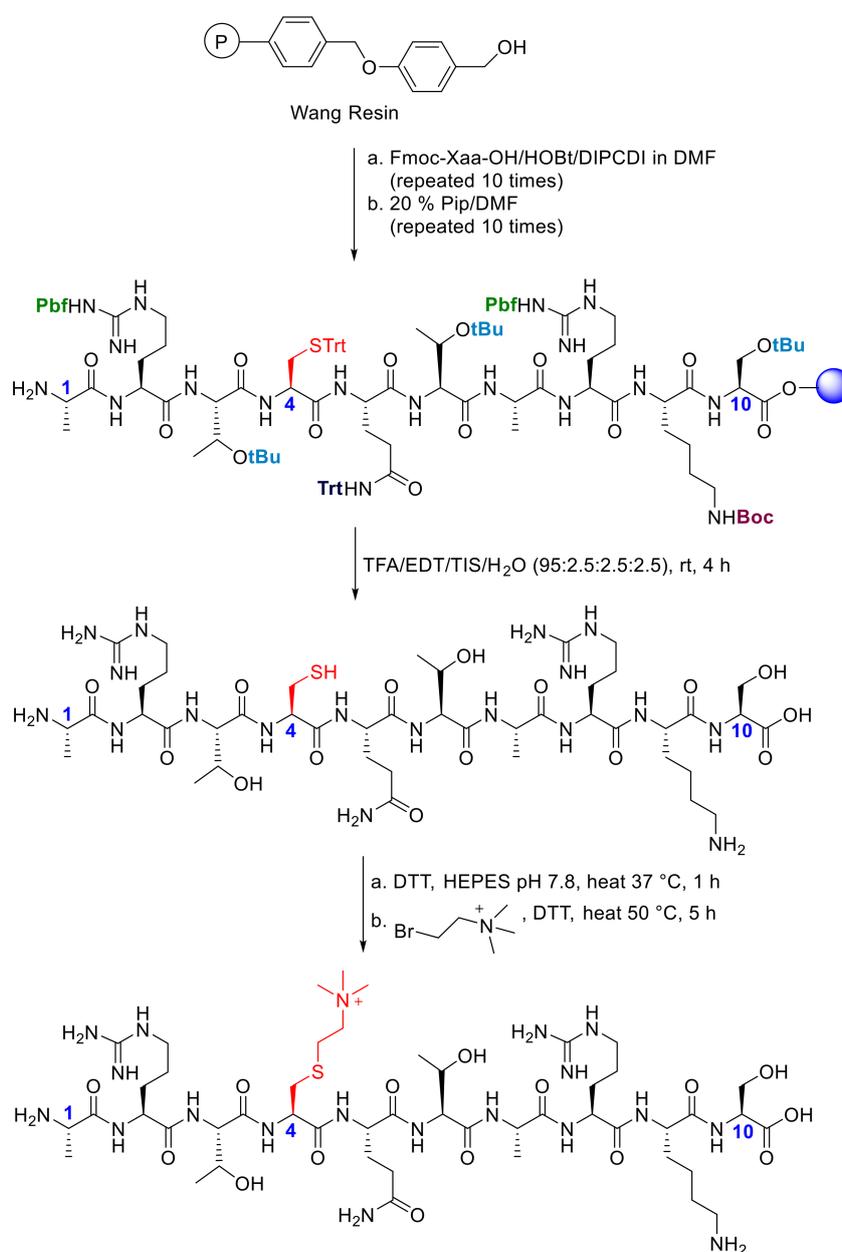
Meanwhile, several synthetic and semi-synthetic methods have been developed to allow a site-specific incorporation of natural and unnatural amino acids in histones. Native chemical ligation can be used to develop fully synthetic proteins bearing the desired posttranslationally modified amino acids site-specifically [16]. This strategy has been shown to be viable in obtaining fully synthetic histone 3 [17]. Full protein synthesis can still be challenging and laborious, however, therefore, a desired alternative is found in synthetically simpler methods. The unique properties of the thiol group of cysteine have been used to this end. Cysteine can be selectively alkylated to obtain analogues that mimic naturally occurring amino acids. Among these are arginine [18], lysine [19,20] and different posttranslationally modified variants of lysine, including acetylated [21] and methylated analogues [22]. For the purpose of generating methylated lysine, simple bromides can be used for chemoselective reaction with the cysteine thiol to obtain intact histone proteins that possess simplest methylated lysine analogues (MLAs) [23].

Cysteine alkylation is an especially valuable method for examinations of histones, as only one or two native cysteine residues exist in four histone proteins; mutating C110 in H3 into an alanine (C110A) does not lead to any loss of function [22]. It has been shown that trimethylthialysine (K<sub>C</sub>me3, Figure 1A), the alkylated cysteine analogue of trimethyllysine, is well recognized by different epigenetic reader proteins [24], but as of yet, it is still up for debate whether the thioether bond created in the backbone of the trimethylthialysine is actually a good way to mimic the natural C–C bond in trimethyllysine, and somewhat conflicting reports in literature have surfaced [24,25]. When comparing the C–C bond to the C–S bond present in these analogues, it is evident that there are some differences in the bond angle (C–S–C is about 12 degrees smaller than C–C–C), while the C–S bond length is about 20% longer than the length of the C–C bond [25]. These two factors counteract each other to some extent, but still on average there is ~0.3 Å expansion of the distance between the geminal methylene units in the alkane and the thioether variant. Whether these subtle alterations contribute to differences in molecular recognition of

methylated lysines by reader proteins seems to be heavily sequence and binder dependent, as there have been reports on practically no loss of binding affinity to 13-fold decreases [24,25]. The objective of this work is to systematically compare the origin of molecular recognition of trimethyllysine and trimethylthialysine-possessing histones by reader proteins, using a combination of thermodynamic analyses, molecular dynamics simulations and quantum chemical analyses.

## 2. Results

We synthesized 10-mer histone H3 peptides that possessed natural trimethyllysine (H3K4me3) and unnatural trimethylthialysine (H3K<sub>C</sub>4me3) by solid-phase peptide synthesis and purified them by preparative HPLC (Figure 2 and Scheme S1). A cysteine residue was introduced at position 4 of the histone peptide, which was site-specifically reacted with (2-bromoethyl)trimethylammonium bromide to form trimethylthialysine. Both synthetic histone peptides were purified by preparative HPLC.



**Figure 2.** Solid-phase peptide synthesis of histone peptide H3K<sub>C</sub>4me3.

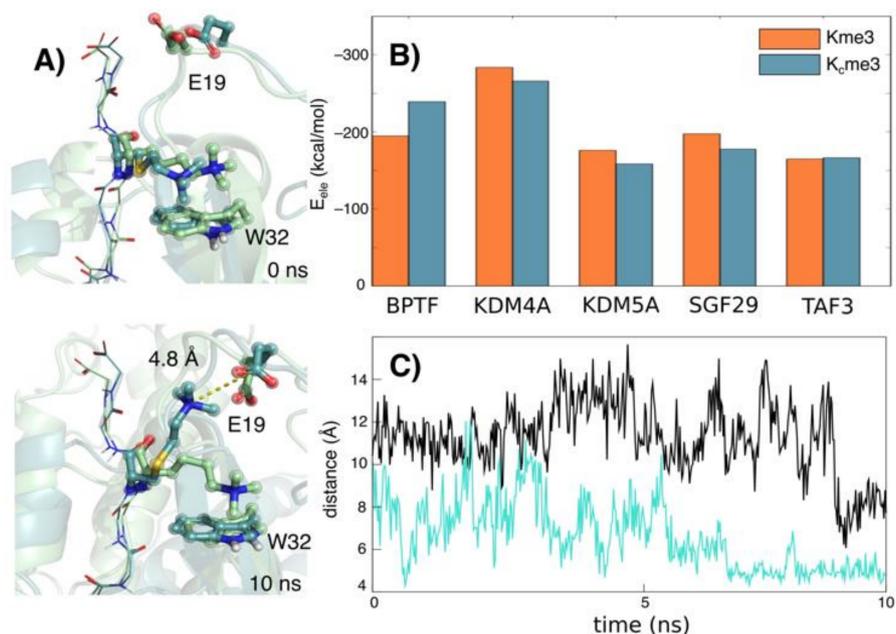
We then examined both peptides for association with five human reader proteins (KDM5A<sub>PHD3</sub>, PDB: 2KGI; TAF3<sub>PHD</sub>, PDB: 2K17; BPTF<sub>PHD</sub>, PDB: 2F6J; SGF29<sub>TTD</sub>, PDB: 3ME9; KDM4A<sub>TTD</sub>, PDB: 2GFA) using isothermal titration calorimetry (ITC). This panel of epigenetic reader proteins has been characterized by means of structural determination and binding with H3K4me<sub>3</sub>, and provides a diverse composition and architecture of the aromatic cage [9]. ITC studies provided thermodynamic parameters (Gibbs free energy of binding  $\Delta G^\circ$ , enthalpy of binding  $\Delta H^\circ$ , entropy of binding  $\Delta S^\circ$ ) for the association between five reader proteins and the two histone peptides (Table 1, Table S1 and Figure S1). We found that all examined reader proteins associate with the H3K<sub>C</sub>4me<sub>3</sub> peptide with comparable dissociation constants to the natural H3K4me<sub>3</sub> sequence, and that for both histone peptides the association with reader proteins is enthalpy-driven, a result that we attribute to the presence of several energetically favorable noncovalent interactions. In three cases, a slight decrease (~2-fold) in binding affinity was observed, whereas for TAF3<sub>PHD</sub> and KDM4A<sub>TTD</sub> a 2-fold increase in binding affinity was found. Differences in  $\Delta G^\circ$  were small, in the range of  $-0.5$  to  $0.5$  kcal mol<sup>-1</sup>. In all cases, except for KDM4A<sub>TTD</sub>, it was observed that  $\Delta H^\circ$  was more unfavorable, while  $\Delta S^\circ$  is more favorable when going from H3K4me<sub>3</sub> to H3K<sub>C</sub>4me<sub>3</sub>. With TAF3<sub>PHD</sub>, the gain in  $\Delta S^\circ$  is large enough to obtain an overall stronger binding affinity for H3K<sub>C</sub>4me<sub>3</sub>. These results are in agreement with earlier binding studies on different binding proteins for trimethyllysine and trimethylthialysine [24], and differ from recent findings showing that binding of BPTF<sub>PHD</sub> leads to a much larger decrease in binding affinity (~13-fold) when comparing H3K4me<sub>3</sub> and H3K<sub>C</sub>4me<sub>3</sub> [25]. In support of the former observation, we found that even longer trimethylhomolysine is very well recognized by a panel of the same five reader proteins [10], indicating that the slightly longer C–S–C moiety (compared to C–C–C core) does not significantly alter the readout process, supporting the finding that BPTF<sub>PHD</sub> indeed should recognize trimethylthialysine well.

Following thermodynamic analyses of reader–H3K4me<sub>3</sub> and reader–H3K<sub>C</sub>4me<sub>3</sub> binding, we carried out molecular dynamics (MD) simulations to provide an insight on flexibility of these five readers when complexed with H3K<sub>C</sub>4me<sub>3</sub> and H3K4me<sub>3</sub> (Figure 3). Starting structures were built by manually replacing the Kme<sub>3</sub> residue of H3K4me<sub>3</sub> with K<sub>C</sub>me<sub>3</sub> in crystal structures of reader proteins, solvated in a 10 Å truncated octahedral box of TIP3P water [26], and neutralized explicitly with either sodium or chloride ions. AMBER12 [27] was then used to simulate the systems for 10 ns each, as previously described [11].

Molecular mechanics-generalized born surface area (MM-GBSA) binding free energy calculations were performed at 500 ps intervals over 10 ns to examine effects on electrostatically dominated cation– $\pi$  interactions among the systems (Figure 3B, Table S1). The electrostatic contribution ( $\Delta E_{\text{ele}}$ ) to the binding free energy was found to be slightly more favorable for the Kme<sub>3</sub> compared to K<sub>C</sub>me<sub>3</sub> for all systems except BPTF<sub>PHD</sub> and TAF3<sub>PHD</sub>. For BPTF<sub>PHD</sub>, this can be explained by an unexpected stabilization of the K<sub>C</sub>me<sub>3</sub> terminal ammonium group ~10 Å away from the  $\pi$ -face of W32. Interestingly, the positively charged residue was instead observed interacting with the negatively charged Glu19 residue side chain located on a flexible loop region of BPTF<sub>PHD</sub> (Figure 3A). After 5 ns, K<sub>C</sub>me<sub>3</sub> formed the interaction with Glu19 for the remainder of the simulation, unlike Kme<sub>3</sub>, shown by a distance vs. time plot of the N <sub>$\epsilon$</sub> <sup>+</sup> atom distance of both methylated lysines to the Glu19 carboxy group (Figure 3C). In the case of TAF3<sub>PHD</sub>, similar  $\Delta E_{\text{ele}}$  values for K<sub>C</sub>me<sub>3</sub> and Kme<sub>3</sub> were likely a result of a closer interaction of K<sub>C</sub>me<sub>3</sub> with W868–W891, rather than a residue located externally from the aromatic cage (Figure S5). A high degree of flexibility observed by the H3 backbone atoms in the simulation with TAF3<sub>PHD</sub>, but could have facilitated prioritization of the cation– $\pi$  interaction.

**Table 1.** Thermodynamic parameters for association of the 10-mer H3K4me3 and H3K<sub>C</sub>4me3 peptides (ART(Kme3/K<sub>C</sub>me3)QTARKS) with epigenetic reader proteins (values and errors were obtained from 3–5 repeated ITC experiments carried out at 298.15 K).

	H3K4me3				H3K <sub>C</sub> 4me3			
	$K_d$ ( $\mu\text{M}$ )	$\Delta G^\circ$ ( $\text{kcal mol}^{-1}$ )	$\Delta H^\circ$ ( $\text{kcal mol}^{-1}$ )	$-T\Delta S^\circ$ ( $\text{kcal mol}^{-1}$ )	$K_d$ ( $\mu\text{M}$ )	$\Delta G^\circ$ ( $\text{kcal mol}^{-1}$ )	$\Delta H^\circ$ ( $\text{kcal mol}^{-1}$ )	$-T\Delta S^\circ$ ( $\text{kcal mol}^{-1}$ )
KDM5A <sub>PHD</sub>	0.071 ± 0.008	−9.7 ± 0.1	−10.7 ± 0.1	1.0 ± 0.1	0.15 ± 0.2	−9.3 ± 0.1	−9.6 ± 0.1	0.3 ± 0.1
TAF3 <sub>PHD</sub>	0.084 ± 0.012	−9.6 ± 0.1	−10.7 ± 0.1	1.1 ± 0.1	0.042 ± 0.007	−10.1 ± 0.1	−10.8 ± 0.1	0.7 ± 0.1
BPTF <sub>PHD</sub>	1.9 ± 0.2	−7.8 ± 0.1	−12.4 ± 0.1	4.6 ± 0.1	3.8 ± 0.5	−7.4 ± 0.1	−9.0 ± 0.1	1.6 ± 0.1
SGF29 <sub>TTD</sub>	2.6 ± 0.3	−7.6 ± 0.1	−8.0 ± 0.1	0.4 ± 0.1	6.1 ± 0.7	−7.1 ± 0.1	−5.8 ± 0.1	−1.3 ± 0.2
KDM4A <sub>TTD</sub>	6.6 ± 0.8	−7.1 ± 0.1	−13.0 ± 0.2	5.9 ± 0.2	3.1 ± 0.6	−7.5 ± 0.1	−14.8 ± 0.2	7.3 ± 0.2



**Figure 3.** (A) Snapshots of BPTF<sub>PHD</sub> complexed with H3 tail backbone (lines) containing K<sub>C</sub>me3 (cyan) and Kme3 (green) active sites at 0 and 10 ns; (B)  $\Delta E_{ele}$  contributions of ligands Kme3 and K<sub>C</sub>me3 from MM-GBSA binding free energy calculations; (C) distance vs. time plot of side N $\epsilon^+$  atoms of Kme3 (black) and K<sub>C</sub>me3 (cyan) to carboxyl group of BPTF<sub>PHD</sub> residue Glu19 (E19).

The binding poses of residues Kme3 and K<sub>C</sub>me3 when complexed with all five reader proteins are shown at times 0, 5 and 10 ns, with corresponding distance vs. time plots of each N $\epsilon^+$  to the centroid of the aromatic cage residue side chains (Figures S2–S6 and Tables S2–S3). All K<sub>C</sub>me3 complexes maintained a cation– $\pi$  interaction for a significant part of the simulation time with each of the aromatic cage residues, based on an established geometric cut-off of 6 Å [28], apart from BPTF<sub>PHD</sub>, and SGF29<sub>TTD</sub> residue F264, for which neither K<sub>C</sub>me3 nor Kme3 formed this interaction (Figure S6). Overall, the binding poses were extremely similar, likely due to the structural similarities of the K<sub>C</sub>me3 and Kme3 trimethylammonium placement in the aromatic cages.

Next, we have quantum-chemically analyzed the energetics and bonding mechanism of TRP2 (a model for the two tryptophan residues of KDM5A<sub>TTD</sub>) with K<sub>C</sub>me3, and for comparison, Kme3, using dispersion-corrected density functional theory at BLYP-D3BJ/TZ2P and COSMO for simulating the aqueous solution [29]. Our model complexes cover those moieties of the KDM5A–H3K4me3 X-ray structure that contribute to the intermolecular interaction in the full reader–histone complexes. TRP2–Kme3 was terminated with one hydrogen at C $\beta$  of the Kme3 side chain and one hydrogen at each C $\beta$  of the TRP2 fragment. For K<sub>C</sub>me3, we have used the same X-ray structure as for Kme3, but, with one CH<sub>2</sub> substituted by an S atom. To simulate the structural rigidity that is imposed by the protein backbone in the full protein system, the TRP2 fragment was kept frozen to the X-ray structure, both as a separate fragment and in the complexes. The Kme3 fragment was fully optimized, not only as the isolated molecule but also as the molecular fragment in its complex with TRP2. The K<sub>C</sub>me3 fragment was also fully optimized as the isolated molecule; but in the complex with TRP2, the carbon of the  $\alpha$ -methyl group was kept constrained at the same position with respect to TRP2 as the  $\alpha$ -methyl group in the TRP2–Kme3. The latter constraint simulates that, in the real complex, the remainder of K<sub>C</sub>me3, which is not present in our simple model, is kept at its position relative to KDM5A<sub>TTD</sub> through intermolecular interactions in the same way as Kme3. The geometry of the optimized Kme3 model system differs only very slightly from the X-ray structure.

The new TRP2–K<sub>C</sub>me3 model complex presents a similar, although 2.1 kcal mol<sup>−1</sup> weaker, bonding interaction than TRP2–Kme3:  $\Delta E(aq) = -8.1$  and  $-10.2$  kcal mol<sup>−1</sup> for TRP2–K<sub>C</sub>me3 and TRP2–Kme3 complexes, respectively (Table 2). This finding supports experimental observations and MD simulations

that assign a quite comparable binding affinity for both of them. The geometries of the two model systems are similar, with  $\text{NMe}_3^+$  in TRP2–Kme3 a bit closer to the TRP2 tryptophan cage than that in TRP2– $\text{K}_\text{C}\text{me}_3$ . The shortest  $\text{H}\cdots\text{C}$  distance between a H atom of the  $\text{K}_\text{C}\text{me}_3$   $\text{NMe}_3^+$  group and a C atom of the tryptophan (in the 5-membered ring) of the TRP2– $\text{K}_\text{C}\text{me}_3$  model is 2.88 Å, which has to be compared with the corresponding shortest  $\text{H}\cdots\text{C}$  distance of 2.78 Å in the TRP2–Kme3 complex (Table 2 and Figure S7). Likewise, another H atom of the same methyl group yields the shortest  $\text{H}\cdots\text{C}$  distance to the other TRP unit (6-membered ring), and this distance is also slightly longer in TRP2– $\text{K}_\text{C}\text{me}_3$  (2.94 Å) compared to 2.88 Å for TRP2–Kme3. Thus, the slightly longer intermolecular  $\text{H}\cdots\text{C}$  distances reflect the slightly weaker interaction in TRP2– $\text{K}_\text{C}\text{me}_3$  as compared to TRP2–Kme3.

**Table 2.** Quantum-chemical bonding analysis (energies in  $\text{kcal mol}^{-1}$ , distances in Å) in TRP2–Kme3 and TRP2– $\text{K}_\text{C}\text{me}_3$  systems in aqueous solution. <sup>1</sup>

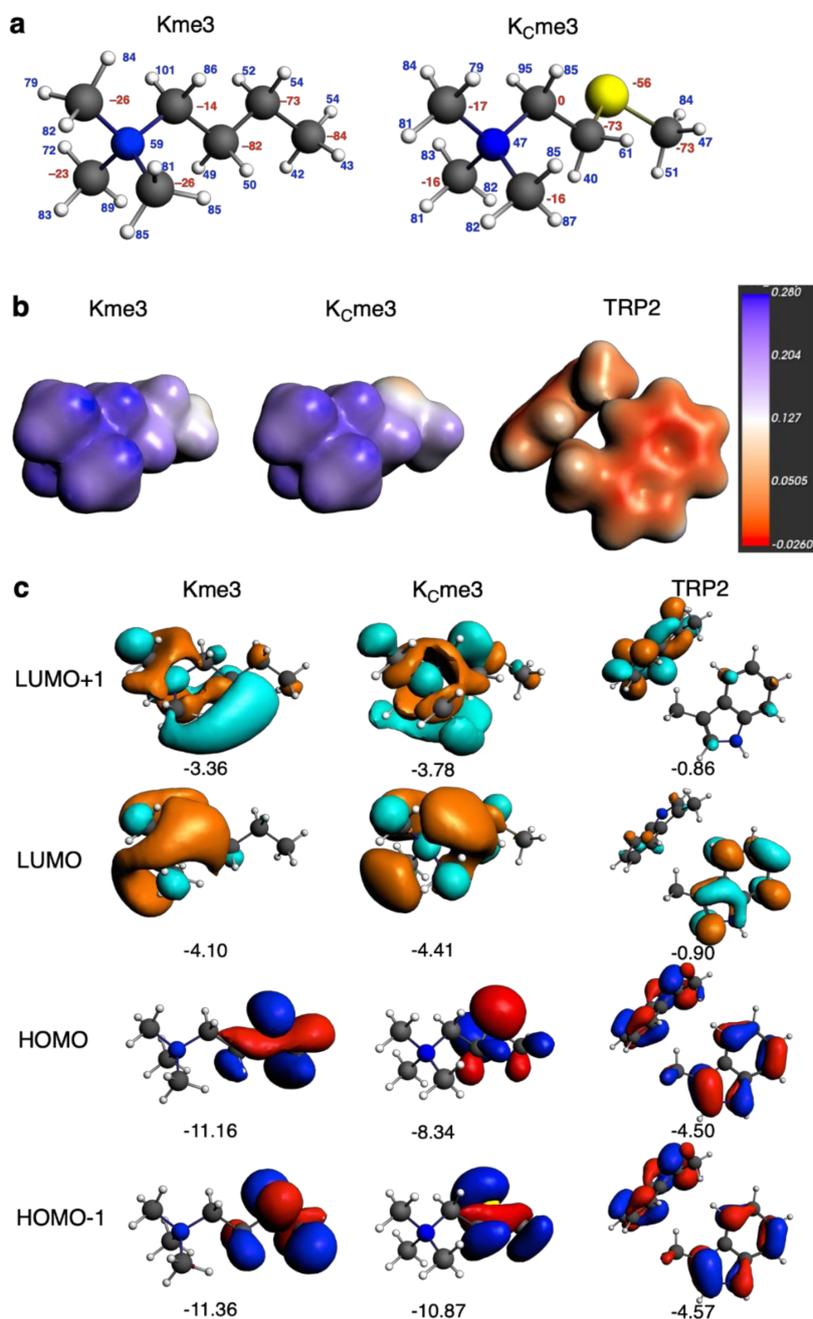
	TRP2–Kme3 <sup>2</sup>	TRP2– $\text{K}_\text{C}\text{me}_3$ <sup>3</sup>
$\Delta E(\text{aq})$	−10.2	−8.1
$\Delta E_{\text{strain}}(\text{aq})$	0.1	2.6
$\Delta E_{\text{int}}(\text{aq})$	−10.3	−10.7
$\Delta E_{\text{int}}(\text{desolv})$	17.3	18.8
$\Delta E_{\text{int}}$	−27.6	−29.5
$\Delta E_{\text{Pauli}}$	20.8	24.3
$\Delta V_{\text{elstat}}$	−15.0	−17.0
$\Delta E_{\text{oi}}$	−13.0	−14.4
$\Delta E_{\text{disp}}$	−20.4	−22.3
$d(\text{H}_{\text{Me}}-\text{C}_{\text{TRP-6MR}})$	2.88	2.94
$d(\text{H}_{\text{Me}}-\text{C}_{\text{TRP-5MR}})$	2.78	2.88

<sup>1</sup> Computed at BLYP-D3BJ/TZ2P with COSMO to simulate aqueous solution. Structural rigidity imposed by the protein backbone is simulated through constrained geometry optimizations. See also Equations (1)–(3) in the Experimental section. <sup>2</sup> TRP2 frozen, Kme3 entirely free. <sup>3</sup> TRP2 frozen,  $\alpha$ -methyl carbon fixed to position TRP2–Kme3 optimization.

Interestingly, our bonding analyses reveal that the instantaneous interaction energies  $\Delta E_{\text{int}}(\text{aq})$  are in reverse order as compared to the net bond energies, although differences are relatively small:  $\Delta E_{\text{int}}(\text{aq})$  amounts to −10.7 and −10.3  $\text{kcal mol}^{-1}$  for TRP2– $\text{K}_\text{C}\text{me}_3$  and TRP2–Kme3, respectively (Table 2). This order in instantaneous interaction energies is inverted by the slightly more destabilizing strain energy  $\Delta E_{\text{strain}}(\text{aq})$  associated with geometrical deformations of  $\text{K}_\text{C}\text{me}_3$  in TRP2– $\text{K}_\text{C}\text{me}_3$ . Thus, while this strain is negligible (only 0.1  $\text{kcal mol}^{-1}$ ) for TRP2–Kme3, it becomes 2.6  $\text{kcal mol}^{-1}$  in the case of TRP2– $\text{K}_\text{C}\text{me}_3$ . The larger  $\Delta E_{\text{strain}}(\text{aq})$  in the latter case is a consequence of the position of the S atom, which escapes from the linearity of the zig-zag shape of the  $\text{K}_\text{C}\text{me}_3$  system when it interacts with TRP2, at variance with Kme3 (Figure S8). When  $\text{K}_\text{C}\text{me}_3$  is allowed to fully relax, without TRP2, the same linearity as for Kme3 is achieved. For comparison, the CCCC dihedral angle in Kme3 in the complex is 177.6°, whereas the equivalent CCSC dihedral in  $\text{K}_\text{C}\text{me}_3$  is 168.2°. This deviation is responsible for the larger strain energy in TRP2– $\text{K}_\text{C}\text{me}_3$  and the corresponding weaker  $\Delta E_{\text{int}}(\text{aq})$  discussed above. This deviation also occurs if  $\text{K}_\text{C}\text{me}_3$  is allowed to fully relax, i.e., without frozen  $\text{C}_\alpha$  position. The interaction energy  $\Delta E_{\text{int}}$  between the exact same structures but in the absence of aqueous solvation is again in favor of TRP– $\text{K}_\text{C}\text{me}_3$ , namely, by 1.9  $\text{kcal mol}^{-1}$ . The desolvation incurred upon binding is 1.5  $\text{kcal mol}^{-1}$  more destabilizing for the  $\text{K}_\text{C}\text{me}_3$  than for the Kme3 complex, because of the presence of a sulfur atom, instead of  $\text{CH}_2$  group, in the former.

The more stabilizing intrinsic interaction energy of TRP2 with  $\text{K}_\text{C}\text{me}_3$  ( $\Delta E_{\text{int}} = -29.5 \text{ kcal mol}^{-1}$ ) than with Kme3 ( $\Delta E_{\text{int}} = -27.6 \text{ kcal mol}^{-1}$ ) is further analyzed using quantitative Kohn–Sham molecular orbital (KS-MO) and an associated canonical energy decomposition analysis (EDA); see Table 2. This analysis reveals more favorable electrostatic, orbital and dispersion interactions as the origin of the stronger interaction  $\Delta E_{\text{int}}$  term in TRP2– $\text{K}_\text{C}\text{me}_3$ . The more attractive electrostatic interaction  $\Delta V_{\text{elstat}}$  is due to the presence of the negatively charged sulfur atom, which comes in close proximity

to the positively charged H atoms of one of the TRP units (see the Voronoi Deformation Density (VDD) charges in Figure 4A). This is also supported by the molecular electrostatic potential isosurfaces for Kme3, K<sub>C</sub>me3 and TRP2 (Figure 4B) [30,31]. The S atom in K<sub>C</sub>me3 appears redder (more towards negative), whereas its trimethylated group is bluer (more positive) than for Kme3, which favors the interaction of the former with TRP2.



**Figure 4.** Computational analysis of Kme3, K<sub>C</sub>me3 and TRP2, computed at BLYP-D3BJ/TZ2P using frozen on X-ray structures for TRP2 and positions of C<sub>α</sub> and full geometrical relaxation for all other portions: (a) Voronoi Deformation Density (VDD) atomic charges (in milli-a.u.; red = negative, blue = positive); (b) molecular electrostatic potential isosurfaces for Kme3, K<sub>C</sub>me3 and TRP2 (in a.u.); (c) frontier orbitals (isosurface at 0.03) and orbital energies (in eV).

Likewise, a favorable spatial configuration also enhances the attractive  $\Delta E_{oi}$  term in TRP2–K<sub>C</sub>me3 through a larger orbital overlap and thus more stabilizing donor–acceptor orbital interactions between

occupied TRP2  $\pi$  orbitals and empty acceptor  $\sigma^*_{C-H}$  type orbitals on the  $K_Cme3$  side chain (Figure 4C; three out of the four overlaps between the frontier FMOs are larger for TRP2- $K_Cme3$ , as can be seen in Table S5) [32,33]. The torsion of the zig-zag chain caused by the S atom makes the three central members of the side chain of  $K_Cme3$  (S-CH<sub>2</sub>-CH<sub>2</sub>) to be closer to one of the TRP2 units in TRP2- $K_Cme3$  than the equivalent CH<sub>2</sub>-CH<sub>2</sub>-CH<sub>2</sub> side chain for  $Kme3$  in TRP2- $Kme3$ . This translates into a more destabilizing closed-shell-closed-shell Pauli repulsion  $\Delta E_{Pauli}$  in the TRP2- $K_Cme3$  complex. Taken together, the TRP2- $K_Cme3$  complex differs only slightly in stability from the TRP2- $Kme3$  complex. The former appears to be only slightly less stable due to somewhat more pronounced structural deformations reflected by  $\Delta E_{strain}(aq)$  and a somewhat more endothermic desolvation  $\Delta E_{int}(desolv)$  (Table 2).

### 3. Discussion

To assess whether trimethylthialysine can serve as an easily accessible and functional analogue of trimethyllysine when studying molecular recognition processes in epigenetics, we have carried out comparative thermodynamic analyses, molecular dynamics simulations and quantum chemical analyses for binding of epigenetic reader proteins with histone peptides bearing trimethyllysine and trimethylthialysine. ITC data showed that H3K4me3 and H3K<sub>C</sub>4me3 exhibit similar binding affinities for a panel of five human reader proteins. These observations were further supported by molecular dynamics simulations that demonstrated that the binding poses for the two residues were very similar, and by quantum chemical analyses that showed that similar bonding interactions are present for  $Kme3$  and  $K_Cme3$  with TRP2. The study further establishes trimethylthialysine as a widely applicable trimethyllysine mimic, which can be used to study genuinely important biomolecular processes that involve the methylation of lysine in a more precise fashion, in vitro and in the cellular environment, thereby helping to unravel the complex language of post-translational modifications and its role on the nucleosome and higher order chromatin structures in biology.

### 4. Materials and Methods

#### 4.1. Preparation of H3K4me3 and H3K<sub>C</sub>4me3

H3K4me3 and H3K<sub>C</sub>4 were synthesized by solid-phase peptide synthesis (Schemes S1 and S2). Cysteine was alkylated to produce trimethylated lysine analogue ( $K_Cme3$ ) as described previously [22]. The reaction was performed using thermomixer (Eppendorf Thermomixer R, Hamburg, Germany). For synthesizing the 1–10 H3K<sub>C</sub>4me3: 50 mg of purified and lyophilized unalkylated peptide was dissolved in 4.9 mL alkylation buffer (4 M GuHCl, 1 M HEPES pH 7.8 and 10 mM D/L-methionine) and allowed to incubate for 1 h at 37 °C under reducing conditions by adding 100  $\mu$ L 1 M DTT. (2-Bromoethyl) trimethylammonium bromide was directly dissolved into the reaction mixture and allowed to react at 50 °C. After 2.5 h reaction time, 10  $\mu$ L 1 M DTT was added to the reaction mixture and the reaction was allowed to proceed for another 2.5 h. The reaction was quenched by incubating the reaction mixture with 25  $\mu$ L of 2-mercaptoethanol for 30 min at room temperature, and then directed to freeze-dryer overnight followed by prep-HPLC purification (Figures S9–S11).

#### 4.2. Reader Domain expression and Purification

The methylated reader proteins were expressed and purified as previously described [29].

#### 4.3. Isothermal Titration Calorimetry

Reader proteins used for ITC experiments were produced as described [9]. ITC experiments were carried out at 298.15 K using a fully automated MicroCal Auto-iTC200 (GE Healthcare, Northampton, MA, USA). Histone peptides and reader proteins were dissolved in the same buffer (used in size exclusion chromatography). Each ITC titration consisted of 19 injections of histone peptide (0.3–1.2 mM) to reader protein (22–100  $\mu$ M). Experiments were repeated 3 to 5 times. Heats of dilution for histone

peptides were determined in control experiments, and were subtracted from the titration binding data before curve fitting. Curve fitting was performed using Origin 6.0 (Microcal Inc., Northampton, MA, USA) with a one-site note.

#### 4.4. Molecular Dynamics Simulations

Ten MD simulations were carried out for 10 ns each. PDB structures for the models representing TAF3<sub>PHD</sub> (PDB: 2K17), KDM4A<sub>TTD</sub> (PDB: 2GFA), KDM5A<sub>PHD3</sub> (PDB: 2KGI), BPTF<sub>PHD</sub> (PDB: 2F6J), and SGF29<sub>TTD</sub> (PDB: 3ME9) reader proteins were used as a template for building the reader bound to K<sub>C</sub>me<sub>3</sub> and Kme<sub>3</sub> systems.

Hydrogen atom addition was performed with *tLeap* [34]. Systems were solvated in a truncated octahedral box of TIP3P [25] that extended at least 10 Å from protein atoms and neutralized explicitly with either Na<sup>+</sup> or Cl<sup>-</sup> counterions.

AMBER12 (San Francisco, CA, USA, 2012) [26,27] was used with the Amberff12SB force field to define protein partial charges. The force constants for bond, angle, and torsions for the atoms bonded to zinc derived from the Zinc AMBER Force Field (ZAFF) developed by the Merz group [35]. Atomic partial charges for each atomic center in K<sub>C</sub>me<sub>3</sub> correspond to those derived using the Restrained Electrostatic Potential (RESP) [36] module in AmberTools [37] calculated using HF/6–31G(d), shown in Table S3. Parameters for Kme<sub>3</sub> were previously derived also using the RESP methodology [11].

The ten systems were then minimized in two steps. First came 1000 steps of steep descent and 1000 steps of conjugate gradient minimization wherein the protein was held fixed by using position restraints with a force constant of 500 kcal mol<sup>-1</sup> Å<sup>-2</sup>. This was repeated without the position restraints. The system was then heated for 1 ns from 0 to 310 K under constant volume periodic boundary conditions (NVT). Then, 1 ns of equilibration under constant pressure and temperature (NPT) was performed. Following this, 10 ns molecular dynamics simulations were then performed.

Langevin thermostat [38] was used to simulate a constant temperature of 310 K with collision frequency of 1 ps<sup>-1</sup>. The SHAKE algorithm [39] was turned on to constrain all bonds involving hydrogen and 2 fs was defined as the time step for numerical integration. Isotropic position scaling was used to maintain the pressure of 1 atm ( $\tau_p = 2$  ps). The particle mesh Ewald summation method [40] was employed to enforce an 8.0 Å cutoff for non-bonded long-range and electrostatic interactions.

Trajectories were visualized using visual molecular dynamics (VMD 1.9.2., Champaign, IL, USA) [41]. Endstate free energies and electrostatic energies were calculated using MM-GBSA calculations [42]. Energy values were measured every 500 ps over 10 ns. A salt concentration of 0.15 M was used to parallel physiological conditions. For cation distance calculations, the  $\pi$ -system was defined for aromatic cage residues as the centroid of the side chain aromatic (non-H) atoms.

#### 4.5. Quantum Chemical Analysis

All calculations were carried out with the Amsterdam Density Functional (ADF 2018, Amsterdam, The Netherlands) program using dispersion-corrected density functional theory at the BLYP-D3BJ/TZ2P level of theory (Table S5) [43]. The effect of aqueous solvation was simulated by means of the conductor like screening model (COSMO) of solvation, as implemented in ADF. The approach has been benchmarked against highly correlated post-Hartree–Fock methods and experimental data and was found to work reliably [44–49].

The bonding mechanism in our model complexes have been further analyzed using quantitative (Kohn–Sham) molecular orbital (MO) theory in combination with an energy decomposition analysis (EDA) [50–52]. The bond energy in aqueous solution  $\Delta E(\text{aq})$  consists of two major components; namely, the strain energy  $\Delta E_{\text{strain}}(\text{aq})$  associated with deforming the Kme<sub>3</sub> and the reader from their own equilibrium structure to the geometry they adopt in the complex, plus the interaction energy  $\Delta E_{\text{int}}(\text{aq})$  between these deformed solutes in the complex (see Equation (1)):

$$\Delta E(\text{aq}) = \Delta E_{\text{strain}}(\text{aq}) + \Delta E_{\text{int}}(\text{aq}). \quad (1)$$

To arrive at an understanding of the importance of desolvation phenomena during the complexation process, we separate the solute–solute interaction  $\Delta E_{\text{int}}(\text{aq})$  into the effect caused by the change in solvation  $\Delta E_{\text{int}}(\text{desolv})$  and the remaining intrinsic interaction  $\Delta E_{\text{int}}$  between the unsolvated fragments in vacuum  $\Delta E_{\text{int}}$ :

$$\Delta E_{\text{int}}(\text{aq}) = \Delta E_{\text{int}}(\text{desolv}) + \Delta E_{\text{int}}. \quad (2)$$

In the EDA, the intrinsic interaction energy  $\Delta E_{\text{int}}$  can be further decomposed, as shown in Equation (3):

$$\Delta E_{\text{int}} = \Delta V_{\text{elstat}} + \Delta E_{\text{Pauli}} + \Delta E_{\text{oi}} + \Delta E_{\text{disp}}. \quad (3)$$

Here,  $\Delta V_{\text{elstat}}$  corresponds to the classical electrostatic interaction between the unperturbed charge distributions of the deformed fragments, which is usually attractive. The Pauli repulsion  $\Delta E_{\text{Pauli}}$  comprises the destabilizing interactions between occupied orbitals and is responsible for the steric repulsions. The orbital interaction  $\Delta E_{\text{oi}}$  accounts for charge transfer (donor–acceptor interactions between occupied orbitals on one moiety with unoccupied orbitals of the other, including the HOMO–LUMO interactions) and polarization (empty/occupied orbital mixing on one fragment due to the presence of another fragment). Finally, the  $\Delta E_{\text{disp}}$  term accounts for the dispersion interactions based on Grimme’s DFT-D3BJ correction. Furthermore, the charge distribution has been analyzed using the Voronoi deformation density (VDD) method [53].

**Supplementary Materials:** The following are available online. Scheme 1: Solid-phase synthesis of histone peptide H3K4me3. Figure S1: ITC data. Figure S2: MD simulations of BPTF<sub>PHD</sub>. Figure S3: MD simulations of KDM4A<sub>TTD</sub>. Figure S4: MD simulations of KDM5A<sub>PHD3</sub>. Figure S5: MD simulations of TAF3<sub>PHD</sub>. Figure S6: MD simulations of SGF29<sub>TTD</sub>. Figure S7: Top view of structure of TRP2–Kme3 and TRP–K<sub>C</sub>me3 model complexes. Figure S8: Front view of the structure of TRP–Kme3 and TRP2–K<sub>C</sub>me3 model complexes. Figure S9: LC-MS analysis of 1–10 H3C4 after RP-HPLC purification. Figure S10: LC-MS analysis of 1–10 H3K<sub>C</sub>4me3 after RP-HPLC purification. Figure S11: LC-MS analysis of 1–10 H3K4me3 after RP-HPLC purification. Table S1 Concentrations of protein and peptide, with C-value and N binding sites in ITC binding studies. Table S2: MM-GBSA binding free energies and electrostatic contributions calculated for Kme3 and K<sub>C</sub>me3 complexed with reader proteins over 10 ns at 500 ps intervals. Table S3: Average root mean square deviation (RMSD) and error of C<sub>α</sub> atoms of reader proteins. Table S4: Cartesian coordinates and charges calculated using the RESP method HF/6–31G\* of modified K<sub>C</sub>me3. Table S5: Overlaps between the MOs of TRP and Kme3 or K<sub>C</sub>me3. Table S6: Cartesian coordinates (in Å) of TRP2–Kme3 and TRP2–K<sub>C</sub>me3 complexes, computed at BLYP-D3BJ/TZ2P using COSMO to simulate aqueous solvation and a constrained optimization to simulate the effect of the protein backbone.

**Author Contributions:** Supervision, J.M.; finalization of project, J.C.J.H.; quantum chemical analyses, J.P. and F.M.B.; molecular dynamics simulations, K.K. and R.S.P.; peptide synthesis, A.H.K.A.T.; protein preparation, B.J.G.E.P.; thermodynamic analyses, J.M.; writing—original draft preparation, J.C.J.H., J.P., K.K., R.S.P., F.M.B. and J.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the European Research Council (ERC Starting Grant, ChemEpigen–715691 to J.M.), Spanish MINECO (CTQ2016–77558-R and MDM–2017–0767 to J.P.), the Generalitat de Catalunya (2017SGR348 to J.P.) and a World Bank Education Grant to K.K.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Strahl, B.D.; Allis, C.D. The language of covalent histone modifications. *Nature* **2000**, *403*, 41–45. [PubMed]
2. Bannister, A.J.; Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Res.* **2011**, *21*, 381–395. [PubMed]
3. Black, J.C.; van Rechem, C.; Whetstine, J.R. Histone lysine methylation dynamics. *Mol. Cell* **2012**, *48*, 491–507. [PubMed]
4. Qian, C.; Zhou, M.M. SET domain protein lysine methyltransferases: Structure, specificity and catalysis. *Cell. Mol. Life Sci.* **2006**, *63*, 2755–2763. [PubMed]
5. Kooistra, S.M.; Helin, K. Molecular mechanisms and potential functions of histone demethylases. *Nat. Rev. Mol. Cell Biol.* **2012**, *13*, 297–311.

6. Taverna, S.D.; Li, H.; Ruthenburg, A.J.; Allis, C.D.; Patel, D.J. How chromatin-binding modules interpret histone modifications: Lessons from professional pocket pickers. *Nat. Struct. Mol. Biol.* **2007**, *14*, 1025–1040.
7. Nadal, S.; Raj, R.; Mohammed, S.; Davis, B.G. Synthetic post-translational modifications of histones. *Curr. Opin. Chem. Biol.* **2018**, *45*, 35–47.
8. Hughes, R.M.; Wiggins, K.R.; Khorasanizadeh, S.; Waters, M.L. Recognition of trimethyllysine by a chromodomain is not driven by the hydrophobic effect. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 11184–11188.
9. Kamps, J.J.; Huang, J.; Poater, J.; Xu, C.; Pieters, B.J.G.E.; Dong, A.; Min, J.; Sherman, W.; Beuming, T.; Bickelhaupt, F.M.; et al. Chemical basis for the recognition of trimethyllysine by epigenetic reader proteins. *Nat. Commun.* **2015**, *6*, 8911.
10. Al Temimi, A.H.K.; Belle, R.; Kumar, K.; Poater, J.; Betlem, P.; Pieters, B.J.G.E.; Paton, R.S.; Bickelhaupt, F.M.; Mecinovic, J. Recognition of shorter and longer trimethyllysine analogues by epigenetic reader proteins. *Chem. Commun.* **2018**, *54*, 2409–2412.
11. Belle, R.; Al Temimi, A.H.K.; Kumar, K.; Pieters, B.J.G.E.; Tumber, A.; Dunford, J.E.; Johansson, C.; Oppermann, U.; Brown, T.; Schofield, C.J.; et al. Investigating d-lysine stereochemistry for epigenetic methylation, demethylation and recognition. *Chem. Commun.* **2017**, *53*, 13264–13267.
12. Lang, K.; Chin, J.W. Cellular incorporation of unnatural amino acids and bioorthogonal labeling of proteins. *Chem. Rev.* **2014**, *114*, 4764–4806.
13. Wang, L.; Brock, A.; Herberich, B.; Schultz, P.G. Expanding the genetic code of *Escherichia coli*. *Science* **2001**, *292*, 498–500. [[PubMed](#)]
14. Wang, Q.; Parrish, A.R.; Wang, L. Expanding the genetic code for biological studies. *Chem. Biol.* **2009**, *16*, 323–326. [[PubMed](#)]
15. Hendrickson, T.L.; de Crecy-Lagard, V.; Schimmel, P. Incorporation of unnatural amino acids into proteins. *Annu. Rev. Biochem.* **2004**, *73*, 147–176. [[PubMed](#)]
16. Dawson, P.E.; Muir, T.W.; Clark-Lewis, I.; Kent, S.B. Synthesis of proteins by native chemical ligation. *Science* **1994**, *266*, 776–779.
17. Shimko, J.C.; North, J.A.; Bruns, A.N.; Poirier, M.G.; Ottesen, J.J. Preparation of fully synthetic histone H3 reveals that acetyl-lysine 56 facilitates protein binding within nucleosomes. *J. Mol. Biol.* **2011**, *408*, 187–204.
18. Le, D.D.; Cortesi, A.T.; Myers, S.A.; Burlingame, A.L.; Fujimori, D.G. Site-specific and regiospecific installation of methylarginine analogs into recombinant histones and insights into effector protein binding. *J. Am. Chem. Soc.* **2013**, *135*, 2879–2882.
19. Al Temimi, A.H.K.; van der Wekken-de Bruijne, R.; Proietti, G.; Guo, H.; Qian, P.; Mecinovic, J.  $\gamma$ -Thialysine versus lysine: An insight into the epigenetic methylation of histones. *Bioconjugate Chem.* **2019**, *30*, 1798–1804.
20. Proietti, G.; Rainone, G.; Hintzen, J.C.J.; Mecinovic, J. Exploring the histone acylome through incorporation of  $\gamma$ -thialysine on histone tails. *Bioconjugate Chem.* **2020**, *31*, 844–851.
21. Huang, R.; Holbert, M.A.; Tarrant, M.K.; Curtet, S.; Colquhoun, D.R.; Dancy, B.M.; Dancy, D.C.; Hwang, Y.; Tang, Y.; Meeth, K.; et al. Site-specific introduction of acetyl-lysine mimic into peptides and proteins by cysteine alkylation. *J. Am. Chem. Soc.* **2010**, *132*, 9986–9987. [[PubMed](#)]
22. Simon, M.D.; Chu, F.; Racki, L.R.; Cecile, C.; Burlingame, A.L.; Panning, B.; Narlikar, G.J.; Shokat, K.M. The site-specific installation of methyl-lysine analogs into recombinant histones. *Cell* **2007**, *128*, 1003–1012. [[PubMed](#)]
23. Pieters, B.J.G.E.; Hintzen, J.C.J.; Grobber, Y.; Al Temimi, A.H.K.; Kamps, J.J.A.G.; Mecinovic, J. Installation of Trimethyllysine analogs on intact histones via cysteine alkylation. *Bioconjugate Chem.* **2019**, *30*, 952–958.
24. Seeliger, D.; Soeroes, S.; Klingberg, R.; Schwarzer, D.; Grubmuller, H.; Fischle, W. Quantitative assessment of protein interaction with methyl-lysine analogues by hybrid computational and experimental approaches. *ACS Chem. Biol.* **2012**, *7*, 150–154. [[PubMed](#)]
25. Chen, Z.; Notti, R.Q.; Ueberheide, B.; Ruthenburg, A.J. Quantitative and structural assessment of histone methyllysine analogue engagement by cognate binding proteins reveals affinity decrements relative to those of native counterparts. *Biochemistry* **2018**, *57*, 300–304.
26. Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.
27. Case, D.A.; Darden, T.A.; Cheatham, T.E., III; Simmerling, C.L.; Wang, J.; Duke, R.E.; Luo, R.; Walker, R.C.; Zhang, W.; Merz, K.M.; et al. *AMBER 12*; University of California: San Francisco, CA, USA, 2012.

28. Gallivan, J.P.; Dougherty, D.A. Cation- $\pi$  interactions in structural biology. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 9459–9464.
29. te Velde, G.; Bickelhaupt, F.M.; Baerends, E.J.; Fonseca Guerra, C.; van Gisbergen, S.J.A.; Snijders, J.G.; Ziegler, T. Chemistry with ADF. *J. Comput. Chem.* **2001**, *22*, 931–967.
30. Motherwell, W.B.; Moreno, R.B.; Pavlakos, I.; Arendorf, J.R.T.; Arif, T.; Tizzard, G.J.; Coles, S.J.; Aliev, A.E. Noncovalent interactions of  $\pi$  systems with sulfur: The atomic chameleon of molecular recognition. *Angew. Chem. Int. Ed.* **2018**, *57*, 1193–1198.
31. Petraglia, R.; Corminboeuf, C. A Caveat on SCC-DFTB and noncovalent interactions involving sulfur atoms. *J. Chem. Theory Comput.* **2013**, *9*, 3020–3025.
32. van der Lubbe, S.C.C.; Fonseca Guerra, C. The nature of hydrogen bonds: A delineation of the role of different energy components on hydrogen bond strengths and lengths. *Chem. Asian J.* **2019**, *14*, 2760–2769.
33. Wolters, L.P.; Bickelhaupt, F.M. Halogen bonding versus hydrogen bonding: A molecular orbital perspective. *ChemistryOpen* **2012**, *1*, 96–105.
34. Pieters, B.J.G.E.; Belle, R.; Mecinovic, J. The effect of length on histone H3K4me3 on recognition by reader proteins. *ChemBioChem* **2013**, *14*, 2408–2412.
35. Chen, V.B.; Arendall, W.B.; Headd, J.J.; Keedy, D.A.; Immormino, R.M.; Kapral, G.J.; Murray, L.W.; Richardson, J.S.; Richardson, D.C. MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D.* **2010**, *66*, 12–21.
36. Peters, M.B.; Yang, Y.; Wang, B.; Fusti-Molnar, L.; Weaver, M.N.; Merz Jr, K.M. Structural survey of zinc containing proteins and the development of the zinc AMBER force field (ZAFF). *J. Chem. Theory Comput.* **2010**, *6*, 2935–2947.
37. Bayly, C.I.; Cieplak, P.; Cornell, W.; Kollman, P.A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: The RESP model. *J. Phys. Chem.* **1993**, *97*, 10269–10280.
38. Case, D.A.; Darden, T.A., III. *AMBER11*; University of California: San Francisco, CA, USA, 2010.
39. Loncharich, R.J.; Brooks, B.R.; Pastor, R.W. Langevin dynamic of peptides: The frictional dependence of isomerization rates of N-acetylanlyl-N'-methylamide. *Biopolymers* **1992**, *32*, 523–535.
40. Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H.J.C. Numerical integration of the cartesian equations of a motion of a system with constraints: Molecular dynamics of *n*-alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.
41. Wang, H.; Fang, J.; Gao, X. The optimal particle-mesh interpolation basis. *J. Chem. Phys.* **2017**, 124107.
42. Philips, J.C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R.D.; Kale, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802.
43. Miller III, B.R.; McGee Jr, T.D.; Swails, J.M.; Homeyer, N.; Gohlke, N.; Roitberg, A.E. MMBPSPy: An efficient program for end-state free energy calculations. *J. Chem. Theory Comput.* **2012**, *8*, 3314–3321.
44. Fonseca Guerra, C.; van der Wijst, T.; Poater, J.; Swart, M.; Bickelhaupt, F.M. Adenine versus guanine quartets in aqueous solution: Dispersion-corrected DFT study on the differences in  $\pi$ -stacking and hydrogen-bonding behavior. *Theor. Chem. Acc.* **2010**, *125*, 245–252.
45. van der Wijst, T.; Fonseca Guerra, C.; Swart, M.; Bickelhaupt, F.M.; Lippert, B.A. A ditopic ion-pair receptor based on stacked nucleobase quartets. *Angew. Chem. Int. Ed.* **2009**, *121*, 3285–3287.
46. Padial, J.S.; de Gelder, R.; Fonseca Guerra, C.; Bickelhaupt, F.M.; Mecinovic, J. Stabilisation of 2,6-diarylpyridinium cation by through-space polar- $\pi$  interactions. *Chem-Eur. J.* **2014**, *20*, 6268–6271.
47. Mardirossian, N.; Head-Gordon, M. Thirty years of density functional theory in computational chemistry: An overview and extensive assessment of 200 density functionals. *Mol. Phys.* **2017**, *19*, 2315–2372.
48. Brauer, B.; Kesharwani, M.K.; Kozuch, S.; Martin, J.M.L. The S66  $\times$  8 benchmark for noncovalent interactions revisited: Explicitly correlated ab initio methods and density functional theory. *Phys. Chem. Chem. Phys.* **2016**, *18*, 20905–20925.
49. Goerigk, L.; Hansen, A.; Bauer, C.; Ehrlich, S.; Najibi, A.; Grimme, S.A. A look at the density functional theory zoo with the advanced GMTKN55 database for general main group thermochemistry, kinetics and noncovalent interactions. *Phys. Chem. Chem. Phys.* **2017**, *19*, 32184–32215.
50. Bickelhaupt, F.M.; Baerends, E.J. *Reviews in Computational Chemistry*; Wiley-VCH: New York, NY, USA, 2000; Volume 15, pp. 1–86.
51. Bickelhaupt, F.M.; Diefenbach, A.; de Visser, S.P.; de Koning, L.J.; Nibberning, N.M.N. Nature of the three-electron bond in  $\text{H}_2\text{S} \cdot \text{SH}_2^+$ . *J. Phys. Chem. A* **1998**, *102*, 9549–9553.

52. Wolters, L.P.; Bickelhaupt, F.M. The activation strain model and molecular orbital theory. *WIREs. Comput. Mol. Sci.* **2015**, *5*, 324–343.
53. Fonseca Guerra, C.; Handgraaf, J.W.; Baerends, E.J.; Bickelhaupt, F.M. Voronoi deformation density (VDD) charges: Assessment of Mulliken, Bader, Hirshfeld, Weindhold, and VDD methods for charge analysis. *J. Comput. Chem.* **2004**, *25*, 189–210.

**Sample Availability:** Not available.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).