JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

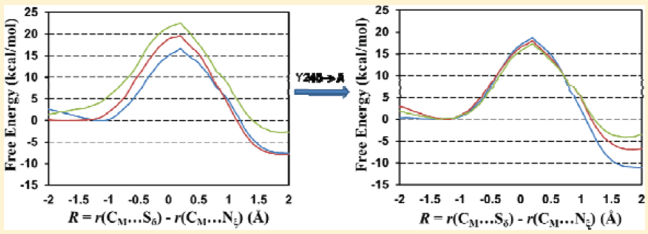
pubs.acs.org/jcim

Understanding Product Specificity of Protein Lysine Methyltransferases from QM/MM Molecular Dynamics and Free Energy Simulations: The Effects of Mutation on SET7/9 beyond the Tyr/Phe Switch

Jianzhuang Yao,[†] Yuzhuo Chu,[†,‡] Ran An,[†] and Hong Guo*[,†,‡]

[†]Department of Biochemistry and Cellular and Molecular Biology and [‡]National Institute for Mathematical and Biological Synthesis, University of Tennessee, Knoxville, Tennessee 37996, United States

ABSTRACT: The results of hybrid quantum mechanical/molecular mechanical (QM/MM) free energy (potential of mean force) simulations for methyl-transfer processes in SET7/9 and its Y245A mutant are compared to address the question concerning the change of the product specificity as well as catalytic efficiency due to the mutation. One of the key questions is whether or not the free energy profiles of methyl transfers may be used to predict the change of the product specificity as a result of the mutations for the residues that are not located at the Tyr/Phe switch position. The simulations show that while the wild-type SET7/9 is a monomethylase, the Y245→A mutation increases the ability of the enzyme to add more methyl groups on the target lysine (i.e., acting as a trimethylase). However, the first methyl-transfer process seems to become less efficient in the mutant compared to that in wild-type. All these results are consistent with experimental observations concerning the effects of the mutation on the product specificity and catalytic efficiency. Thus, the previous suggestion that the energetics of the methyl-transfer reactions may determine the product specificity, at least in some cases, is confirmed by the present work. Moreover, the dynamic information of the reactant complexes obtained from the QM/MM molecular dynamics simulations shows that the ability of the reactant complexes to form the reactive transition-state-like configurations may be used as an important indicator for the prediction of the product specificity of PKMTs, consistent with previous computational studies.



INTRODUCTION

Lysine residues on the tails of histone proteins may be methylated by protein lysine methyltransferases (PKMTs) using *S*-adenosyl-L-methionine (AdoMet) as the methyl donor.[1,2] This and some other post-translational modifications generate important epigenetic marks of the histone code for regulation of chromatin.[3] One of the key properties of PKMTs is their product specificities that determine how many methyl groups that the enzyme may "write" on the target lysine.[2,4,5] Different methylation states produced by different PKMTs can lead to specific patterns of histone modifications and networks of interacting proteins for the regulation of chromatin structure and gene expression.[6−8] Determining the factors that control the product specificity of the enzymes and understanding the reasons that stop the continuation of methylation process (up to a maximum of three methyl groups) are therefore of considerable importance.

The product specificity of PKMTs has been a subject of extensive experimental[9−20] and computational investigations[21−30] in the last several years. Previous structural and biochemical studies[5] have identified some active-site structural features that seem to be of importance in controlling the product specificity of the enzymes.[10,14,15] One important feature is the general occurrence of Tyr/Phe residues at certain positions of the active site.[10,11,13,15] It has been found that replacement of some of the Try/Phe residues could lead to changes of product specificity and/or enzyme's catalytic efficiency. For instance, a single Tyr or Phe residue occupies a structurally similar position in the active sites of several PKMTs with this residue normally being Phe for di- and trimethylases (e.g., Phe-281 in DIM-5[10]) and Tyr for monomethylases (e.g., Y305 in SET7/9[19] and Y334 in SET8).[11,14] Interestingly, the trimethylase DIM-5 can be converted into a mono/dimethylase by the F281→Y mutation, whereas the monomethylases SET7/9 and SET8 may be changed to dimethylases through the corresponding Tyr→Phe mutation (e.g., Y305→F and Y334→F mutations for SET7/9 and SET8, respectively).[10,11,14] In other words, the Tyr→Phe mutation at this position (termed as Tyr/Phe switch position) of the active site tends to increase the ability of the enzyme to add more methyl group(s) to the target lysine (if the residue is Tyr in wild-type), while the Phe→Tyr mutation tends to decrease this ability (if the residue is Phe in wild-type). The similar observations have been made for some other PKMTs.[15,17] Different suggestions have been proposed to

Received: August 2, 2011
Published: January 16, 2012

explain how this Tyr/Phe switch residue could affect the product specificity of PKMTs,[14,17,19] including the proposal that the Tyr/Phe switch may regulate the product specificity through changing the size of the active site or altering the affinity of water molecule(s) at the active site.

Previous investigations have demonstrated that mutations of residues at some other positions of the active sites of PKMTs may also lead to changes of product specificity and/or catalytic efficiency.[10,17,19] However, unlike the Tyr/Phe switch mentioned above, only limited studies have been performed for determining the role of these residues on the product specificity and for understanding the origin for the effects of these mutations. The enzymes for which such changes of product specificity have been observed include, but not limited to, SET7/9,[13,19] G9a methyltransferase and its close relative G9a-like-protein (GLP).[15,17] For instance, there is a Tyr residue (Y245) hydrogen-bonding to the methyl accepting nitrogen of the target lysine in the X-ray structures of SET7/9 complexed with substrates.[13,19] It has been found that the Y245→A mutation is able to increase the enzyme's ability to add two additional methyl groups to the target lysine (even though Y245 is not really located at the Tyr/Phe switch position), and the Y245A mutant is therefore a trimethylase.[13,19] Nevertheless, the monomethylation process seems to become significantly less efficient for the Y245A mutant compared to that in wild-type, probably due in part to the reduction of transition-state stabilization for the methyl transfer as a result of this mutation[23] and/or less optimum orientations between the methyl donor and acceptor.[28]

We and some other groups[21−30] have studied PKMTs by computer simulations and tried to understand their product specificities as well as changes of the product specificities as a result of certain mutations. The energetic origin for the existence of specific product specificity was examined for some PKMTs and their mutants. It was proposed based on systematic investigations of the energetics for the first, second, and third methyl transfers in the enzymes that the reason for the stop of methylation process at certain stage (i.e., before the completion of all the three methyl additions) was probably due to a significant increase of the activation barrier for one of the methyl-transfer reactions.[21−23,26] Therefore, the relative efficiencies of the chemical steps involving the three methyl transfers from AdoMet to the ε-amino group of the target lysine/methyl lysine in PKMTs may control, at least in some cases, how the epigenetic marks of lysine methylation are written.[21−23] For instance, for SET8 the simulations showed that the Y334→F mutation delayed the first occurrence of an inefficient methyl transfer (i.e., with a relatively high free-energy barrier) from the second methyl transfer to the third methyl transfer.[22] Thus, in contrast to the wild-type enzyme that may not be able to catalyze the second methyl transfer (i.e., a monomethylase), the Y334F mutant becomes a dimethylase. Interestingly, it has been shown that this principle for understanding product specificity might even be applied to some of the inactive mutants (e.g., the F281W mutant of DIM-5) that have a relatively high activation barrier even for the first methyl transfer.[21] However, question remains as to whether the previous conclusions[21,22] could be generalized to the cases involving residues other than that occupying the Tyr/Phe switch position. As a first step of answering this question, the quantum mechanical/molecular mechanical (QM/MM) free energy simulations are applied to SET7/9 and its Y245A mutant with the TAF10 substrate in this work. The free energy barriers for the methyl transfers obtained from the simulations are found to be well correlated with the experimental observations on the product specificities for these cases, supporting our earlier suggestions.[21−23] Specifically, unlike wild-type, the Y245A mutant is found to act like a trimethylase based on the results of the simulations. However, the first methyl-transfer process is less efficient in Y245A than in wild-type, consistent with the results of previous experimental studies[13,19] and molecular dynamics (MD) simulations.[28] Furthermore, the QM/MM MD simulations are performed on the reactant complexes of the first, second, and third methyl transfers for each of the systems systematically, and a good correlation is obtained between the ability of the reactant complexes to form the reactive transition-state-like configurations for the methyl transfers and product specificity.

METHODS

QM/MM free energy (potential of mean force) simulations were applied in the determination of the free energy profiles for the first, second, and third methyl transfers in SET7/9 and its Y245A mutant using the models generated from the newly determined X-ray structures containing AdoHcy and unmodified or methylated TAF10 peptide (see below);[19,20] the MD simulations were also performed on the reactant complexes of the methyl transfers for each system to study their dynamic properties. AdoMet/AdoHcy and the side chains of the target lysine/methyl lysine were treated by QM, and the rest of the systems were treated by MM. The link-atom approach[31] implemented with CHARMM[32] was used to separate the QM and MM regions. The reason for this QM and MM partition is to be consistent with our previous studies[21,22] which seem to be able to reproduce experimental results. This may also make it easier for comparison. The stochastic boundary MD method[33] was used for the QM/MM MD and free energy simulations. The system was separated into a reaction and a reservoir zone, and the reaction zone was further divided into a reaction and a buffer region. The reaction region was a sphere with radius r of 22 Å, and the buffer region had r equal to 20 Å $\leq r \leq$ 22 Å. The reference center for partitioning the system was chosen to be the C_δ atom of the target lysine/methyl lysine. The resulting systems contained around 5500 atoms, including about 900 water molecules.

The self-consistent charge-density functional tight-binding (SCC-DFTB) method[34] implemented in CHARMM[32] was used for the QM atoms, and the all-hydrogen CHARMM potential function (PARAM27)[35] was used for the MM atoms. A modified TIP3P water model[36,37] was employed for the solvent. In our earlier studies,[21−23] the results of the SCC-DFTB and B3LYP/6-31G** methods for the description of the methyl transfer in a small model system were compared using an energy minimization-based approach. This comparison allowed us to understand the performance of the semi-empirical method in the description of the bond breaking and making for the methyl transfer and to derive an empirical formula for the correction of the errors in the free energy curves from potential of mean force simulations. It was shown that that the energy curves from the corrected SCC-DFTB and B3LYP/6-31G** were very close,[21−23] supporting the use of this approach to make the first-order correction to the bond breaking and making events involving the simple and similar S_N2 methyl-transfer processes. The similar approach of this first-order correction to the free energy curves was also adopted in this work. For a more detailed discussion, see refs 21−23. It should

dx.doi.org/10.1021/ci200364m | J. Chem. Inf. Model. 2012, 52, 449−456

be pointed out that the key energetic properties used for understanding the product specificity are the relative free energy barriers for the different methyl transfers that depend only on the free energy profiles from the reactant complexes to the transition states.

The initial coordinates of the reactant complexes for the first, second, and third methyl transfers were based on the crystallographic complexes of SET7/9 and its Y245A mutant containing AdoHcy and unmodified or methylated TAF10 peptide (PDB codes: 3M53, 3M57, 3M58, 3M59, and 2F69).[19,20] In each case, a methyl group was added to AdoHcy to form AdoMet, the methyl donor. For the reactant complexes of the first, second, and third methyl transfers in Y245A, the X-ray structures containing K189 (3M57), K189me (3M58), and K189Me2 (3M59) were used, respectively, to generate the corresponding reactant complexes. For the second and third methyl transfers in wild-type, the reactant complexes were generated by replacing one and two hydrogen atoms on the target lysine in the X-ray structure (3M53) by one and two methyl groups, respectively, to generate the corresponding reactant complexes. The initial structures for the entire stochastic boundary system 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 and equilibrated at 298.15 K for 500 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. After 500 ps of equilibration were performed, 1 ns QM/MM MD simulations were carried out for each of the reactant complexes of the first, second, and third methyl transfers. As was discussed in the previous studies,[21−23] the S_N2 methyl transfer from AdoMet to lysine/methyl lysine is presumably more efficient if the S−CH_3 group of AdoMet is well aligned with the lone pair of electrons on N_ξ in the reactant complex, i.e., with a small θ angle and relatively short C_M–N_ξ distance (about 3 Å). Therefore, we determined the distributions of $r(C_M–N_\xi)$ and θ from the MD trajectories; θ was defined as the angle between the direction of the C_M–S_δ bond ($\mathbf{r}_2$) and the direction of the electron lone pair ($\mathbf{r}_1$) (see Figure 1). Moreover, the histogram method was used in the calculation of the probability density distributions for $r(C_M–N_\xi)$ and θ as well as the relative free energies as a function of $r(C_M–N_\xi)$ (For details, see refs 21−23.).

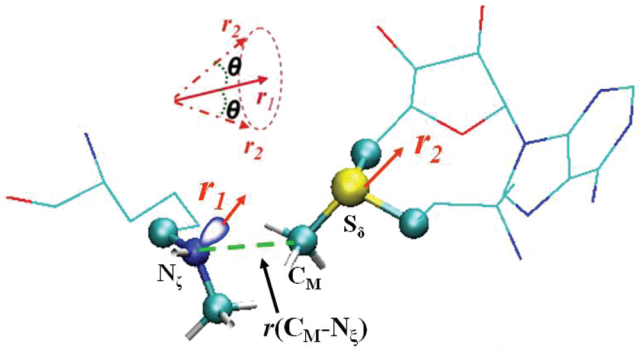


Figure 1. Definition of the structural parameters for monitoring the relative orientation of AdoMet and K189me1 [K189 and K189(me)$_2$] in the reactant complex. The efficiency of the methyl transfer may be related to the distributions of $r(C_M\cdots N_\zeta)$ and θ in the reactant complexes. θ is defined as the angle between the two vectors $\mathbf{r}_1$ and $\mathbf{r}_2$. Here $\mathbf{r}_1$ is the direction of the lone pair of electrons on N_ζ, and $\mathbf{r}_2$ is the vector pointing from C_M to S_δ. The reaction coordinate for calculating the free energy profiles for the methyl transfers is $R = r(C_M\cdots S_\delta) - r(C_M\cdots N_\zeta)$.

The umbrella sampling method[38] implemented in the CHARMM program along with the weighted histogram analysis method (WHAM)[39] was applied to determine the change of the free energy (potential of mean force) as a function of the reaction coordinate for the methyl transfer from AdoMet to lysine/methyl lysine in SET7/9 and Y245A. The reaction coordinate was defined as a linear combination of $r(C_M–N_\xi)$ and $r(C_M–S_\delta)$ [$R = r(C_M–S_\delta) - r(C_M–N_\xi)$] (see Figure 1). For each methyl-transfer process, 20 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 from 50 to 400 kcal mol^{-1} Å^{-2}.

RESULTS AND DISCUSSION

The average active-site structure of the reactant complex and the structure near the transition state for the first methyl transfer in SET7/9 are shown in Figure 2A,C, respectively. Figure 2A shows that the average active-site structure of the reactant complex has the lone pair of electrons on N_ξ of the target lysine well aligned with the methyl group of AdoMet, in agreement with previous MD studies.[21−23,26−28] These previous simulations already showed that there seemed to be a good correlation between the alignment of the methyl group with the lone pair of electrons and the efficiency of the methyl transfer. Y245 and an active-site water molecule (W1) accept stable hydrogen bonds from the ε-amino group of target lysine, and these interactions may help to orientate the electron lone pair toward the methyl group of AdoMet in the reactant complex. It has been observed in our earlier study[23] that these hydrogen bonds became generally stronger as the system changed from the reactant complex to transition state. The similar observation was also made in the present simulations. Indeed, the average distances for the two hydrogen bonds decrease by 0.2−0.6 Å in going from the reactant complex to the transition state. Such strengthening of the hydrogen-bonding interactions may play a role in the transition-state stabilization and increase the efficiency of the methyl transfer, as proposed previously.[23]

Consistent with the average structure in Figure 2A and previous investigations,[21−23,26−28] the distribution plot (Figure 2B, left) shows that there is a large population of the structures with relative short $r(C_M–N_\xi)$ distances and small values of the θ angle; the average $r(C_M–N_\xi)$ distance in the reactant complex is around 2.9 Å, and the average θ angle is less than 30°. Figure 2C shows that the $r(C_M–N_\xi)$ and $r(C_M–S_\delta)$ distances change to around 2.1 and 2.3 Å, respectively, as the system reaches the structure near the transition state for the methyl transfer. The similar structural parameters have also been obtained for the transition states of other methyl transfers, including those for the second and third methyl transfers in SET7/9 and all the three methyl transfers in the Y245A mutant of SET7/9 (see below).

It has been proposed that the CH···O hydrogen bonding between the AdoMet methyl group and oxygen atoms within the SET domain active site may play an important role in AdoMet binding and catalysis.[40,41] The interactions involving the C−H group have attracted considerable attention recently in the structural biology and computational chemistry communities due in part to the frequent occurrence of short

dx.doi.org/10.1021/ci200364m | J. Chem. Inf. Model. 2012, 52, 449−456

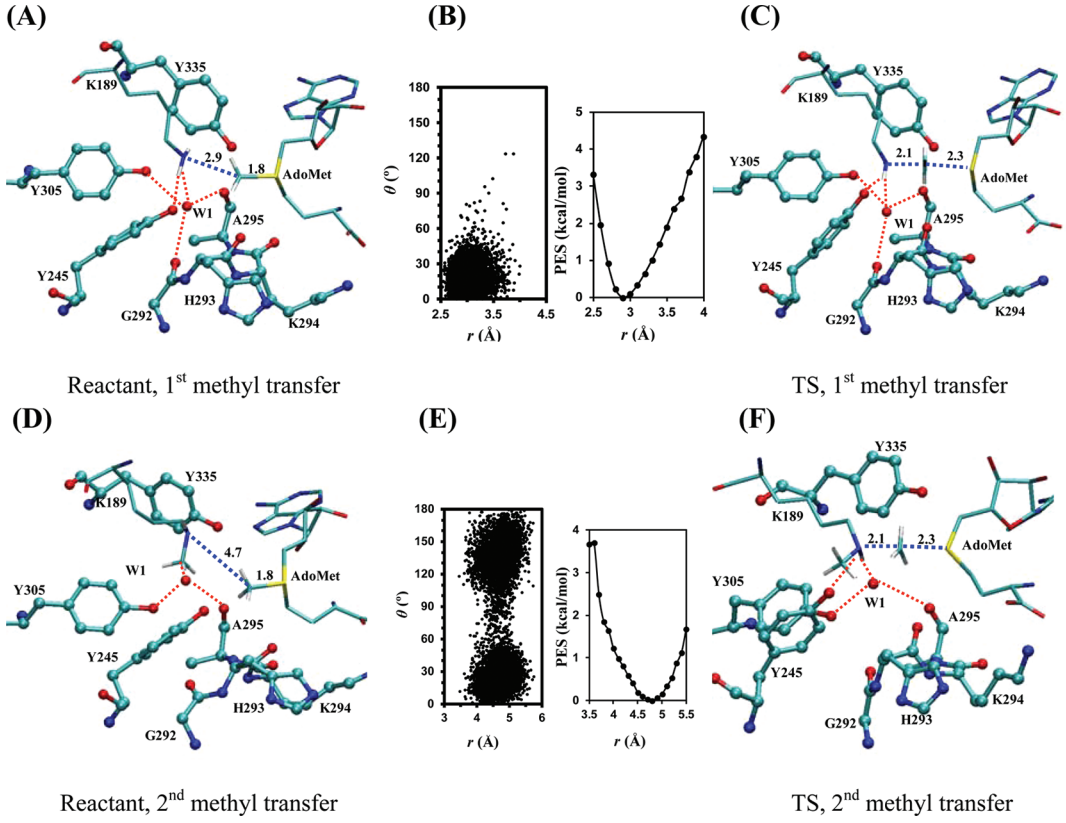


Figure 2. MD results for the wild-type enzyme (SET7/9) with TAF10-K189 peptide as the substrate. (A) The average active-site structure of the reactant complex for the first methyl transfer. SET7/9 is shown in balls and sticks, and AdoMet and TAF10-K189 are in sticks. Hydrogen atoms are not shown for clarity except for those on N_ζ and the transferable methyl group. Hydrogen bonds are indicated by red dotted lines, and the distances related to the reactant coordinate are also shown. (B) Two-dimensional plot of $r(C_M\cdots N_\zeta)$ and θ distributions based on the 1 ns simulations of the reactant complex for the first methyl transfer along with the corresponding free energy change as a function of $r(C_M\cdots N_\zeta)$ obtained from the distributions. (C) The average structure near the transition state for the first methyl transfer obtained from the free energy (potential of mean force) simulations. (D) The average active-site structure of the reactant complex for the second methyl transfer. (E) The two-dimensional plot of $r(C_M\cdots N_\zeta)$ and θ distributions based on the 1 ns simulations of the reactant complex for the second methyl transfer along with the corresponding free energy change as a function of $r(C_M\cdots N_\zeta)$. (F) The average structure near the transition state for the second methyl transfer obtained from the free energy simulations.

C−H···O contacts in the structures of proteins, peptides, and DNA. For a review of the C−H···O interactions and the results of high-level ab initio QM calculations based on protein structures as well as possible polarization effects,[42] see refs 43 and 44. For SET7/9, one of the possible interactions involves Y335 (see Figure 2), a conserved residue.[40,41] The results of the simulations show that Y335 is indeed within the accepted carbon−oxygen hydrogen-bond distances;[43] for the first methyl transfer, the average distance from oxygen atom of Y335 to carbon atom of the transferable methyl group is about 3.3 Å in the both structures of the reactant complex and transition state.

The average active-site structure of the reactant complex for the second methyl transfer in SET7/9 is shown in Figure 2D; the structure near the transition state is plotted in Figure 2F. In contrast to the case of the first methyl transfer, the lone pair of electrons on N_ζ of the methylated lysine cannot be well aligned with the methyl group of AdoMet in the reactant complex; the average $r(C_M-N_\zeta)$ distance is 4.7 Å. This fact is further demonstrated by the distribution plot in Figure 2E (left) which shows the $r(C_M-N_\zeta)$ distance is generally long during the MD simulations with a broad distribution of θ; the free energy cost for generating the structures with shorter $r(C_M-N_\zeta)$ distance (e.g., 3 Å) more suitable for the reaction is expected to be quite high based on the free energy plot (Figure 2E, right). These results are consistent with previous investigations.[21,23,26−28] Figure 2D,E shows that W1 is present at the active site in the both reactant complex and transition state, and previous structural studies seem to suggest that the failure for W1 to dissociate may inhibit the further methylation of the monomethyl lysine by wild-type SET7/9 and SET8 (an enzyme that has the similar product specificity as SET7/9).[14,19] Consistent with this suggestion, our earlier QM/MM MD and free energy simulations on SET8[22] showed that while W1 was present at the active site of wild-type SET8 during the second methyl transfer, it moved away from its original position in the Y334F mutant (the Tyr/Phe switch mutation) of SET8 during the second methyl transfer. This seems to make the second methyl transfer more efficient in Y334F and prevents the stop of the methylation process after the first methyl transfer. Comparison of the structures in Figure 2A,C shows that the structure of the reactant complex for the first methyl transfer in wild-type is rather similar to the corresponding structure near the transition state. This is in contrast to the case for the second methyl transfer where the reactant structure (Figure 2D) is significantly distorted from the corresponding transition-state structure (Figure 2F).The results support the earlier suggestion that one of the reasons for the existence of the relatively low free-energy barrier for the first methyl transfer

dx.doi.org/10.1021/ci200364m | J. Chem. Inf. Model. 2012, 52, 449−456

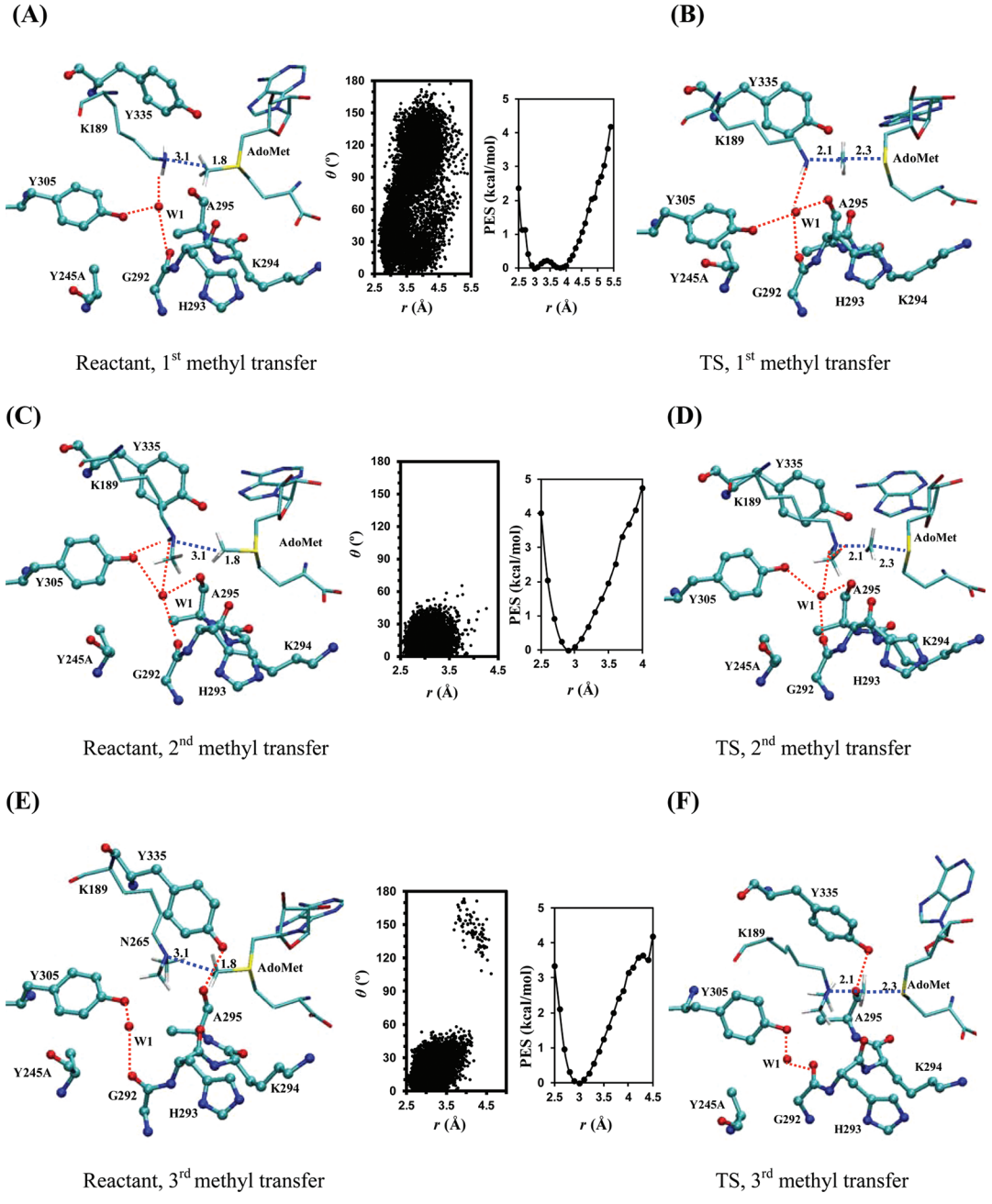


Figure 3. MD results for Y245A. (A) Left: The average active-site structure of the reactant complex for the first methyl transfer in the mutant. Right: the two-dimensional plot of $r(C_M \cdots N_\zeta)$ and θ distributions based on the 1 ns simulations of the reactant complex for the first methyl transfer along with the corresponding free energy change as a function of $r(C_M \cdots N_\zeta)$ obtained from the distributions. (B) The average structure near the transition state for the first methyl transfer obtained from the free energy simulations. (C) Left: The average active-site structure of the reactant complex for the second methyl transfer. Right: The two-dimensional plot of $r(C_M \cdots N_\zeta)$ and θ distributions based on the 1.5 ns simulations of the reactant complex for the second methyl transfer along with the corresponding free energy change as a function of $r(C_M \cdots N_\zeta)$. (D) The average structure near the transition state for the second methyl transfer obtained from the free energy simulations. (E) Left: The average active-site structure of the reactant complex for the third methyl transfer. Right: The two-dimensional plot of $r(C_M \cdots N_\zeta)$ and θ distributions of the reactant complex for the third methyl transfer along with the corresponding free energy change as a function of $r(C_M \cdots N_\zeta)$. (F) The average structure near the transition state for the third methyl transfer obtained from the free energy simulations.

compared to those for the second and third methyl transfers (see below) is likely owed to the fact that a part of transition-state stabilization is already reflected on the reactant state of the first methyl transfer through the generation of the transition-state-like conformation.[21,22] For the second methyl transfer in wild-type, the average distance from oxygen atom of Y335 to carbon atom of the transferable methyl group is significantly longer than that for the first methyl transfer (about 4.3 and 3.8 Å for the reactant complex and transition state of the second methyl transfer, respectively).

Figure 3A,C,E shows the average structures of the reactant complexes for the first, second, and third methyl transfers, respectively, in the Y245A mutant of SET7/9 along with the distribution plots and free energy analyses; the corresponding structures near the transition states are given in Figure 3B,D,F, respectively. The lone pair of electrons on N_ζ of the target

dx.doi.org/10.1021/ci200364m | J. Chem. Inf. Model. 2012, 52, 449−456

lysine/monomethyl lysine/dimethyl lysine seems to be able to align with the methyl group of AdoMet reasonably well for the first, second, and third methyl transfers; the average $r(C_M-N_\xi)$ distances are around 3.1 Å. The distribution is broad for the first methyl transfer. Nevertheless, the population for the structures with the well aligned lysine and methyl group of AdoMet is still quite high from our simulations, as indicated by the free energy plot in Figure 3A (right). Indeed, the free energy cost for generating structures with $r(C_M-N_\xi) \sim 3$ Å is negligible. It appears that additional space at the active site generated by the substitution of Y245 by Ala would make it easier for the addition of more methyl groups (see below).

The free energy profiles for the first, second, and third methyl transfers in SET7/9 and Y245A are shown in Figure 4A, B, respectively. As is evident from Figure 4A, for wild-type, the free energy barriers for the second and third methyl transfers are significantly higher than that for the first methyl transfer (about 3 and 6 kcal/mol higher, respectively). Therefore, addition of the second methyl would probably become more difficult, and the methylation process would stop after the first methyl transfer, consistent with the observation that the enzyme is a monomethylase. It is of interest to note that the difference (about 3 kcal/mol) for the free energy barriers of the first and second methyl transfers is quite similar to an earlier result (about 3.7 kcal/mol) from a quite different QM/MM approach;[26] some of the difference might be due to the use of different X-rays structures with different substrates. For Y245A, the free energy barriers for all the three methyl-transfer steps are not very significantly different. Figure 4 shows that the free energy barrier for the first methyl transfer is about 2 kcal/mol higher in Y245A than that in wild-type, and as a result, the methyl-transfer process probably becomes slower in the mutant. This is consistent with previous experimental results[13,19] which showed that the k_{cat} value for the monomethylation in Y245A was about 30-fold lower than that in wild-type. The increase of the free energy barrier for the first methyl transfer in going from wild-type to Y245A could be related to the loss of the transition-state stabilization involving Y245 (see above), although other factors may be involved as well.[19,28] Interestingly, the experimental study also suggested that for Y245A, the k_{cat} values for dimethylation and trimethylation are higher than that for monomethylation. Our simulations results given in Figure 4B show that the barriers for the second and third methyl transfers are indeed 0.7 and 1.3 kcal/mol lower than that for the first methyl transfer. In our earlier studies,[21,22] we proposed the use of two free energy triplets, (0, Δ_{2-1W}, Δ_{3-1W}) and (Δ_{M-W}, Δ_{2-1M}, Δ_{3-1M}), to describe the energetics of methyl transfers in PKMTs and their mutants (see Figure 4 for explanation). The energy triplets for wild-type and the Y245A mutant can be written as (0, 3, 5.9) and (2.1, −0.7, −1.3), respectively. It should be pointed out that the relative free energy barriers, as opposed to the absolute barriers, are expected to be more reliable and important in the determination of the product specificity. The relative free energy barriers are expected to be less sensitive to the choice of the QM method and experimental structural differences due to the cancellation of the errors. This is one of the reasons that we proposed the use of two free energy triplets to describe the energetics of methyl transfers in PKMTs and their mutants that are based on the relative free energy barriers.

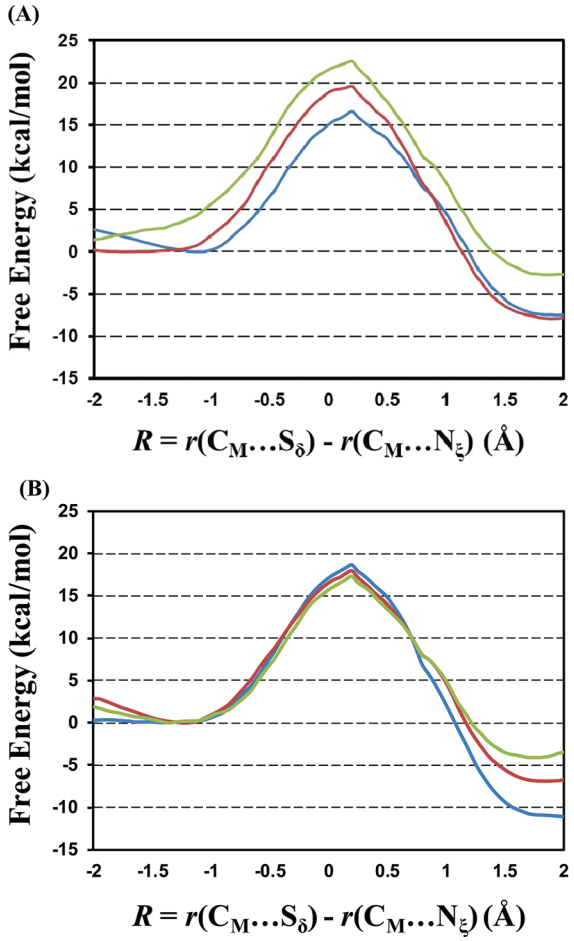


Figure 4. (A) Free energy (potential of mean force) changes for the first, second, and third methyl transfers, respectively, as a function of the reaction coordinate [$R = r(C_M\cdots S_\delta) - r(C_M\cdots N_\zeta)$] in the wild-type SET7/9 with TAF10-K189 peptide as the substrate. The first methyl transfer: blue line with a free energy barrier of 16.6 kcal/mol; the second methyl transfer: red line with a free energy barrier of 19.6 kcal/mol; the third methyl transfer: green line with a free energy barrier of 22.5 kcal/mol. Two free energy triplets, (0, Δ_{2-1W}, Δ_{3-1W}) and (Δ_{M-W}, Δ_{2-1M}, Δ_{3-1M}) for wild-type and mutated enzyme, respectively, have been proposed to describe the product specificity of PKMTs and their mutants. For wild-type enzyme, the second (Δ_{2-1W}) and third (Δ_{3-1W}) parameters are the differences in the free-energy barriers between the second and first and between the third and first methyl transfers, respectively. For the mutated enzyme, the first parameter (Δ_{M-W}) is the difference in the free-energy barriers for the first methyl transfer in the wild-type and mutant. The second (Δ_{2-1M}) and third (Δ_{3-1M}) parameters are the differences in the free-energy barriers between the second and first and between the third and first methyl transfers, respectively, for the mutant. The energy triplet for SET7/9 with TAF10-K189 substrate was therefore calculated to be (0, 3, 5.9). (B) The free energy changes for the first, second, and third methyl transfers as a function of the reaction coordinate in the Y245A mutant. The first methyl transfer: blue line with a free energy barrier of 18.7 kcal/mol; the second methyl transfer: red line; the third methyl transfer: green line. The energy triplet for Y245A is therefore (2.1, −0.7, −1.3).

CONCLUSIONS

In this study, the QM/MM MD and free energy simulations have been performed on different methyl-transfer steps from AdoMet to target lysine/methyl-lysine in SET7/9 and its Y245A mutant. As far as the product specificity is concerned,

dx.doi.org/10.1021/ci200364m | J. Chem. Inf. Model. 2012, 52, 449−456

one of the key questions is what causes the stop of further methyl addition during histone lysine methylation (which is different from the question concerning the rate-limiting step of the enzyme-catalyzed methyl-transfer process). In the earlier publications,[21−23] DIM-5, SET7/9, and SET8 as well as their mutants at the Tyr/Phe switch position were studied, and it was shown that the stop of methylation (product specificity) of PKMTs could be identified based on the relative free energy barriers for the three methyl transfers (the energy triplets) as well as the dynamic properties of the reactant complexes. The next question that needs to be addressed is whether or not the free energy profiles and the energetic triplets may be used to understand and predict the changes of the product specificity as a result of the mutations at other positions. This question was addressed in the present study. The free energy barriers for the methyl transfers obtained here were found to be well correlated with the experimental observations on the change of product specificity due to the Y245→A mutation. Indeed, the simulations showed that while the wild type SET7/9 may act like a monomethylase, the Y245→A mutation could increase the ability of SET7/9 to add two more methyl groups on the target lysine and lead to a trimethylase. The simulations also suggested that the first methyl-transfer process in Y245A might become less efficient compared to the first methyl transfer in wild-type, consistent with experimental observations and previous MD simulations (see above). The QM/MM MD simulations were also performed on the reactant complexes of the first, second, and third methyl transfers for each of the systems, and a good correlation was obtained between the ability of the reactant complexes to form the reactive configurations for the methyl transfers and product specificity.

AUTHOR INFORMATION

Corresponding Author

*E-mail: hguo1@utk.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work is supported in part by the National Science Foundation Award (grant no. 0817940 to H.G.) and the NSF TeraGrid resources provided by University of Texas at Austin. Y.Z.C is supported by a fellowship from the National Institute for Mathematical and Biological Synthesis (NIMBioS), an Institute sponsored by the National Science Foundation, the U.S. Department of Homeland Security, and the U.S. Department of Agriculture through NSF Award EF-0832858, with additional support from The University of Tennessee, Knoxville. We thank Prof. Martin Karplus for a gift of CHARMM program and Prof. Xiaodong Cheng and Drs. Qin Xu and Haobo Guo for useful discussions.

REFERENCES

(1) Jenuwein, T. The epigenetic magic of histone lysine methylation. *FEBS J.* **2006**, *273*, 3121−3135.

(2) Martin, C.; Zhang, Y. The diverse functions of histone lysine methylation. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 838−849.

(3) Strahl, B. D.; Allis, C. D. The language of covalent histone modifications. *Nature* **2000**, *403*, 41−45.

(4) Cheng, X.; Collins, R. E.; Zhang, X. Structural and sequence motifs of protein (histone) methylation enzymes. *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 267−294.

(5) Xiao, B.; Wilson, J. R.; Gamblin, S. J. SET domains and histone methylation. *Curr. Opin. Struct. Biol.* **2003**, *13*, 699−705.

(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) Lall, S. Primers on chromatin. *Nat. Struct. Mol. Biol.* **2007**, *14*, 1110−1115.

(8) Turner, B. M. Reading signals on the nucleosome with a new nomenclature for modified histones. *Nat. Struct. Mol. Biol.* **2005**, *12*, 110−112.

(9) Zhang, X.; Tamaru, H.; Khan, S. I.; Horton, J. R.; Keefe, L. J.; Selker, E. U.; Cheng, X. Structure of the Neurospora SET domain protein DIM-5, a histone H3 lysine methyltransferase. *Cell* **2002**, *111*, 117−27.

(10) Zhang, X.; Yang, Z.; Khan, S. I.; Horton, J. R.; Tamaru, H.; Selker, E. U.; Cheng, X. Structural basis for the product specificity of histone lysine methyltransferases. *Mol. Cell* **2003**, *12*, 177−85.

(11) Couture, J. F.; Collazo, E.; Brunzelle, J. S.; Trievel, R. C. Structural and functional analysis of SET8, a histone H4 Lys-20 methyltransferase. *Genes Dev.* **2005**, *19*, 1455−1465.

(12) Xiao, B.; Jing, C.; Kelly, G.; Walker, P. A.; Muskett, F. W.; Frenkiel, T. A.; Martin, S. R.; Sarma, K.; Reinberg, D.; Gamblin, S. J.; Wilson, J. R. Specificity and mechanism of the histone methyltransferase Pr-Set7. *Genes Dev.* **2005**, *19*, 1444−1454.

(13) Xiao, B.; Jing, C.; Wilson, J. R.; Walker, P. A.; Vasisht, N.; Kelly, G.; Howell, S.; Taylor, I. A.; Blackburn, G. M.; Gamblin, S. J. Structure and catalytic mechanism of the human histone methyltransferase SET7/9. *Nature* **2003**, *421*, 652−656.

(14) Couture, J. F.; Dirk, L. M. A.; Brunzelle, J. S.; Houtz, R. L.; Trievel, R. C. Structural origins for the product specificity of SET domain protein methyltransferases. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 20659−20664.

(15) Collins, R. E.; Tachibana, M.; Tamaru, H.; Smith, K. M.; Jia, D.; Zhang, X.; Selker, E. U.; Shinkai, Y.; Cheng, X. In vitro and in vivo analyses of a Phe/Tyr switch controlling product specificity of histone lysine methyltransferases. *J. Biol. Chem.* **2005**, *280*, 5563−5570.

(16) Tachibana, M.; Sugimoto, K.; Fukushima, T.; Shinkai, Y. SET domain-containing protein, G9a, is a novel lysine-preferring mammalian histone methyltransferase with hyperactivity and specific selectivity to lysines 9 and 27 of histone H3. *J. Biol. Chem.* **2001**, *276*, 25309−25317.

(17) Wu, H.; Min, J. R.; Lunin, V. V.; Antoshenko, T.; Dombrovski, L.; Zeng, H.; Allali-Hassani, A.; Campagna-Slater, V.; Vedadi, M.; Arrowsmith, C. H.; Plotnikov, A. N.; Schapira, M. Structural Biology of Human H3K9Methyltransferases. *PLoS One 5*, e8570.

(18) Xu, S. T.; Wu, J.; Sun, B. F.; Zhong, C.; Ding, J. P. Structural and biochemical studies of human lysine methyltransferase Smyd3 reveal the important functional roles of its post-SET and TPR domains and the regulation of its activity by DNA binding. *Nucleic Acids Res.* **2011**, *39*, 4438−4449.

(19) Del Rizzo, P. A.; Couture, J. F.; Dirk, L. M. A.; Strunk, B. S.; Roiko, M. S.; Brunzelle, J. S.; Houtz, R. L.; Trievel, R. C. SET7/9 Catalytic Mutants Reveal the Role of Active Site Water Molecules in Lysine Multiple Methylation. *J. Biol. Chem.* **2010**, *285*, 31849−31858.

(20) Couture, J. F.; Collazo, E.; Hauk, G.; Trievel, R. C. Structural basis for the methylation site specificity of SET7/9. *Nat. Struct. Mol. Biol.* **2006**, *13*, 140−146.

(21) Xu, Q.; Chu, Y-Z; Guo, H-B; Smith, J. C.; Guo, H. Energy Triplets for Writing Epigenetic Marks: Insights from QM/MM Free-Energy Simulations of Protein Lysine Methyltransferases. *Chemistry* **2009**, *15*, 12596−12599.

(22) Chu, Y. Z.; Xu, Q.; Guo, H. Understanding Energetic Origins of Product Specificity of SET8 from QM/MM Free Energy Simulations: What Causes the Stop of Methyl Addition during Histone Lysine Methylation? *J. Chem. Theory Comput.* **2010**, *6*, 1380−1389.

(23) Guo, H. B.; Guo, H. Mechanism of histone methylation catalyzed by protein lysine methyltransferase SET7/9 and origin of

product specificity. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 8797−8802.
(24) Zhang, X. D.; Bruice, T. C. Product specificity and mechanism of protein lysine methyltransferases: Insights from the histone lysine methyltransferase SET8. *Biochemistry* **2008**, *47*, 6671−6677.
(25) Bai, Q. F.; Shen, Y. L.; Yao, X. J.; Wang, F.; Du, Y. P.; Wang, Q.; Jin, N. Z.; Hai, J.; Hu, T. J.; Yang, J. B. Modeling a New Water Channel That Allows SET9 to Dimethylate p53. *PLoS One* **2011**, *6*, e19856.
(26) Hu, P.; Wang, S.; Zhang, Y. How do SET-domain protein lysine methyltransferases achieve the methylation state specificity? Revisited by ab initio QM/MM molecular dynamics simulations. *J. Am. Chem. Soc.* **2008**, *130*, 3806−3813.
(27) Wang, S. L.; Hu, P.; Zhang, Y. K. Ab initio quantum mechanical/molecular mechanical molecular dynamics simulation of enzyme catalysis: The case of histone lysine methyltransferase SET7/9. *J. Phys. Chem. B* **2007**, *111*, 3758−3764.
(28) Hu, P.; Zhang, Y. K. Catalytic mechanism and product specificity of the histone lysine methyltransferase SET7/9: An ab initio QM/MM-FE study with multiple initial structures. *J. Am. Chem. Soc.* **2006**, *128*, 1272−1278.
(29) Zhang, X.; Bruice, T. C. Enzymatic mechanism and product specificity of SET-domain protein lysine methyltransferases. *Proc. Natl. Acad. Sc.i U. S. A.* **2008**, *105*, 5728−5732.
(30) Georgieva, P.; Himo, F. Quantum Chemical Modeling of Enzymatic Reactions: The Case of Histone Lysine Methyltransferase. *J. Comput. Chem.* **2010**, *31*, 1707−1714.
(31) Field, M. J.; Bash, P. A.; Karplus, M. A combined quantum-mechanical and molecular mechanical potential for molecular-dynamics simulations. *J. Comput. Chem.* **1990**, *11*, 700−733.
(32) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. Charmm - a program for macro-molecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **1983**, *4*, 187−217.
(33) Brooks, C. L.; Brunger, A.; Karplus, M. Active-site dynamics in protein molecules - a stochastic boundary molecular-dynamics approach. *Biopolymers* **1985**, *24*, 843−865.
(34) Cui, Q.; Elstner, M.; Kaxiras, E.; Frauenheim, T.; Karplus, M. A QM/MM implementation of the self-consistent charge density functional tight binding (SCC-DFTB) method. *J. Phys. Chem. B* **2001**, *105*, 569−585.
(35) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(36) 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.
(37) Neria, E.; Fischer, S.; Karplus, M. Simulation of activation free energies in molecular systems. *J. Chem. Phys.* **1996**, *105*, 1902−1921.
(38) Torrie, G. M.; Valleau, J. P. Monte-carlo free-energy estimates using non-boltzmann sampling - application to subcritical lennard-jones fluid. *Chem. Phys. Lett.* **1974**, *28*, 578−581.
(39) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. The weighted histogram analysis method for free-energy calculations on biomolecules 0.1. The method. *J. Comput. Chem.* **1992**, *13*, 1011−1021.
(40) Horowitz, S.; Yesselman, J. D.; Al-Hashimi, H. M.; Trievel, R. C. Direct Evidence for Methyl Group Coordination by Carbon-Oxygen Hydrogen Bonds in the Lysine Methyltransferase SET7/9. *J. Biol. Chem.* **2011**, *286*, 18658−18663.
(41) Couture, J. F.; Hauk, G.; Thompson, M. J.; Blackburn, G. M.; Trievel, R. C. Catalytic roles for carbon-oxygen hydrogen bonding in SET domain lysine methyltransferases. *J. Biol. Chem.* **2006**, *281*, 19280−19287.
(42) Guo, H.; Karplus, M. Solvent influence on the stability of the peptide hydrogen-bond - a supramolecular cooperative effect. *J. Phys. Chem.* **1994**, *98*, 7104−7105.
(43) Guo, H. B.; Beahm, R. F.; Guo, H. Stabilization and destabilization of the C^{δ}−H···O═C hydrogen bonds involving proline residues in helices. *J. Phys. Chem. B* **2004**, *108*, 18065−18072.
(44) Guo, H. B.; Gorin, A.; Guo, H. A peptide-linkage deletion procedure for estimate of energetic contributions of individual peptide groups in a complex environment: Application to parallel β-Sheets. *Interdiscip. Sci.* **2009**, *1*, 12−20.