

A EUROPEAN JOURNAL OF CHEMICAL BIOLOGY

CHEM **BIO** CHEM

SYNTHETIC BIOLOGY & BIO-NANOTECHNOLOGY

Accepted Article

Title: Lysine Ethylation by Histone Lysine Methyltransferases

Authors: Abbas H. K. Al Temimi, Michael Martin, Qingxi Meng, Danny C. Lenstra, Ping Qian, Hong Guo, Elmar Weinhold, and Jasmin Mecinovic

This manuscript has been accepted after peer review and appears as an Accepted Article online prior to editing, proofing, and formal publication of the final Version of Record (VoR). This work is currently citable by using the Digital Object Identifier (DOI) given below. The VoR will be published online in Early View as soon as possible and may be different to this Accepted Article as a result of editing. Readers should obtain the VoR from the journal website shown below when it is published to ensure accuracy of information. The authors are responsible for the content of this Accepted Article.

To be cited as: *ChemBioChem* 10.1002/cbic.201900359

Link to VoR: <http://dx.doi.org/10.1002/cbic.201900359>

WILEY-VCH

www.chembiochem.org

A Journal of



Lysine Ethylation by Histone Lysine Methyltransferases

Abbas H. K. Al Temimi,^[a] Michael Martin,^[b] Qingxi Meng,^[c] Danny C. Lenstra,^[a] Ping Qian,^[c] Hong Guo,^{*[d,e]} Elmar Weinhold,^{*[b]} and Jasmin Mecinović^{*[a,f]}

^a Institute for Molecules and Materials, Radboud University, Heyendaalseweg 135, 6525 AJ Nijmegen, The Netherlands.

^b Institute of Organic Chemistry, RWTH Aachen University, Landoltweg 1, D-52056 Aachen, Germany. elmar.weinhold@oc.rwth-aachen.de

^c Chemistry and Material Science Faculty, Shandong Agricultural University, Tai'an, P.R. China.

^d Department of Biochemistry and Cellular and Molecular Biology, University of Tennessee, Knoxville, USA. hguo1@utk.edu

^e UT/ORNL Center for Molecular Biophysics, Oak Ridge National Laboratory, USA.

^f Department of Physics, Chemistry and Pharmacy, University of Southern Denmark, Campusvej 55, 5230 Odense, Denmark. mecinovic@sdu.dk

Biomedicinally important histone lysine methyltransferases (KMTs) catalyze the transfer of methyl group from *S*-adenosylmethionine (AdoMet) cosubstrate to lysine residues in histones and other proteins. Here we report experimental and computational investigations on human KMT-catalyzed ethylation of histone peptides using *S*-adenosylethionine (AdoEth) and *Se*-adenosylselenoethionine (AdoSeEth) cosubstrates. MALDI-TOF MS experiments revealed that, unlike monomethyltransferases SETD7 and SETD8, methyltransferases G9a and GLP do have a capacity to ethylate lysine residues in histone peptides, and that cosubstrates follow the efficiency trend AdoMet > AdoSeEth > AdoEth. G9a and GLP can also catalyze AdoSeEth-mediated ethylation of ornithine, and produce histone peptides bearing lysine residues with different alkyl groups, such as H3K9me₂et and H3K9me₂et. Molecular dynamics and free energy simulations based on quantum mechanics/molecular mechanics (QM/MM) potential supported the experimental findings by providing an insight into the geometry and energetics of the enzymatic methyl/ethyl transfer process.

Introduction

Histone proteins are subject of diverse posttranslational modifications (PTMs), including methylation, acetylation, crotonylation, phosphorylation, citrullination and ubiquitination, which regulate the activity of human genes via epigenetic mechanisms.^[1] Methylation of lysine residues in unstructured histone tails is associated with both gene activation and repression, depending on the histone, methylation state and methylation site.^[2] Histone lysine methylation is catalyzed by *S*-adenosylmethionine (AdoMet or SAM)-dependent histone lysine methyltransferases (KMTs) that install one (Kme), two (Kme₂) or three (Kme₃) methyl groups on the *N*^ε-amino group of lysine (Figure 1A).^[3] Histone lysine methylation is removed by flavin-dependent lysine specific demethylases and Fe(II)/2-oxoglutarate (2OG)-dependent

histone demethylases (KDMs),^[4] and recognized by a large number of *N*^ε-methyllysine binding epigenetic reader proteins,^[5] which collectively spread the epigenetic landscape of posttranslational modifications.

Histone KMTs contain the conserved SET (Su(var)3-9, enhancer-of-zeste, trithorax) domain responsible for the enzymatic activity; DOT1L is the only member of the histone KMT family, known to date, that does not contain the SET domain.^[6] Structural analyses revealed that KMTs possess distinct binding pockets for AdoMet cosubstrate and histone substrate (Figure 1B).^[6c] In the ternary complex, the nucleophilic *N*^ε-amino group of lysine is well aligned with the electrophilic methyl group of AdoMet for an efficient S_N2 reaction that takes place in a narrow hydrophobic channel typically comprised of side chains of Tyr and Phe residues (Figure 1B). The presence of Tyr and Phe in the active sites of KMTs appears to define the methylation state of the product; Tyr to Phe substitutions result in the formation of higher methylation states of lysine.^[6c] The target lysine needs to be deprotonated for nucleophilic attack and an active site Tyr may also be responsible for deprotonation of the protonated lysine, although a water channel has also been suggested to play a role as a general base.^[7]

Despite recent success in structural, mechanistic and inhibition studies on histone lysine methyltransferases,^[3a, 8] the biocatalytic potential of KMTs remains to be established.^[9] Enzymatic assays revealed that human KMTs exhibit a high degree of specificity for methylation of lysine analogues that differ in stereochemistry, side chain length and main chain.^[10] On the other hand, KMTs appear to have a limited ability to catalyze other alkylations of histones, including transfer of allyl, propargyl and larger alkyl groups from AdoMet analogues bearing methyl group replacements.^[11] Here we report investigations on KMT-catalyzed ethylation of histone peptides employing *S*-adenosylethionine (AdoEth) and *Se*-adenosylselenoethionine (AdoSeEth) cosubstrates (Figure 1C).

Results and Discussion

Analogues of AdoMet with methyl group replacements, including AdoEth and AdoSeEth, can be enzymatically synthesized from L-methionine derivatives and adenosine triphosphate (ATP) using methionine adenosyltransferases (MAT) from different organisms.^[12] A pronounced product inhibition of the MAT enzymes, however, often limits the synthesis to small amounts with isolated enzymes.^[13] Larger amounts of cosubstrate analogues can be obtained by chemical synthesis (Scheme 1). *S*-Adenosylhomocysteine (AdoHcy)^[14] or *Se*-adenosylselenohomocysteine (AdoSeHcy)^[11a, b, h] are typically reacted with alkylating agents under slightly acidic conditions using a mixture of formic and acetic acid. These conditions guide regioselective alkylation of the sulfur or selenium atom because all other nucleophilic positions are transiently protected by protonation. During the synthesis of the more reactive selenonium analogue AdoSeEth, we noticed many byproducts. Fortunately, formation of these byproducts could be efficiently suppressed by adding water to the mixture of formic and acetic acid (Scheme 1).

Generally, alkylations of AdoHcy under acidic conditions lead to both diastereoisomers (epimers) at sulfur in almost equal amounts.^[14a] In the case of AdoEth the two epimers formed in a 45:55 ratio (*S*-epimer to *R*-epimer). Both epimers were separated by reverse-phase HPLC and only the *S*-epimer (corresponding to the biologically active *S*-epimer of AdoMet) was used in this study. However, in the case of AdoSeEth separation of both epimers by reverse-phase HPLC was not possible and AdoSeEth was used as an epimeric mixture.

We then carried out comparative enzymatic assays for KMT-catalyzed methylation (using AdoMet) and ethylation (using AdoEth and AdoSeEth) of synthetic histone peptides using MALDI-TOF MS, as recently described;^[10a, b] histone H3₁₋₁₅ was used for studies with SETD7 (also known as KMT7), G9a (also known as KMT1C and EHMT2) and GLP (also

known as KMT1D and EHMT1), and histone H4₁₃₋₂₇ was used for studies with SETD8 (also known as KMT5A). MALDI-TOF MS data confirmed that human KMTs catalyze nearly quantitative methylation of histone peptides in the presence of AdoMet: H3K4me, H4K20me, H3K9me₂ and H3K9me₃ were formed in the presence of SETD7, SETD8, G9a and GLP, respectively (Figure 2, top panels). Unlike monomethylation, SETD7 and SETD8 did not catalyze ethylation of H3K4 and H4K20 with AdoEth or more reactive AdoSeEth within detection limits (Figure 2A–B and Figures S1–S2). Human G9a and GLP, however, were able to catalyze ethylation of H3K9, with AdoSeEth being a superior ethylation agent over AdoEth (Figure 2C–D). Although G9a and GLP catalyze di- and trimethylation of H3K9, both enzymes were only able to catalyze monoethylation of H3K9, with no di- and triethylation detected (Figure 2C–D). A longer incubation time (5 h) with AdoEth led to a formation of 56% and 30% of H3K9et by G9a and GLP, respectively (Figure S3), whereas almost complete (87%) formation of H3K9et was observed after 5 h when AdoSeEth and G9a were used (Figure S4). Control experiments in the absence of AdoEth/AdoSeEth or G9a/GLP verified that ethylation reactions are KMT-catalyzed (Figures S5–S6).

Having shown that G9a and GLP have an ability to catalyze monoethylation of H3K9, we next investigated potential enzymatic ethylation of biologically important methylated histones H3K9me and H3K9me₂. G9a and GLP both poorly catalyzed ethylation of H3K9me (traces detected) in the presence of AdoEth, even upon prolonged incubation (Figure 3A–B and Figure S7). Both enzymes, however, produced detectable amounts of H3K9meet in the presence of more reactive AdoSeEth (Figure 3A–B and Figure S8). G9a, in particular, produced significant amounts (55% after 3 hours and 75% after 5 hours) of H3K9meet (Figure 3A–B and Figure S8). An observation that G9a and GLP have a capacity to catalyze the formation of H3K9meet is interesting, because functionally related histone lysine demethylases PHF8, FBXL11 and JMJD2E were found to catalyze removal of methyl and ethyl groups in

H3K9meet, thus producing an unmodified H3K9 (a substrate for G9a and GLP).^[15] Moreover, our MALDI-TOF MS assays revealed that G9a and GLP also catalyze ethylation of H3K9me2, producing approximately 25% of bulky H3K9me2et in the presence of AdoSeEth; only traces of H3K9me2et were observed when AdoEth was used as a cosubstrate (Figure 3C–D). Prolonged incubation (5 h at 37 °C) led to increased amounts (41%) of H3K9me2et in the presence of G9a and AdoSeEth, whereas AdoEth did not enhance formation of the trialkylated product (Figures S9–S10). Control experiments in the absence of G9a/GLP or AdoSeEth showed no ethylation of H3K9me and H3K9me2, implying that the reactions are KMT-catalyzed and that the ethyl group in the H3K9meet and H3K9me2et product derives from AdoSeEth cosubstrate (Figures S11–S12).

Despite the fact that AdoSeEth and AdoEth can both act as ethylating agents in enzymatic assays, they do exhibit significant differences with respect to reactivity. In analogy with AdoMet/AdoSeMet,^[11b, 16] AdoSeEth appears to be more reactive (i.e. better alkylation agent) than AdoEth. One notable difference between the two molecules is the bond length; the C-Se bond is longer (2.0 Å) than the C-S bond (1.8 Å) making the selenonium analogues more reactive. Due to longer C-Se bond and higher reactivity of AdoSeEth, we hypothesized that KMTs might have an ability to catalyze ethylation of ornithine, the one-methylene shorter lysine analogue. In line with our earlier observation,^[10b] we observed that G9a and GLP do not methylate H3Orn9, in the presence of AdoMet, within limits of detection (Figure 4, top panels). Similarly, no G9a/GLP-catalyzed ethylation of H3Orn9 was observed when AdoEth was used as a cosubstrate, even upon longer incubation (Figure 4, middle panels and Figure S13). Interestingly, our MALDI-TOF MS data showed that G9a and GLP predominantly catalyze monoethylation of H3Orn9 in the presence of AdoSeEth; as in the case of H3K9, no diethylation of H3Orn9 was observed (Figure 4, bottom panels). A longer incubation time (5 h at 37 °C) led to nearly complete (96%) and significant (78%) formation of H3Orn9et with G9a

and GLP, respectively (Figure S14). Controls without G9a/GLP or AdoSeEth again verified that ethylation reactions are KMT-catalyzed (Figure S15). We also examined potential G9a/GLP-catalyzed ethylation of H3Dab9 peptide that possesses the two-methylene shorter lysine analogue 2,4-diaminobutyric acid (Dab). However, we did not detect any ethylated products in the presence of AdoEth and AdoSeEth within detection limits (Figure S16–S17). Finally, we also carried out enzymatic assays with ornithine-containing histone peptides H3Orn4 and H4Orn20 with human SETD7 and SETD8. Unlike monomethylation of H3K4, human SETD7 did not catalyze monoethylation of H3Orn4 in the presence of AdoEth or AdoSeEth (Figure S18). Similarly, despite high degree of monomethylation of H4K20 by SETD8, the enzyme did not yield any H4Orn20et in the presence of AdoEth or AdoSeEth (Figure S19). This is in line with the absence of ethylation of H3K4 and H4K20 with AdoEth and AdoSeEth as discussed above (Figure 2A–B). Collectively, our enzymatic assays revealed that human G9a and GLP possess a biocatalytic potential for ethylation of the shorter ornithine residue.

To obtain a better understanding of the G9a- and GLP-catalyzed ethylation of H3K9 and H3Orn9 we performed kinetic experiments with different cosubstrate concentrations (Table 1, Figures S20 and S21). However, G9a- and GLP-catalyzed ethylation reactions with AdoEth were so slow that no multiple turnovers were obtained within the extended reaction time. Comparing the single turnover rate constant $k = 0.046 \text{ min}^{-1}$ of G9a in the presence of saturating AdoEth concentrations with $k_{cat} = 11.6 \text{ min}^{-1}$ obtained with AdoMet under multiple turnover (steady state) conditions indicates that alkylation with the natural cosubstrate is at least 200-fold faster. Very similar kinetics results were obtained with GLP. Such a two to three orders reduced activity with AdoEth compared with AdoMet was also observed with DNA MTases,^[14a] and can be attributed to an increased steric strain in the S_N2 -type transition state for ethyl compared to methyl transfer. With AdoSeEth the ethylation rate under saturating

cosubstrate concentrations is about 3-4 times faster for both G9a and GLP, however, the true rate enhancement upon going from AdoEth to AdoSeEth might be even larger because AdoSeEth was employed as an epimeric mixture at the selenonium center (separation of the epimers by reverse-phase HPLC was not possible, *vide supra*) and the epimer with the non-natural *R*-configuration might act as an inhibitor as observed for (*R*)-AdoMet.^[17] Furthermore, the ethylation rate of the side chain shortened substrate H3Orn9 with AdoSeEth is only reduced by 50% compared to that of the H3K9 substrate with the natural target amino acid.

We then carried out quantum mechanics/molecular mechanics (QM/MM) investigations to rationalize experimental observations on the KMT-catalyzed ethylation of lysine residues. Because the parameters for selenium are still not available in the semi-empirical QM DFTB3 method, only the simulations for the ethylation reactions involving AdoEth were performed in this work. The average active-site structures of the reactant complexes for the first methylation and first ethylation in SETD8 are compared in Figure 5 (active-site structures near the transition state are shown in Figure S22); the distribution map of $r(C_M-N^{\epsilon})/r(C_{M1}-N^{\epsilon})$ and θ is also given in each case. As can be seen from these figures, the alignment of the electron lone pair on N^{ϵ} of the target lysine with the transferable ethyl groups (Figure 5B) is significantly worse than that with the transferable methyl group (Figure 5A). For instance, the average distance between N^{ϵ} of lysine and C_{M1} is 4.3 Å in Figure 5B, while it is 3.4 Å between N^{ϵ} and C_M in Figure 5A. The poor alignment between the ethyl donor and acceptor positions is also reflected in the distribution map. The free-energy profiles for the first methylation and first ethylation reactions in SETD8 are given in Figure 5C. Consistent with the structures of the reactant complexes, the free energy barrier for the ethylation reaction (22.7 kcal mol⁻¹) is considerably higher than that of methylation (19.4 kcal mol⁻¹). The simulation results indicate that although SETD8 can catalyze the monomethylation of H4K20, it may not be able to do so for the ethylation reaction, consistent with the experimental data (Figure 2B).

The average active-site structures of the reactant complexes for the first and second ethyl transfers in GLP are given in Figure 6A–B, respectively (active-site structures near the transition state are shown in Figure S23). Figure 6A–B shows that the alignment of the electron lone pair on N^ϵ of the target lysine with the transferable ethyl group in GLP is significantly better for the first ethyl transfer (with a relatively shorter average $C_{M1}\cdots N^\epsilon$ distance of 4.08 Å and more populations of the near attack conformations) compared to that for the second ethyl transfer (with the average $C_{M1}\cdots N^\epsilon$ distance of 4.56 Å). This is in contrast with the cases for the first and second methyl transfers in GLP where the target lysine and methyl lysine can be well aligned, respectively, with the transferable methyl group (Figure S24). The free-energy profiles for the first and second ethylation reactions in Figure 6C demonstrate that while GLP may catalyze the first ethylation reaction (with the free energy barrier of 18.3 kcal mol⁻¹ that is higher than 17.0 kcal mol⁻¹ for the first methylation reaction^{10b}), it is unlikely to be able to catalyze the second ethylation reaction using AdoEth due to the fact that the free energy barrier becomes significantly higher (23.4 kcal mol⁻¹). These results are in agreement with the observed monoethylated product H3K9et and the lack of the diethylated product H3K9et2 in MALDI-TOF assays, as shown Figure 2D.

Conclusions

We have demonstrated that human histone lysine methyltransferases G9a and GLP have a capacity to catalyze monoethylation of H3K9 in the presence of AdoEth and AdoSeEth cosubstrates. Enzymatic assays revealed that AdoSeEth is a superior ethylating agent over AdoEth, but comparatively poorer cosubstrate than the natural AdoMet. An ability of AdoEth to act as an ethylating agent for histone ethylation in cells^[18] and in vitro as investigated here might have some biological relevance, because its precursor ethionine is a toxic compound that can be converted to AdoEth by eukaryotic MAT enzymes.^[19] Computational work unveiled

that the molecular origin for the more efficient enzymatic methylation over ethylation of lysine residues in histones lies in more optimal alignment of smaller methyl group of AdoMet when compared to larger ethyl group of AdoEth. Our examinations also revealed that G9a and GLP catalyze ethylation of histone H3 bearing biologically relevant methylated lysine residues (H3K9me and H3K9me₂), and histone H3 peptide possessing one-methylene shorter ornithine (H3Orn9) in the presence of AdoSeEth. Collectively, this work highlights a biocatalytic potential of selected human KMTs and expands the substrate scope for KMT-catalyzed alkylation of histones. It is envisioned that this work, along with recent investigations of KMT-catalyzed alkylation of proteins, including histones, will advance our basic understanding of KMT catalysis.

Experimental Section

Materials

Se-Adenosylselenohomocysteine (AdoSeHcy)^[11h] and 5'-[(*S*)-[(3*S*)-3-amino-3-carboxypropyl]ethylsulfonio]-5'-deoxyadenosine (*S*-adenosylethionine, AdoEth)^[14a] were prepared as described before.

Synthesis

Ethyl triflate. Ethyl triflate was obtained by a slightly modified literature procedure.^[20] Polyvinylpyridine (1,21 g) was suspended in dichloromethane (27 mL) with ice cooling. Trifluoromethanesulfonic anhydride (1,55 g, 5,49 mmol) was added dropwise and then ethanol (240 mg, 5,21 mmol) in dichloromethane (2.75 mL) was added within 2 min. The ice bath was removed and stirring continued for another 10 min at room temperature. The solid was filtered off, and the solvent was removed under reduced pressure to yield a slightly brown liquid (754

mg, 4.23 mmol, 81 %). ^1H NMR (300 MHz, CDCl_3): δ = 1.45 (t, J = 7.0 Hz, 3H; H2), 4.55 ppm (q, J = 7.2 Hz, 2H; H1).

5'-[(Se)[(3S)-3-Amino-3-carboxypropyl]ethylselenonio]-5'-deoxyadenosine (Se-adenosylselenoethionine, AdoSeEth). Se-Adenosylselenohomocysteine (AdoSeHcy) (20 mg, 46,4 μmol) was dissolved in a 1:1 mixture of formic and acetic acid (1,95 mL) and water (0.98 mL) was added. The solution was supplemented with ethyl triflate (1,49 g, 8.37 mmol) and the reaction mixture was stirred at room temperature for 1 h. The reaction was quenched by adding water (6 mL) and the aqueous phase was extracted three times with diethylether (7.5 mL each time). Purification of the product in the aqueous phase was performed by preparative reverse-phase HPLC (Prontosil-ODS 5 μm , 120 \AA , 250 \times 20 mm, Bischoff, Leonberg, Germany). Compounds were eluted with methanol (linear gradients from 0 % to 7.8 % in 15 min and to 78 % in 5 min) in aqueous trifluoroacetic acid (0.01%) and a flow of 10 mL/min. Compounds were detected at 260 and 272 nm. The two epimers (at selenium) both eluted with a retention time around 9.9 min and could not be separated. Product containing fractions were collected and the solvents were removed by lyophilization. The remaining solid was dissolved in water (1.5 mL) and the yield 8.32 mg (18.1 μmol , 39 %) was determined by UV spectroscopy (ϵ^{260} = 15,400 $\text{L mol}^{-1} \text{cm}^{-1}$). ESI-MS m/z (relative intensity): 461.1 (100) $[\text{M}]^+$, 360.2 (8) $[\text{5'-ethylseleno-5'-deoxyadenosine} + \text{H}]^+$, 250.4 (17) $[\text{M} - \text{ethionine}]^+$, 136,2 (9) $[\text{adenine} + \text{H}]^+$.

Expression and purification of KMTs

The four wild-type human histone lysine methyltransferases (SETD7, SETD8, G9a and GLP) were expressed and purified as earlier described.^[10a, b] The wild-type sequences of human methyltransferases were as follows: SETD8 (aa 186–352), SETD7 (aa 1-366), G9a (aa 913-1193), and GLP (aa 951-1235). Briefly, overnight cultures of *E.coli* Rosetta BL21 (DE3)pLysS harboring expression plasmids were made by growing at 37 °C with shaking for

18 h in LB media supplemented with kanamycin and chloramphenicol. The expression was induced at OD₆₀₀ (approximately 0.5-0.6) by adding isopropyl-beta-D-thiogalactopyranoside (IPTG) and shaking continued at 16 °C for 20 h. Cells were harvested by centrifugation at 4 °C for 15 min, and the cell pellets were resuspended in lysis buffer. Cells were lysed using a Soniprep 150 sonicator for 20 s (8 times) with 90 s intervals keeping the cells chilled in an ice water bath at all times. After centrifugation the supernatant was loaded onto the Ni-charged His-tag binding resin that was equilibrated with column buffer. Resins were washed thoroughly with column buffer, followed by washing buffer and protein was eluted with elution buffer using a linear gradient concentration of imidazole. The protein was then applied to size exclusion chromatography (SEC) using a Superdex 75 column (GE Healthcare). Purified proteins were concentrated employing Amicon Ultra Centrifugal Filter Units (Millipore) with suitable molecular weight cutoffs (10 kD). Protein concentration was determined employing the Nanodrop DeNovix DS-11 spectrophotometer and the purity was monitored by SDS-PAGE on a 4–15 % gradient polyacrylamide gel (Bio-Rad). Enzymes were aliquoted and stored at -80 °C for future use.

Enzymatic assays by MALDI-TOF MS

The reactions were carried out in total volume of 25 μL in an Eppendorf vial using thermomixer. A typical enzymatic assay included histone peptides (40 μM), cosubstrate AdoMet (200 μM with SETD8 and SETD7; 500 μM with GLP and G9a), AdoEth (1 mM) or AdoSeEth (1 mM), and KMT enzyme (2 μM) in a reaction buffer of 50 mM Tris-HCl at pH 8.0. The reactions were incubated at 37 °C and aliquots were taken from the reaction vial at different time points (1 h, 3 h, and 5 h) to measure the conversion of histone peptide substrates to alkylated products. The reaction was stopped by mixing 3 μL of the reaction mixture with 3 μL of MeOH. An aliquot (3 μL) of these samples were directly mixed onto the MALDI target

plate using 3 μL of α -cyano-4-hydroxycinnamic acid (CHCA) matrix (5 mg mL^{-1} in 50% (v/v) acetonitrile/water) and air-dried. Peptide substrates masses were measured using the positive-ion reflector mode. Full mass scans were acquired in the m/z range of 500-4000. Each mass spectrum was generated from the data deriving from 3–5 single laser shots in 200-shot steps from different positions of the sample spot. All the spectra were manually acquired by using a Microflex mass spectrometer and the FlexControl software and the data were annotated employing FlexAnalysis software (Bruker Daltonics, Germany). The methylation and ethylation species were observed as follows; monomethylation (+ 14 Da), dimethylation (+ 28 Da), trimethylation (+ 42 Da), and monoethylation (+ 28 Da). All methylation and ethylation experiments were performed in replicate.

Enzyme kinetics analyses

The kinetics parameters (k , K) were determined by incubating G9a or GLP (2 μM for AdoEth; 1 μM for AdoSeEth) and histone peptide (25 μM 15-mer H3K9 or H3Orn9) in 50 mM TRIS pH 8.0 in the presence of various concentrations of AdoEth (0-250 μM) or AdoSeEth (0-250 μM). The reactions (final volume = 20 μL) were incubated at 37 $^{\circ}\text{C}$ (700 rpm) for 15 min, after which they were quenched by adding an equal volume of MeOH. For analysis by MALDI-TOF MS, 2 μL of the quenched reaction mixture was mixed with 6 μL of saturated CHCA solution (1:1 v/v in MeCN/ddH₂O + 0.1% TFA). From this, 1 μL was spotted onto the MALDI plate for crystallization. The enzymatic activity was determined using the peak areas (including all isotopes) for each alkylation state. To obtain the kinetics parameters, the data were plotted and fitted to the non-linear regression enzyme kinetics k_{cat} function using GraphPad PRISM software (Figure S20). The kinetics profiles for AdoMet (Figure S21) were obtained by incubation of G9a or GLP (50 nM), histone peptide (10 μM 15-mer H3K9) and AdoMet (0-15

μM) in 50 mM TRIS pH 8.0 buffer for 3 min at 37 °C (700 rpm). MALDI-TOF MS measurements and data analysis was performed as for AdoEth and AdoSeEth described above.

QM/MM studies

QM/MM free energy (potential of mean force) and MD simulations were performed to study the active-site dynamics of SETD8 and GLP and to calculate the free energy profiles of the ethyl transfers from AdoEth to lysine and its ethylated form using the CHARMM program.^[21] The $-\text{CH}_2-\text{CH}_2-\text{S}^+(\text{Et})-\text{CH}_2-$ part of AdoEth and lysine/ethyl lysine chain were treated by QM and the rest of the system by MM. The link-atom approach^[22] was applied to separate the QM and MM regions. A modified TIP3P water model^[23] was employed for the solvent, and the stochastic boundary molecular dynamics method^[24] was used for the QM/MM simulations. The system was separated into a reaction zone and a reservoir region, and the reaction zone was further divided into a reaction region and a buffer region. The reaction region was a sphere with radius r of 20 Å, and the buffer region extended over $20 \text{ \AA} \leq r \leq 22 \text{ \AA}$. The reference center for partitioning the system was chosen to be the N^ϵ atom of the target lysine. The resulting systems contained around 5800 atoms for GLP (or 5400 atoms for SETD8), including about 700-900 water molecules. The DFTB3 method^[24-25] implemented in CHARMM was used for the QM atoms. The semi-empirical approach adopted here has been used previously on a number of systems, and the results seem to be quite reasonable.^[26] The all-hydrogen CHARMM potential function (PARAM27)^[27] was used for the MM atoms.

The initial coordinates for the reactant complexes of the methylation/ethylation were based on the crystallographic complexes (PDB codes: 2BQZ and 3HNA for SETD8 and GLP, respectively); a methyl/ethyl group was manually added to SAH to change it to AdoMet/AdoEth and the methyl group on the methyl lysine was manually deleted to generate the target lysine. The initial structures for the entire stochastic boundary systems were

optimized using the steepest descent (SD) and adopted-basis Newton-Raphson (ABNR) methods. The systems were gradually heated from 50.0 to 298.15 K in 50 ps. A 1 fs time step was used for integration of the equation of motion, and the coordinates were saved every 50 fs for analyses. 1.5 ns QM/MM MD simulations were carried out for each of the reactant complexes, and the similar approaches have been used previously.^[7a, 28]

The umbrella sampling method^[29] implemented in the CHARMM program along with the Weighted Histogram Analysis Method (WHAM)^[30] was applied to determine the changes of the free energy (potential of mean force) as a function of the reaction coordinate for the methyl/ethyl transfer from AdoMet/AdoEth to the target lysine/ethyl lysine in SETD8 and in GLP. The reaction coordinate was defined as a linear combination of $r(\text{C}_M\text{-S}_\delta)$ and $r(\text{C}_M\text{-N}^\epsilon)$ for methylation [$R = r(\text{C}_M\text{-S}_\delta) - r(\text{C}_M\text{-N}^\epsilon)$] or $r(\text{C}_{M1}\text{-N}^\epsilon)$ and $r(\text{C}_{M1}\text{-S}_\delta)$ for ethylation [$R = r(\text{C}_{M1}\text{-S}_\delta) - r(\text{C}_{M1}\text{-N}^\epsilon)$] (see Figure 5 for the atom designation). 30 windows were used and for each window 50 ps production runs were performed after 50 ps equilibration. The force constants of the harmonic biasing potentials used in the PMF simulations were 50–400 kcal mol⁻¹ Å⁻².

Keywords

Epigenetics, ethylation, histone, lysine, methyltransferase.

Acknowledgements

We thank the Netherlands Organization for Scientific Research (NWO, NCI-TA grant 731.015.202) for financial support. This work was supported in part by the Natural Science Foundation of Shandong Province in China (No. ZR2017MB048 to P.Q.) and by the Excellence Initiative of the German Federal and State Governments. We thank Professor Martin Karplus for a gift of the CHARMM program.

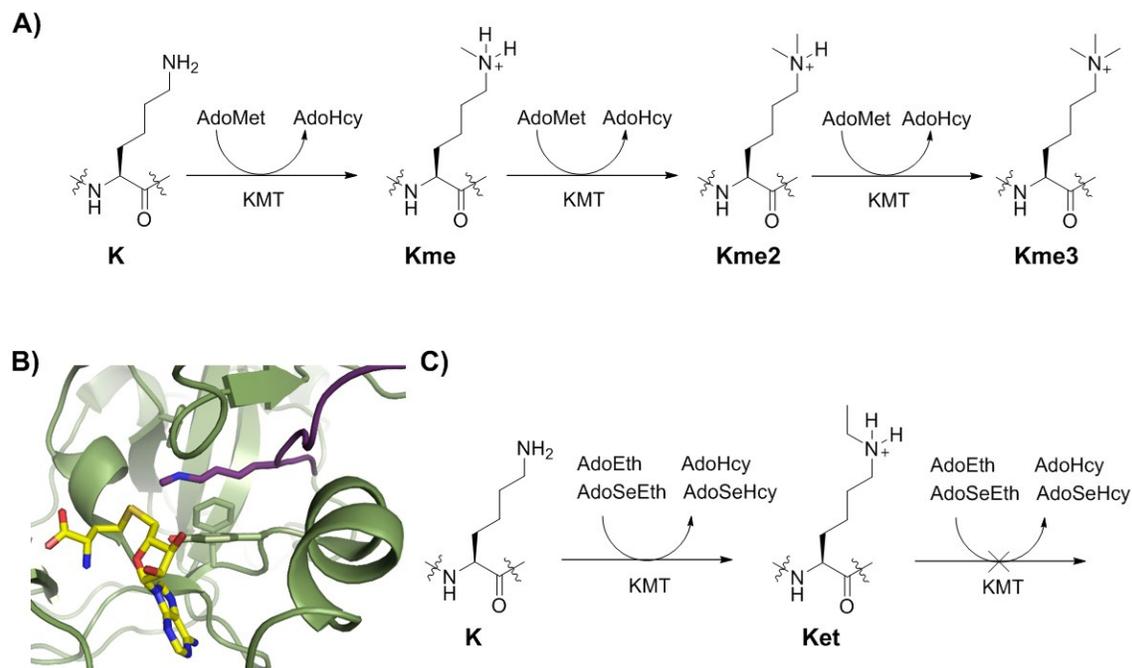
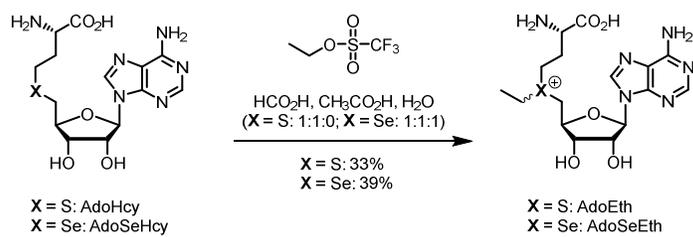


Figure 1. Histone lysine methyltransferase (KMT)-catalyzed alkylation of histones. A) AdoMet-mediated lysine methylation leading to Kme, Kme2 and Kme3. B) View from a crystal structure of GLP (green) in complex with H3K9me (violet) and AdoHcy (yellow) (PDB ID: 3HNA). C) AdoEth- or AdoSeEth-mediated lysine ethylation leading to Ket.



Scheme 1. Chemical synthesis of AdoEth^[14a] and AdoSeEth.

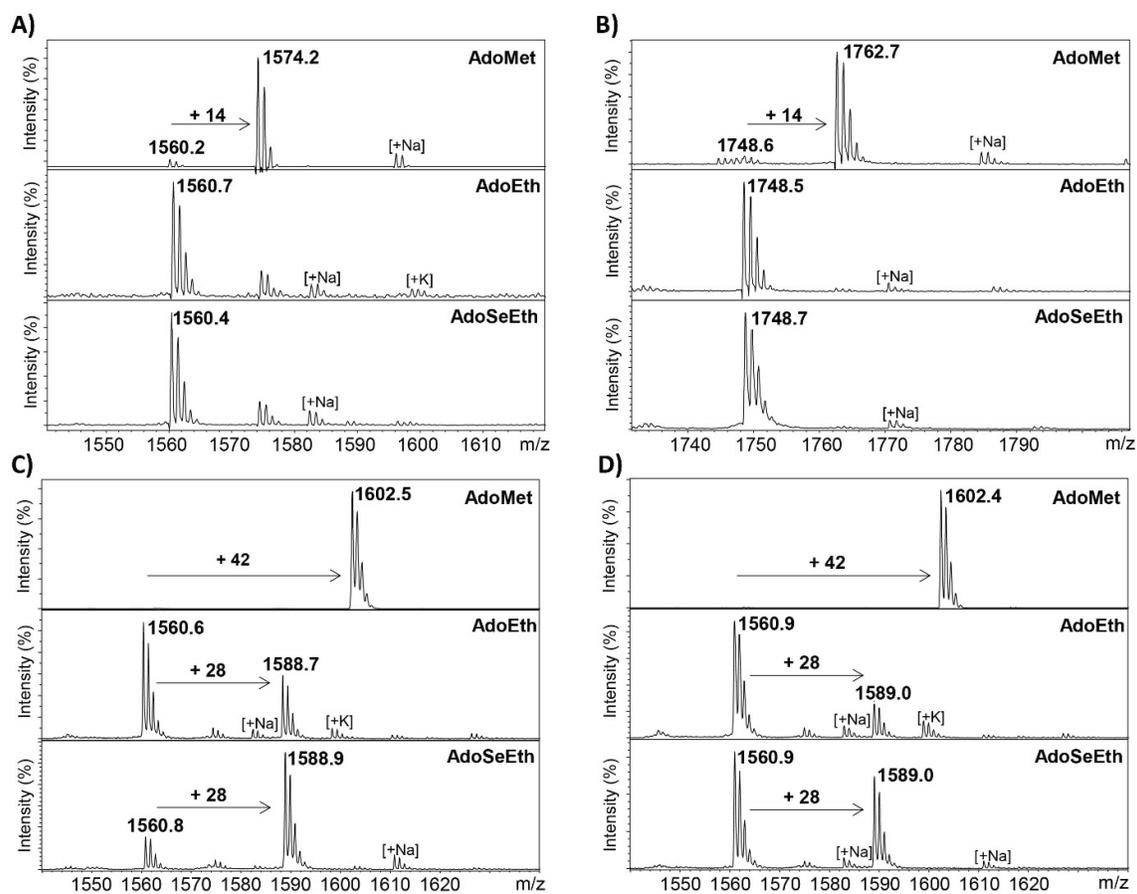


Figure 2. MALDI-TOF MS assays showing methylation (top panel) and ethylation reactions (middle and bottom panels) of A) H3K4 by SETD7; B) H4K20 by SETD8; C) H3K9 by G9a; D) H3K9 by GLP.

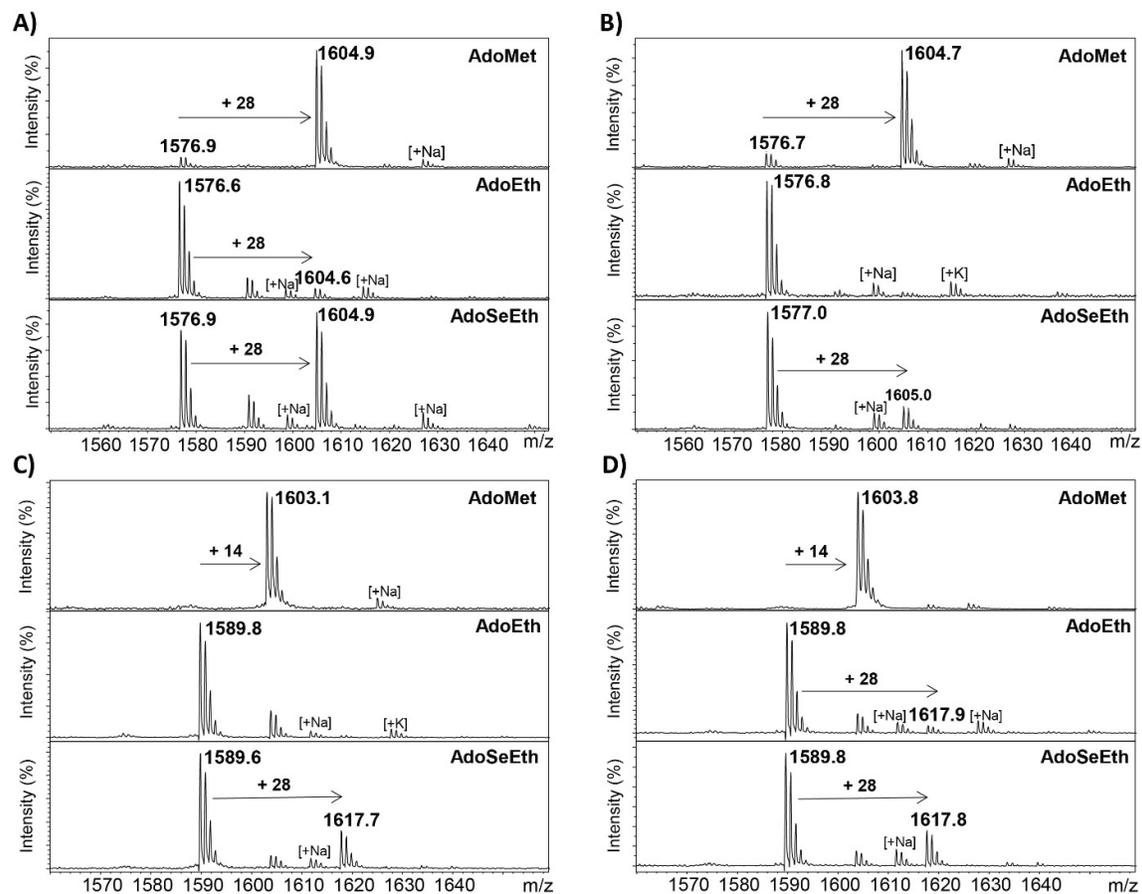


Figure 3. MALDI-TOF MS assays showing methylation (top panel) and ethylation reactions (middle and bottom panels) of A) H3K9me by G9a; B) H3K9me by GLP; C) H3K9me2 by G9a; D) H3K9me2 by GLP.

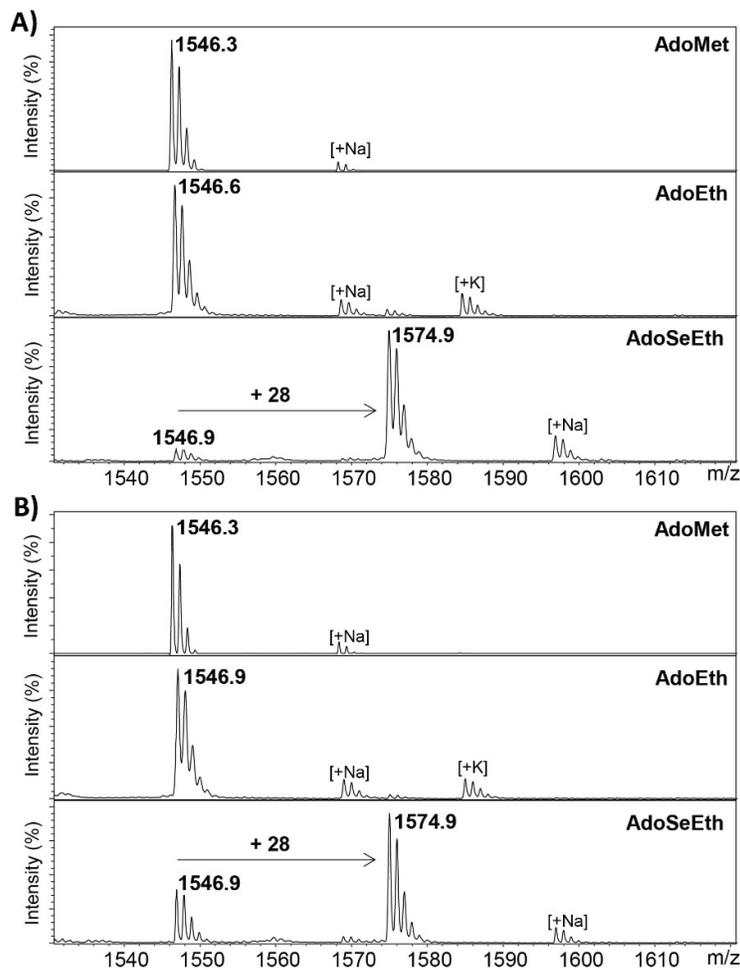


Figure 4. MALDI-TOF MS assays showing methylation (top panel) and ethylation reactions (middle and bottom panels) of A) H3Orn9 by G9a; B) H3Orn9 by GLP.

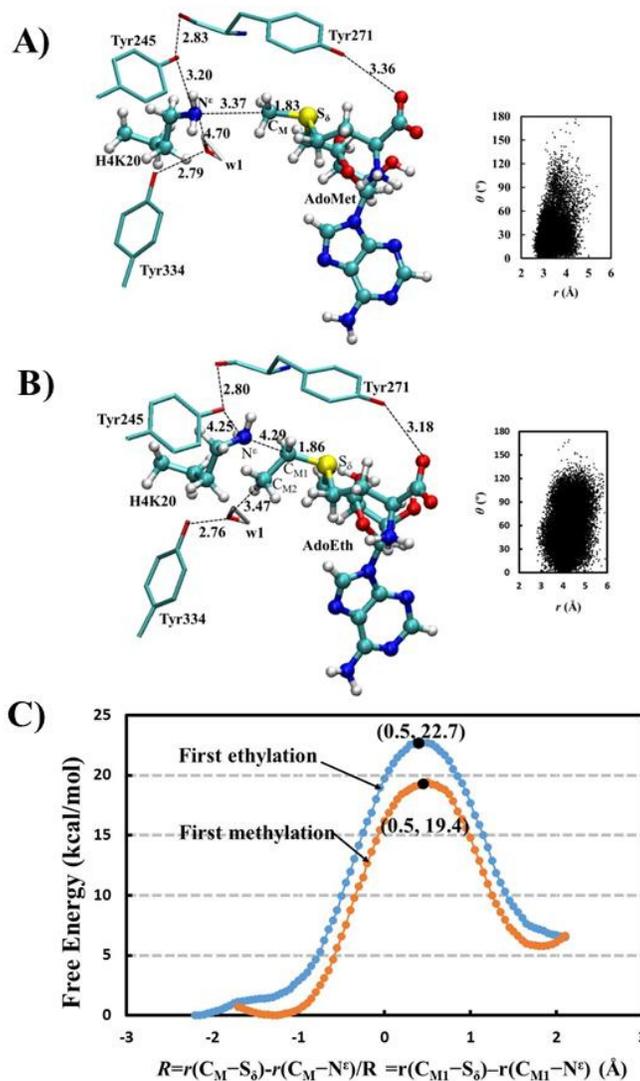


Figure 5. A) Representative active-site structure of the reactant complex of SETD8 for the first methylation containing AdoMet and lysine (H4K20) obtained from the QM/MM MD simulations. The distribution map on the right shows the alignment of N^cH₂ and the transferable methyl group in the reactant complex in terms of the distance (r) between N^c and C_M and the angle (θ) between the direction of electron lone pair on N^c and the C_M-S bond. SET8 is shown in sticks, and AdoMet and lysine are in balls and sticks. Some average distances from the simulations are also given (in angstroms). B) Representative active-site structure of the reactant complex of SETD8 for the first ethylation containing AdoEth and lysine (H4K20) along with the $r(C_{M1}\cdots N^c)$ and θ distribution map obtained from the QM/MM MD simulations. C) Free energy (potential of mean force) profiles for the methylation/ethylation reactions in SETD8 as a function of the reaction coordinate [$R = r(C_M-S_\delta) - r(C_M-N^c) / R = r(C_{M1}-S_\delta) - r(C_{M1}-N^c)$]. The free energy profile for the first methylation: orange line with a free energy barrier of 19.4 kcal mol⁻¹ and the location of the transition state is at around 0.5. The free energy profile for the first ethylation: blue line with a free energy barrier of 22.7 kcal mol⁻¹.

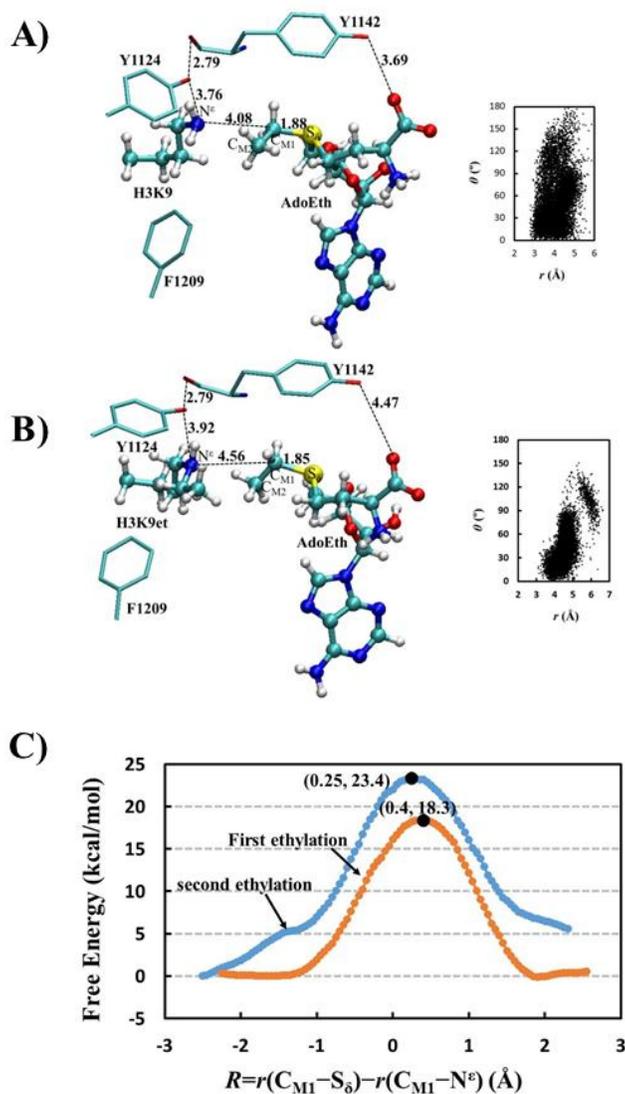


Figure 6. A) Average structure obtained from molecular dynamics simulations for the first ethyl transfer from AdoEth to H3K9 in GLP along with the distribution map during the MD simulations. Some average distances of the potential hydrogen bonds between AdoEth/substrate and nearby residues are shown with units of angstroms. B) Average structure obtained from molecular dynamics simulations for the second ethyl transfer from AdoEth to H3K9et in GLP. C) Free energy changes for mono- and diethylation of H3K9 in GLP. Monoethylation: orange line with a free energy barrier of 18.3 kcal mol⁻¹. Diethylation: blue line with a free energy barrier of 23.4 kcal mol⁻¹.

Table 1. Rate constants k for G9a- and GLP-catalyzed methylation and ethylation of H3K9 and H3Orn9 histone peptides with saturating AdoMet, AdoEth and AdoSeEth concentrations. Reactions with AdoMet were obtained under steady state conditions, while reactions with AdoEth and AdoSeEth were too slow to reach steady state.

		AdoMet ¹	AdoEth ²	AdoSeEth ³	
		H3K9	H3K9	H3K9	H3Orn9
G9a	k (min ⁻¹)	11.6 ± 0.61	0.046 ± 0.005	0.21 ± 0.01	0.12 ± 0.01
GLP	k (min ⁻¹)	9.90 ± 0.44	0.058 ± 0.003	0.16 ± 0.01	0.10 ± 0.003

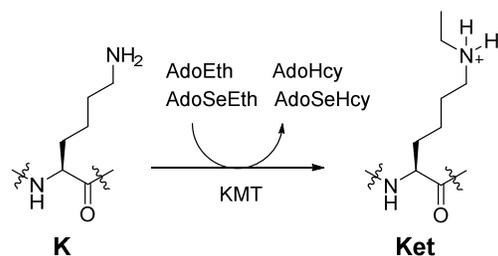
¹No methylation was observed with H3Orn9 and AdoMet. ²No ethylation was observed with H3Orn9 and AdoEth. ³AdoSeEth was employed as epimeric mixture at the selenium center.

References

- [1] a) A. J. Bannister and T. Kouzarides, *Cell Res.* **2011**, *21*, 381-395; b) T. Kouzarides, *Cell* **2007**, *128*, 693-705; c) B. D. Strahl and C. D. Allis, *Nature* **2000**, *403*, 41-45.
- [2] J. C. Black, C. Van Rechem and J. R. Whetstine, *Mol. Cell* **2012**, *48*, 491-507.
- [3] a) M. Luo, *Chem. Rev.* **2018**, *118*, 6656-6705; b) C. Martin and Y. Zhang, *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 838-849.
- [4] S. M. Kooistra and K. Helin, *Nat. Rev. Mol. Cell Biol.* **2012**, *13*, 297.
- [5] S. D. Taverna, H. Li, A. J. Ruthenburg, C. D. Allis and D. J. Patel, *Nat. Struct. Mol. Biol.* **2007**, *14*, 1025-1040.
- [6] a) J.-F. Couture, L. M. A. Dirk, J. S. Brunzelle, R. L. Houtz and R. C. Trievel, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 20659-20664; b) P. A. Del Rizzo and R. C. Trievel, *Epigenetics* **2011**, *6*, 1059-1067; c) S. C. Dillon, X. Zhang, R. C. Trievel and X. Cheng, *Genome Biol.* **2005**, *6*, 227-227; d) H.-M. Herz, A. Garruss and A. Shilatifard, *Trends Biochem. Sci.* **2013**, *38*, 621-639; e) M. Schapira, *Curr. Chem. Genomics* **2011**, *5*, 85-94.
- [7] a) H.-B. Guo and H. Guo, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 8797-8802; b) X. Zhang and T. C. Bruce, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 5728-5732.
- [8] M. Luo, *ACS Chem. Biol.* **2012**, *7*, 443-463.
- [9] M. R. Bennett, S. A. Shepherd, V. A. Cronin and J. Micklefield, *Curr. Opin. Chem. Biol.* **2017**, *37*, 97-106.
- [10] a) R. Belle, A. H. K. Al Temimi, K. Kumar, B. J. G. E. Pieters, A. Tumber, J. E. Dunford, C. Johansson, U. Oppermann, T. Brown, C. J. Schofield, R. J. Hopkinson, R. S. Paton, A. Kawamura and J. Mecinović, *Chem. Commun.* **2017**, *53*, 13264-13267; b) A. H. K. Al Temimi, Y. V. Reddy, P. B. White, H. Guo, P. Qian and J. Mecinović, *Sci. Rep.* **2017**, *7*, 16148; c) A. H. K. Al Temimi, R. S. Teeuwen, V. Tran, A. J. Altunc, D. C. Lenstra, W. Ren, P. Qian, H. Guo and J. Mecinović, *Org. Biomol. Chem.* **2019**, *17*, 5693-5697; d) A. H. K. Al Temimi, R. van der Wekken-de Bruijne, G. Proietti, H. Guo, P. Qian and J. Mecinović, *Bioconjugate Chem.* **2019**, *30*, 1798-1804.
- [11] a) I. R. Bothwell, K. Islam, Y. Chen, W. Zheng, G. Blum, H. Deng and M. Luo, *J. Am. Chem. Soc.* **2012**, *134*, 14905-14912; b) I. R. Bothwell and M. Luo, *Org. Lett.* **2014**, *16*, 3056-3059; c) K. Islam, I. Bothwell, Y. Chen, C. Sengelaub, R. Wang, H. Deng and M. Luo, *J. Am. Chem. Soc.* **2012**, *134*, 5909-5915; d) K. Islam, W. Zheng, H. Yu, H. Deng and M. Luo, *ACS Chem. Biol.* **2011**, *6*, 679-684; e) W. Peters, S. Willnow, M. Duisken, H. Kleine, T. Macherey, K. E. Duncan, D. W. Litchfield, B. Lüscher and E. Weinhold, *Angew. Chem. Int. Ed.* **2010**, *49*, 5170-5173; f) T. Shimazu, J. Barjau, Y. Sohtome, M. Sodeoka and Y. Shinkai, *PLoS ONE* **2014**, *9*, e105394; g) R. Wang, K. Islam, Y. Liu, W. Zheng, H. Tang, N. Lailler, G. Blum, H. Deng and M. Luo, *J. Am. Chem. Soc.* **2013**, *135*, 1048-1056; h) S. Willnow, M. Martin, B. Lüscher and E. Weinhold, *ChemBioChem* **2012**, *13*, 1167-1173.
- [12] a) F. Schlenk and J. L. Dainko, *Biochim. Biophys. Acta* **1975**, *385*, 312-323; b) S. Singh, J. Zhang, T. D. Huber, M. Sunkara, K. Hurley, R. D. Goff, G. Wang, W. Zhang, C. Liu, J. Rohr, S. G. Van Lanen, A. J. Morris and J. S. Thorson, *Angew. Chem. Int. Ed.* **2014**, *53*, 3965-3969.
- [13] J. R. Matos and C.-H. Wong, *Bioorg. Chem.* **1987**, *15*, 71-80.

- [14] a) C. Dalhoff, G. Lukinavicius, S. Klimasauskas and E. Weinhold, *Nat. Chem. Biol.* **2006**, *2*, 31-32; b) C. Dalhoff, G. Lukinavičius, S. Klimašauskas and E. Weinhold, *Nat. Prot.* **2006**, *1*, 1879.
- [15] R. J. Hopkinson, L. J. Walport, M. Münzel, N. R. Rose, T. J. Smart, A. Kawamura, T. D. W. Claridge and C. J. Schofield, *Angew. Chem. Int. Ed.* **2013**, *52*, 7709-7713.
- [16] D. F. Iwig, A. T. Grippe, T. A. McIntyre and S. J. Booker, *Biochemistry* **2004**, *43*, 13510-13524.
- [17] R. T. Borchardt and Y. S. Wu, *J. Med. Chem.* **1976**, *19*, 1099-1103.
- [18] C. S. Baxter and P. Byvoet, *Cancer Res.* **1974**, *34*, 1424-1428.
- [19] J. H. Alix, *Microbiol. Rev.* **1982**, *46*, 281-295.
- [20] S. A. Ross, M. Pitié and B. Meunier, *J. Chem. Soc., Perkin Trans. 1* **2000**, 571-574.
- [21] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus, *J. Comput. Chem.* **1983**, *4*, 187-217.
- [22] M. J. Field, P. A. Bash and M. Karplus, *J. Comput. Chem.* **1990**, *11*, 700-733.
- [23] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926-935.
- [24] C. L. Brooks, A. Brünger and M. Karplus, *Biopolymers* **1985**, *24*, 843-865.
- [25] M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai and G. Seifert, *Phys. Rev. B* **1998**, *58*, 7260-7268.
- [26] a) A. S. Christensen, T. Kubař, Q. Cui and M. Elstner, *Chem. Rev.* **2016**, *116*, 5301-5337; b) Q. Cui, M. Elstner, E. Kaxiras, T. Frauenheim and M. Karplus, *J. Phys. Chem. B* **2001**, *105*, 569-585.
- [27] A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin and M. Karplus, *J. Phys. Chem. B* **1998**, *102*, 3586-3616.
- [28] a) Y. Chu, Q. Xu and H. Guo, *J. Chem. Theory Comput.* **2010**, *6*, 1380-1389; b) Y. Chu, J. Yao and H. Guo, *PLoS ONE* **2012**, *7*, e37674; c) Q. Xu, Y.-z. Chu, H.-B. Guo, J. C. Smith and H. Guo, *Chem. Eur. J.* **2009**, *15*, 12596-12599.
- [29] G. M. Torrie and J. P. Valleau, *Chem. Phys. Lett.* **1974**, *28*, 578-581.
- [30] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen and P. A. Kollman, *J. Comput. Chem.* **1992**, *13*, 1011-1021.

Table of Contents



Histone lysine methyltransferases, a biomedically important class of epigenetic enzymes that catalyze methylation of histones in the presence of *S*-adenosylmethionine (AdoMet) cosubstrate, have a limited capacity to catalyze ethylation of histones in the presence of *S*-adenosylethionine (AdoEth) and *Se*-adenosylselenoethionine (AdoSeEth) cosubstrates.