

What is The Catalytic Mechanism of Enzymatic Histone N-Methyl Arginine Demethylation and Can it be Influenced by an External Electric Field?

Rajeev Ramanan,^{a,b,&} Sodiq O. Waheed,^{a,&} Christopher J. Schofield,^c and Christo Z. Christov^{a,*}

[a] Dr. R. Ramanan, S. O. Waheed, Dr. C. Z. Christov

Department of Chemistry, Michigan Technological University, Houghton, Michigan 49931, United States.

[b] Present address: Dr. R. Ramanan

Department of Chemistry, National Institute of Technology, Rourkela, Odisha 769001, India.

[c] Prof. Dr. C. J. Schofield

The Department of Chemistry and the Ineos Oxford Institute for Antimicrobial Research, The Chemistry Research Laboratory, Mansfield Road, University of Oxford, OX1 5JJ, United Kingdom.

[&]R.R. and S.O.W. made equal contributions to this article.

*Corresponding author: christov@mtu.edu

Abstract

Arginine methylation is an important mechanism of epigenetic regulation. Some Fe(II) and 2-oxoglutarate dependent Jumonji-C (JmjC) N ϵ -methyl lysine histone demethylases also have N-methyl arginine demethylase activity. We report combined molecular dynamic (MD) and Quantum Mechanical/Molecular Mechanical (QM/MM) studies on the mechanism of N-methyl arginine demethylation by human KDM4E and compare the results with those reported for N-methyl lysine demethylation by KDM4A. At the KDM4E active site, Glu191, Asn291, and Ser197 form a conserved scaffold that restricts substrate dynamics; substrate binding is also mediated by an out of active site hydrogen-bond between the substrate Ser1 and Tyr178. The calculations imply that in either C-H or N-H potential bond cleaving pathways for hydrogen atom transfer (HAT) during

N-methyl arginine demethylation, electron transfer occurs via a σ -channel; the transition state for the N-H pathway is ~ 10 kcal/mol higher than for the C-H pathway due to the higher bond dissociation energy of the N-H bond. The results of applying external electric fields (EEFs) reveal EEFs with positive field strengths parallel to the Fe=O bond have a significant barrier-lowering effect on the C-H pathway, by contrast, such EEFs inhibit the N-H activation rate. The overall results imply that KDM4 catalyzed N-methyl arginine demethylation and N-methyl lysine demethylation occur via similar C-H abstraction and rebound mechanisms leading to methyl group hydroxylation, though there are differences in the interactions leading to productive binding of intermediates.

Introduction

Epigenetic processes regulate normal eukaryotic development and dysregulation of them correlate with human diseases, including brain disorders and cancer.^[1,2] Methylation and demethylation of N-methylated lysine- and arginine-residues histone tails in nucleosomes are dynamic events that regulate transcription.^[1,2] The non-heme Fe(II) and 2-oxoglutarate (2OG)-dependent Jumonji-C (JmjC) oxygenases are the largest family of N^ε-methyl lysine histone demethylases (KDMs).^[3-5] In humans, there are 60-70 2OG-oxygenases, which have roles in processes including fatty acid metabolism, the hypoxic response, collagen biosynthesis, and translational regulation.^[6,7] In addition to the histone lysine demethylase (KDM) activities of some JmjC 2OG- oxygenases, recent in vitro studies have identified arginine demethylase (RDM) activities for some, but not all, JmjC KDMs, including for KDM3A, KDM4A, KDM4E, KDM5C, and KDM6B.^[8] Arginine methylation can modulate binding of histones with 'reader' domain containing proteins,^[9] so regulating transcription.^[10,11] Arginine methylation can also inhibit lysine methylation; it is, e.g.,

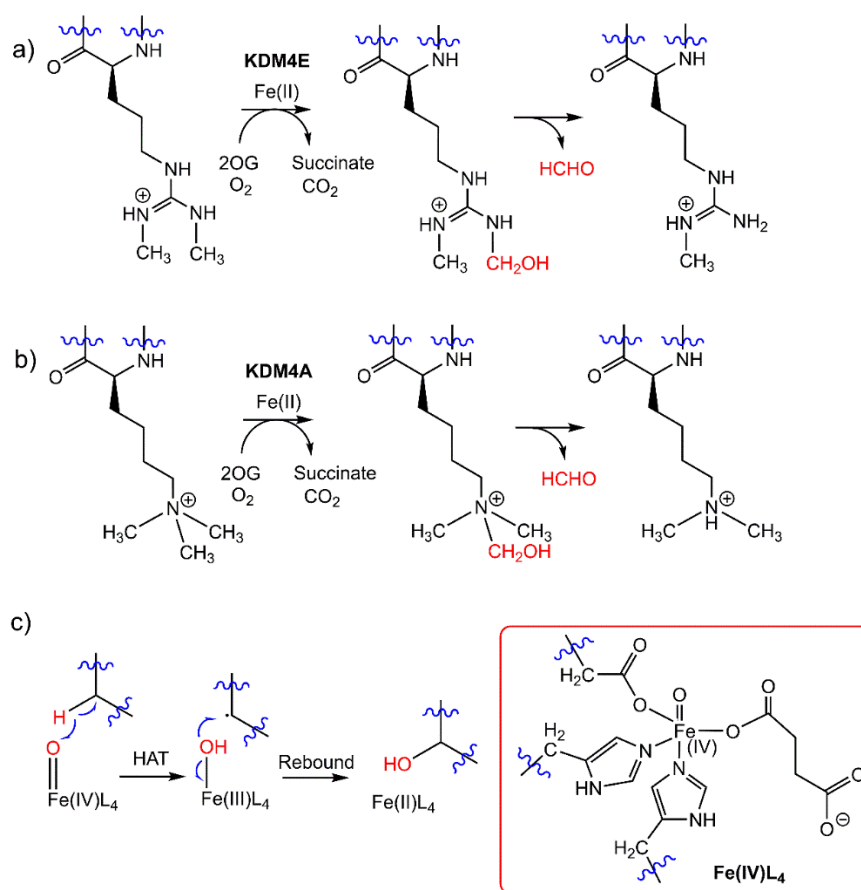
reported that the H3R2me2a mark hinders mixed-lineage leukemia 1 (MLL1) cells from methylating H3K4me3.^[10]

The mechanisms of dioxygen activation by non-heme Fe (II) enzymes have been explored using experimental and computational methods;^[2,12-14] however, there are no such studies reported on the recently reported JmjC KDM catalyzed histone arginine demethylation (RDM) reaction.^[15-17]

The 2OG oxygenases have evolved to activate O₂ to enable production of a ferryl intermediate that effects C-H bond cleavage, typically generating an alcohol product.^[7,12-14] Details of the mechanism leading to the ferryl intermediate and studies on model systems have been reported.^[15-26a] Following substrate binding, the consensus 2OG oxygenase mechanism^[7,12-26a-c] involves initial generation of a Fe(III)-superoxo complex; oxidative decarboxylation of 2OG generates a Fe(II)-peroxysuccinate intermediate which reacts to give a high-spin (S=2) state ferryl intermediate, which abstracts a hydrogen atom from the substrate. Hydrogen atom transfer (HAT) through a high spin surface is favored by spin exchange enhanced reactivity at the transition state,^[27,28a] generating an Fe(III)-hydroxyl intermediate that undergoes a low barrier hydroxyl rebound reaction with the substrate radical carbon to complete hydroxylation [Scheme 1]. Despite the apparent observance of the consensus mechanism by many of its members, the 2OG oxygenase family accepts both small and large molecule substrates, catalyzing a large range of two-electron oxidations with differences in regio- and stereo-selectivities.^[6,12]

KDM4E [Figure 1] is a JmjC 2OG oxygenase that in addition to KDM activity, at least in isolated form, also has RDM activity, including as shown by its demethylation of a histone H4 fragment with a symmetrically demethylated arginine at residue R3 (H4R3me2s).^[8] The mechanisms of histone lysine demethylation (KDM) by 2OG-demethylases have been recently explored,^[15,16,18] providing insights into their reaction mechanisms, alternative mechanistic strategies, residues that

stabilize the transition states, and long-range correlated motions involved in their catalysis. KDM4E is a relatively efficient arginine demethylase (RDM), while KDM4A is a relatively more efficient lysine demethylase (KDM).^[8] The mechanism of histone lysine demethylation catalyzed by KDM4A^[18] has been explored, but the previous study did not include the new RDM activity. Elucidating mechanistic differences between the RDM and KDM reactions may in the long term enable strategies for selective modulation of either the RDM or KDM activities. Here we report investigations on the differences in active site interactions in KDM4E and KDM4A that might in part account for the differences in their KDM and RDM selectivity.



Scheme 1. Outline mechanisms for arginine- and lysine-demethylation by the JmjC 2OG oxygenases. **a)** KDM4E catalyzed demethylation of a symmetrically N-dimethylated arginine residue. **b)** KDM4A catalyzed demethylation of an N-trimethylated lysine residue. **c)** Hydrogen

atom transfer (HAT) and rebound steps in the C-H bond cleavage step by the ferryl intermediate, coordination of which is shown in the inset.

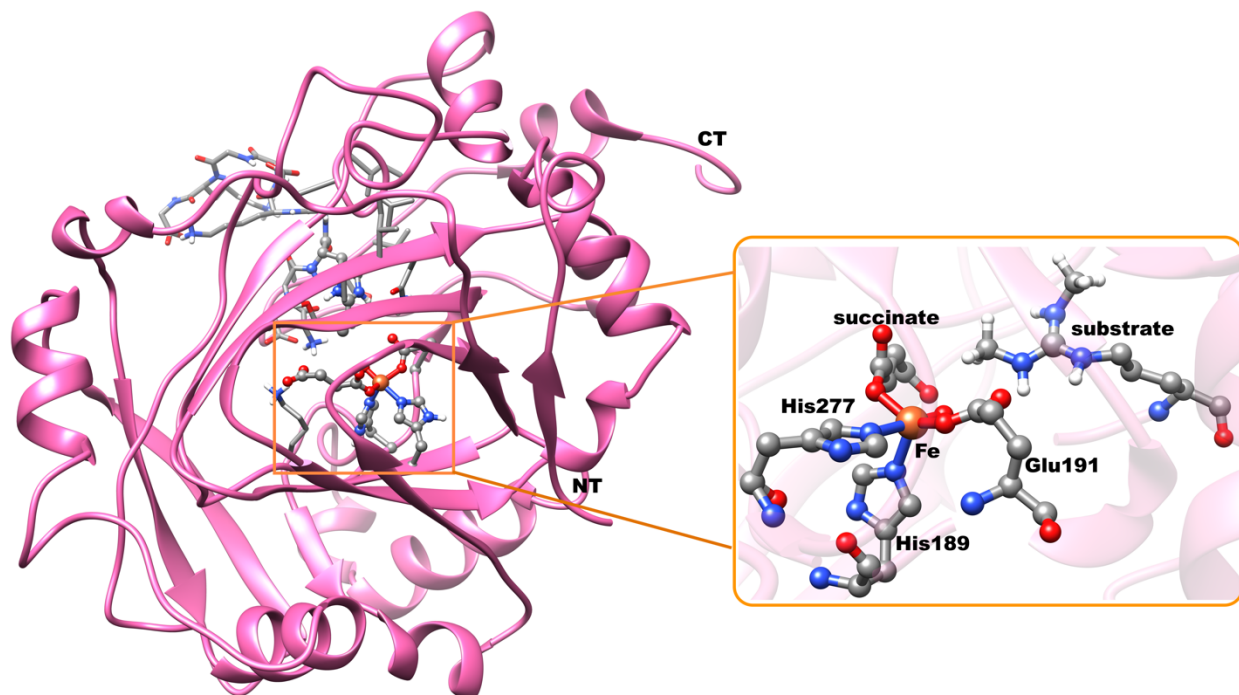


Figure 1. Model of KDM4E at the Fe(IV)=O intermediate state. The active site is shown in the inset. The QM region contains the imidazole groups of His189 and His277, the Fe and oxo atoms of Fe(IV)=O, the Glu191 sidechain, succinate, and the dimethyl guanidinium group of the substrate. NT and CT are the N- and C-termini, respectively.

We performed computational molecular dynamics (MD) and quantum mechanical and molecular mechanical (QM/MM) studies on substrate binding and mechanism of the KDM4E RDM activity. We investigated alternative mechanistic proposals, i.e., C-H or N-H hydrogen abstraction, and explored the effects of long-range interactions. We also explored the effects of applying an external electric field (EEF) on the reaction mechanism and product selectivity.^[28b-g] Comparison of the results for the RDM activity of KDM4E with the KDM activity of KDM4A imply both proceed

via ferryl mediated C-H abstraction, but that there are differences in the interactions leading to the formation of intermediates.

Methods

System Setup and MD Simulation: KDM4E is a human KDM4 family JmjC KDM that catalyzes histone H3 K9me3/2 N^ε-methyl lysine demethylation.^[29,30] The available crystal structure of KDM4E does not contain the histone substrate. However, KDM4A has been crystallized with a fragment of the H4 histone tail (H4(1-9)R3me2s) that contains an N-methylated arginine substrate (PDB ID: 5FWE).^[29] KDM4E and KDM4A are related (66% amino acid identity) with some key differences in the binding sites. Aiming to understand the catalytic mechanism and interactions in the more effective RDM enzyme, we chose to work with KDM4E and complemented the lack of an experimental enzyme-substrate structure of KDM4E by modeling.

The substrate H4(1-9)R3me2s coordinates from the KDM4A structure were superimposed on a KDM4E structure modeled at the ferryl oxidation state (based on a KDM4E crystal structure (PDB ID: 2w2i).^[31] The overlaid structures gave an RMSD of 0.651 Å [Figure S1]. The Ni in the KDM4E crystal structure was replaced with an Fe and the 2OG analogue pyridine-2,4-dicarboxylic acid (PD2) was manually modified to succinate [Figure S2]. Missing hydrogen atoms were added to the protein followed by neutralization with an appropriate number of counter ions were performed using the Leap module of Amber16.^[32] Antechamber as implemented in Amber was used to generate parameters for non-standard ligands /residues such as succinate and substrate. The Metal Center Parameter Builder (MCPB.py) program was used to generate parameters for the metal (Fe(IV)-oxo) intermediate.^[33a] The high spin state of Fe (S=2, M=5) was assigned to the ground

state, based on previous studies with KDM4A, KDM7B(PHF8), AlkB, AlkBH2 and other 2OG oxygenases.^[16-18,33b-c] The active site Fe is ligated by a monodentate chelating succinate, two histidine residues (His189 and His277) and one glutamic acid, Glu191 (monodentate) [Figure S3], as revealed by crystallographic studies of KDM4E.^[29,31]

Note that in addition to the iron binding, active site, the KDM4 KDMs contain a zinc binding site (Zn binding residues are Cys234, His240, Cys306, and Cys308 in KDM4A) as observed crystallographically.^[29] In the case of the KDM4E crystal structure (PDB 2W2I) on which our modeling studies were based, the zinc binding site was not apparent, hence was not included in our studies, which focused on the catalytic importance of the interactions in the active site.

MD simulations were performed with the FF14SB force field.^[34] The protein was fully immersed in a rectangular box made of TIP3P water molecules.^[35] The edge of the box was kept 10 Å away from the protein surface. Periodic boundary conditions were used for MD simulations. Long-range electrostatic interactions were calculated using the Particle Mesh Ewald (PME) method with a direct space and vdW cut-off of 10 Å.^[36] Initial energy minimizations were performed in two steps, i.e., first the steepest descent for 5000 steps, followed by the conjugate gradient method for another 5000 steps. Solute molecules were restrained using a potential of 500 kcal/mol Å² in the first minimization where only the solvent and counterions were allowed to minimize. A full minimization of the entire protein and solvents with steepest descent (5000 steps) and conjugate gradient (5000 steps) was carried out. The CPU version of the Amber16 code (SANDER) was used for energy minimization. The temperature of the system was gradually increased from 0 to 300 K in an NVT ensemble for 100 ps. The system was subsequently run for 1ns in an NPT ensemble. The temperature was kept constant at 300 K and pressure at 1.0 atm, using Langevin-thermostat^[37] and Barendsen barostat^[38] with a pressure relaxation time of 1 ps. Before moving to

the final productive simulations, equilibration for ~3 ns was performed. The SHAKE algorithm was used to constrain all bonds with hydrogen.^[39] Production MD runs were performed for 1000 ns in an NPT ensemble with a target pressure set at 1 bar and a constant pressure coupling of 2ps. Two other replicas simulations were performed at different initial velocities and the production runs were done for 300 ns [Figure S4]. The GPU version of the Amber16 code (PMEMD) was used for the final productive molecular dynamics simulations.^[32]

QM/MM Calculations: The quantum mechanical (QM) region in the hybrid quantum mechanical/molecular mechanics (QM/MM) calculations included the Fe, the ferryl O, the imidazole sidechains of both coordinating histidines, the glutamate sidechain from the gamma carbon, succinate, and the substrate residue (methylated arginine H4(1-9)R3me2s). The QM region contains a total of 60 atoms; 4 linked hydrogen atoms were used to complete the valences of bonds spanning between the QM and the MM regions. The ChemShell suite of programs was used for QM/MM calculations.^[40,41] The QM calculations were performed with Turbomole,^[42] and MM calculations with DL_POLY software.^[43] The Amber force fields generated for the MD simulation were used in DL_POLY. The electronic embedding scheme was employed to calculate the polarizing effect of the enzyme on the QM region.^[44] The B3LYP^[45] functional with the def2-SVP basis set (B1) was used for all geometry optimizations, which can yield accurate barriers and proper electronic structure for Fe metal containing systems.^[16-18,46] As in previous studies, transition states were defined by two step calculations. At first, a relaxed potential energy surface (adiabatic mapping) scan was performed by gradually changing the reaction coordinate.^[16-18] The highest energy point on the scanned surface was used as an initial guess for the final TS optimization. The DL-find optimizer implemented in Chemshell was used for reactant optimization and scanning,^[47] while the HDLC optimizer was used for the final TS optimization.^[48]

The local minimum and the first-order saddle points were verified with frequency calculations. The B3LYP/def2-TZVP level of theory (B2) single point calculations were used to refine the final energies.^[16-18] The Grimme's D3 dispersion correction^[49a] is applied in single-point calculations and the energy profile is presented in the supporting information. The external electric fields (EEFs) were applied for the QM/MM calculations using TITAN code^[49b] as in other studies.^[50,51] The TITAN code was used to generate uniform electric fields on the entire model of the enzyme used for the QM/MM calculations [Figures S5-S6]. To achieve this, two circular parallel charge plates were generated, consisting of 7082 point charges in total with 3541 positive charges and 3541 negative charges. The distance between the center of the plate and the Fe atom is 46.8 Å and the radius of the plate is 92.4 Å. The solvated enzyme-substrate complex with the added point charges were subjected to QM/MM reaction path calculations using ChemShell, accounting for the polarization effect of the added new point charges on the QM Hamiltonian in addition to the already incorporated effect of the point charges of the MM part (protein and water atoms).

Results and Discussion

Dynamics and Substrate Binding in the Histone H4 Tail with N-Methylated Arginine by KDM4E

To investigate KDM4E substrate interactions, we performed a 1 μ s MD simulation and two other replicas simulations for 300 ns each at different initial velocities of the KDM4E.H4(1-9)R3me2s complex at the ferryl intermediate stage [Figure S4], the primary coordination sphere of which is given in Scheme 1. The replicas showed very similar RMSD profiles to the initial simulation. As with the KDM activity of KDM4A, with KDM4E, the Fe(IV)-oxo is coordinated by two histidines (His189 and His277), a glutamate (Glu191), and a succinate oxygen, in each case in a monodentate manner. Also, similarly to KDM4A, the non-coordinating 'C-1' oxygen of succinate is stabilized

by interactions with Tyr133/Lys207; Phe186 makes CH- π interactions with the succinate methylenes. The non-coordinating oxygen of Glu191 is stabilized by a hydrogen bond with Asn291, which in turn is hydrogen bonded with Ser197. Stabilization of succinate and the non-coordinating oxygens of Fe-ligating Glu- or Asp residues is a common feature in 2OG oxygenases. For example, as for KDM4E, with the JmjC KDMs PHF8^[30] and KDM4A,^[52] the non-coordinating oxygen of the Fe-ligating Asp (in PHF8)/Glu (in KDM4A) is stabilized by hydrogen bonding with an Asn residue, while in AlkB and AlkBH2 an Arg residue is employed to do this [Figures S7-S9].^[53-56] The non-coordinating succinate carboxylate is positioned to interact with Lys207, Tyr133, again similarly to the analogous intermediate in KDM catalysis by KDM4A. Overall, the coordinating environment of the iron center employed in the RDM reaction of KDM4E appears to be very similar to that of the KDM reaction of KDM4A.

The symmetrically dimethylated arginine (R3me2s) of the H4 histone tail binds deeply within the KDM4E active site, as observed in structures of KDM4A complexed with substrates including H4(1-9)R3me2s.^[29] The MD simulations [Figure 2b] with KDM4E imply the crystallographically observed binding mode is stable. After the first 100ns of simulation, the symmetrically substituted N-methyl arginine groups occupy distinct regions, with one close to the ferryl oxygen and one directed away from it [Figure 2a and b].

Note that although irrelevant in the case of the symmetrically substituted N-methyl arginine substrate studied here, previous reports show KDM4E, like some other JmjC KDMs, can oxidize ‘asymmetrically substituted’ substrate analogues, e.g., Lys derivatized with N ϵ Me and isopropyl groups, raising issues of chemo- and regio-selectivity (KDM4E selective hydroxylation of the isopropyl occurs).^[57] In such cases, the preferred binding mode of the different alkyl groups likely contributes to the observed regioselectivity (along with relative C-H bond strengths).^[57]

The KDM4E residues preceding Asn291, i.e., Ile290, Ala289, appear to form a rim around the dimethylguanidium group of dimethylated arginine substrate [Figure 3] and likely contribute to the rigidity of the active site restricting substrate arginine side chain movement [Figure 2b]. Lys242, Tyr178, and Asn291 are second sphere residues that help orient the symmetrically substituted N-methyl arginine substrate sidechain in the active site [Figure 3].

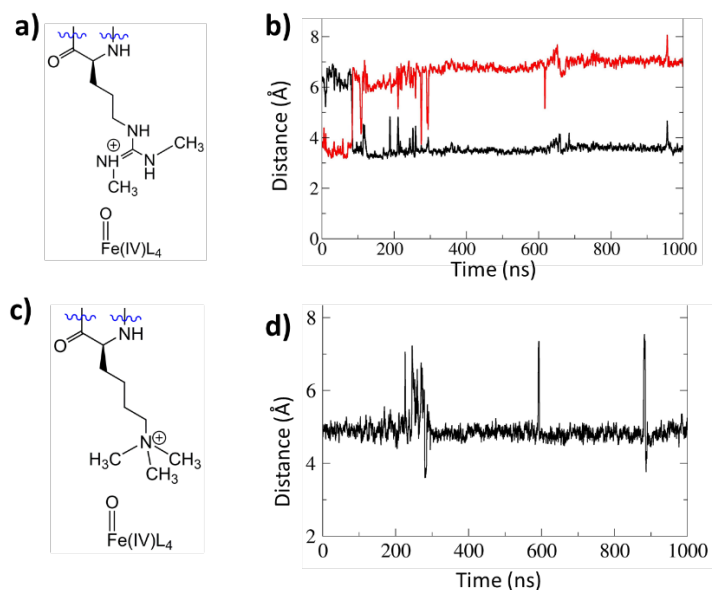


Figure 2. Mechanism of JmjC catalyzed N-methyl arginine demethylation. **a)** The symmetrically dimethylated arginine is positioned with one methyl group adjacent to the ferryl oxygen. **b)** Dynamics of the arginine N-methyl groups at the active site of KDM4E. Distances of the two N-methyl carbon (red and black) to the Fe(IV)-oxygen. **c)** The N ϵ -trimethylated lysine and the ferryl-oxygen. **d)** Distance of the substrate N ϵ -trimethylated N to the ferryl-oxygen in KDM4A during a 1 μ s simulation.^[18]

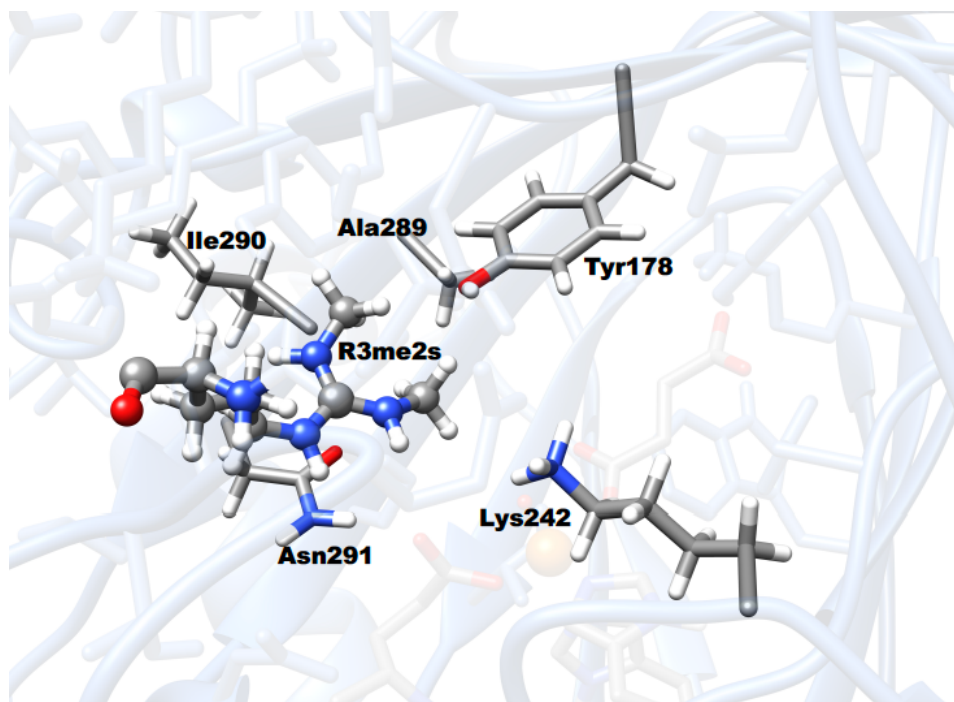


Figure 3. Residues that stabilize the symmetrical dimethylated Arginine (R3me2s) of the H4R3me2s substrate in the KDM4E active site.

Reaction Mechanism of Demethylation of Methylated Arginine in H4 by KDM4E

We then modeled reaction pathways for the KDM4E catalyzed RDM reaction, considering both initial reaction of a methyl group C-H bond or reaction of the N-H bond of the N(H)-methyl group. The distance plots in Figure 2b imply that the active site ensures that one of the methyl groups is always close to the ferryl group. The N-H bond is only rarely close to the ferryl group implying a conformational preference for cleavage of the methyl C-H bond. The C-H activation possibility was modeled from a snapshot where a methyl hydrogen is close to the ferryl oxygen, i.e., at 864ns where there is an O-CH hydrogen distance of 2.55Å. The N-H activation possibility was modeled from an enzyme-substrate complex at 916ns with an initial O-NH hydrogen distance of 2.09Å. The results were tested using another snapshot (945ns) where the substrate N-H and C-H hydrogens are positioned near equidistantly to the ferryl oxygen, i.e., 3.03 and 2.65 Å, respectively.

These structures were chosen in order to investigate factors governing a preference towards C-H or N-H HAT and in particular whether appropriate positioning of either C-H/ or N-H with respect to Fe(IV)=O bond or the difference in the bond strength between C-H and N-H is most important. The analysis of the MD trajectories for the selected snapshot at 864ns used for C-H activation reveals that the HAT distance and angle geometric parameters are present in 45.3% and 6.4% of all MD snapshots, respectively. In the snapshot used for N-H activation (916ns), the geometric parameters are present in 22.2% and 32.5% from the trajectories, respectively. In the snapshot selected at 945ns, where both C-H and N-H possibilities were tested, the HAT distance and angle geometric parameters occurred in 45.3% and 22% from the MD trajectories, respectively, in the C-H pathway while in the N-H pathway, the geometric parameters are present in 22.4% and 18.1% from all MD snapshots, respectively.

The potential energy profile for CH activation [Figure 4b] in KDM4E shows that the reaction proceeds by the generally accepted mechanism of HAT, followed by a rebound step, as occurs for the KDM activity of vs. KDM4A [Figure 4d] [Scheme 1c]. The barrier for the HAT during the RDM activity of KDM4E is 16.3 kcal/mol at the QM(B2+ZPE)/MM level of theory and 17.5 kcal/mol at the QM(B1)/MM level of theory. The energy barrier for the HAT of arginine demethylation in KDM4E is lower than the barrier for lysine demethylation in KDM4A^[18] by 5.3 kcal/mol, which suggests that the more effective catalytic KDM performance of KDM4A is due to events prior to C-H bond cleavage.

The transition state geometry for CH activation in KDM4E is given in Figure 5a. The Fe-O distance is 1.74 Å, which is elongated from the 1.62 Å of the starting **IM1(CH)** ferryl complex and the Fe-O-H angle is 157.2°. The Mulliken spin density at the carbon radical center of the substrate is -0.9, which implies a σ electron transfer process. The spin natural orbital (SNO) of the

KDM4E HAT transition state [Figure 6] shows electron transfer to the σ^*z^2 orbital of the Fe metal, supporting this proposal.

The optimized reactant geometry showed similar interactions as in the MD simulation, with the stabilization of succinate and non-coordinating oxygen of glutamate intact. Glu191, Asn291, and Ser197 bring rigidity to the active site by making a network of hydrogen bonds. The evolution of the hydrogen bonds involving Glu191, Asn291, Ser197, Tyr178 and Ser1 H4 substrate during MD simulation is shown in Figure 4a. Long MD (i.e. 1000 ns) studies reveal that the Ser1 H4 substrate residue makes strong hydrogen bonding stabilization with Tyr178. Overall, the active site interactions involving Glu191, Asn291, and out of the active site hydrogen-bonded stabilization from Ser1(substrate) with Tyr178 help to productively orientate the N-methylated arginine side chain in the KDM4E active site. Generation of the hydrogen bonding network Glu191, Asn291, and H4Ser1 distinguishes KDM4E RDM catalysis from the KDM activities of KDM4A or PHF8 [Figure S7].^[31] In KDM4A and PHF8, steric repulsions from the active site residues contribute to the productive orientation of the substrate. For example, in KDM4A, Lys241 (KDM4A), Tyr177 (KDM4A), and Asn290 (KDM4A), and or in PHF8, Ile191 (PHF8), Arg460 (PHF8) and Phe250 (PHF8) provide the steric repulsion to generate a tailor-made active site for efficient catalysis.

Figures 5a and b show the interactions that stabilize the transition states during the CH hydrogen abstractions in KDM4E (RDM catalysis) and KDM4A (KDM catalysis), respectively. With KDM4E, the non-coordinating oxygen of the coordinating Glu191 is stabilized by hydrogen bonding with Asn291. Networks of hydrogen bonding interactions of second sphere residues, Tyr178, Asp136, H4Ser1 (substrate), and Glu311 enhance the stability of the TS. In KDM4A, the non-coordinating oxygens of both the Fe ligating Glu190 and the succinate are stabilized via hydrogen bonding interactions with Asn290 and Asn198, respectively.^[18] Binding of the non-polar

part of the H3K9me3 substrate in KDM4A is stabilized by the phenyl ring of Tyr175. Overall, the H4R3me2s substrate in KDM4E appears to be more precisely stabilized than the H3K9me3 substrate in KDM4A.^[18]

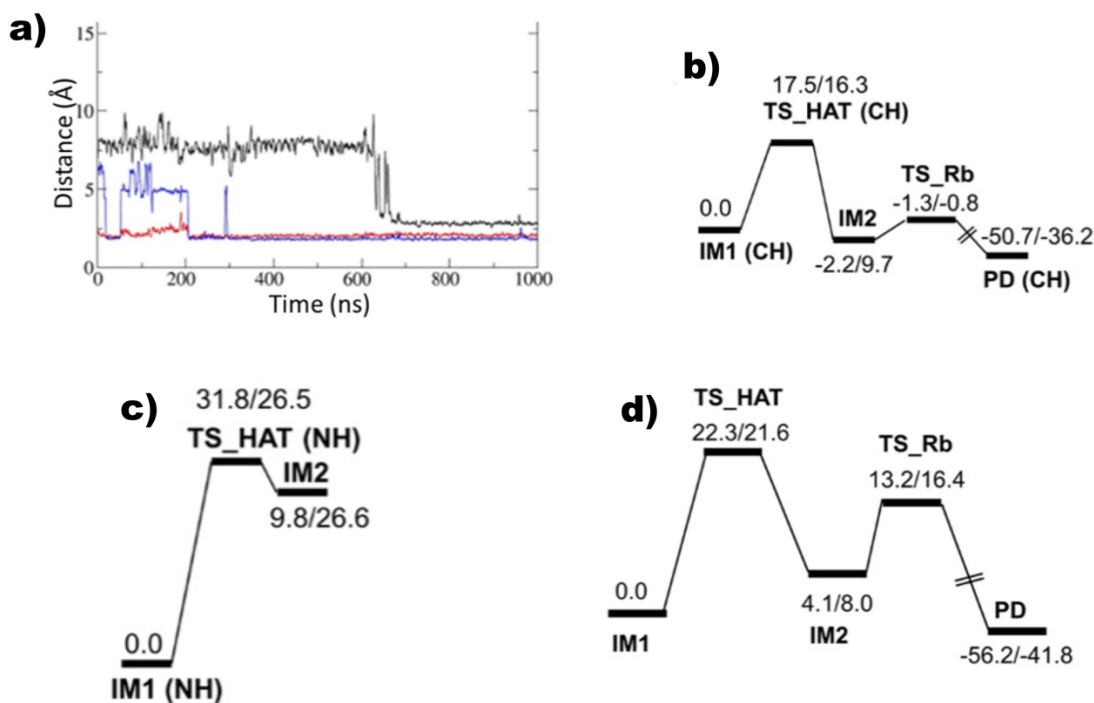


Figure 4. Distance plots for Glu191-Asn291 in red; Ser197-Asn291 in blue; and Tyr178-Ser1 in black for the RDM reaction of KDM4E **(a)**. Potential energy profile for C-H activation **(b)** and N-H activation for the RDM activity of KDM4E **(c)**. Potential energy profile for HAT and the rebound steps during KDM4A KDM catalysis **(d)**.^[18] Energies are given in kcal/mol at the QM(B1)/MM followed by QM(B2+ZPE)/MM level of theory.

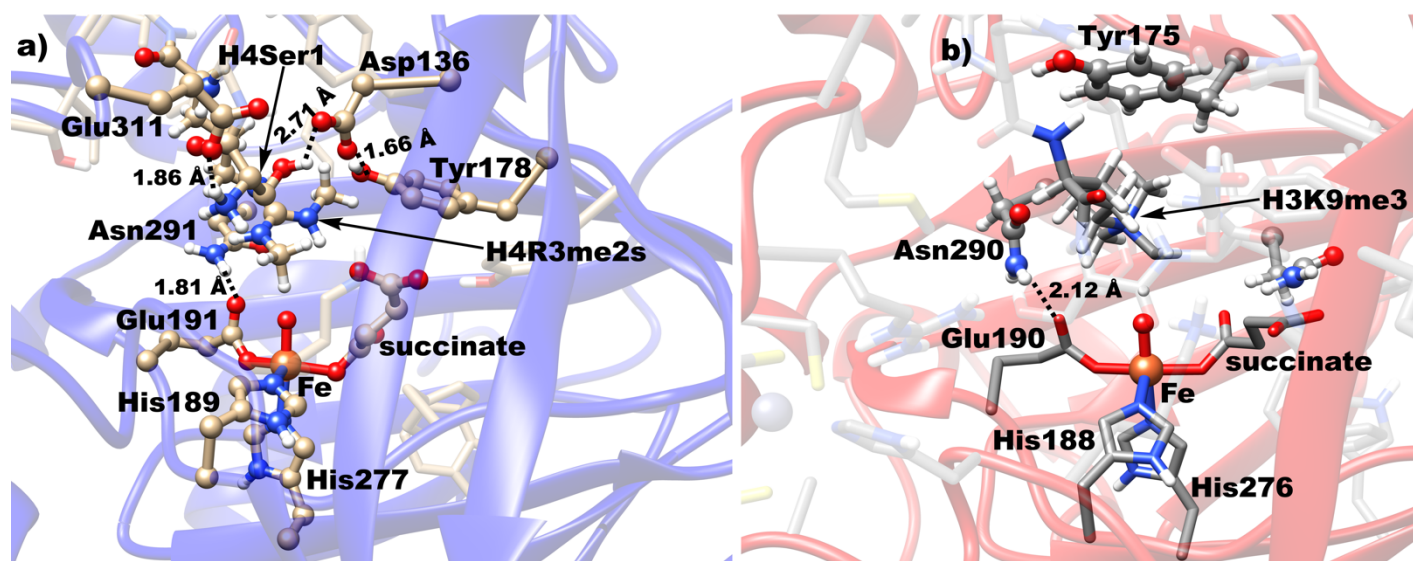


Figure 5. The optimized transition state geometry for C-H hydrogen abstractions in **(a)** KDM4E and **(b)** KDM4A with their respective stabilizing residues.

We explored the possibility for HAT of the NH hydrogen as this has been proposed in model studies of KDMs.^[25b] The barrier for NH hydrogen abstraction is higher than for CH abstraction, being 26.5 kcal/mol at the QM(B2+ZPE)/MM level of theory and 31.8 kcal/mol at the QM(B2)/MM level of theory [Figure 4c]. The DFT-D3 dispersion corrected energies at QM(B2+ZPE)/MM level of theory show the same trend and also support the preference of HAT via C-H over N-H [Figure S10]. The potential NH activation mechanism proceeds via a transition state where the Fe-O bond is elongated to 1.69 Å, with the O—H and N----H distances being 1.19 and 1.34 Å, respectively. The Fe-O-H angle is 142.5°, suggesting a possible σ trajectory where the electron is shifted to the σ^*z^2 orbital. The electronic structure shows an electron density of -0.03 at the N, suggesting the alpha electron of the N-H σ bond shifts to the Fe σ^*z^2 orbital. Thus, both the -CH activation and -NH activation transition states have a similar electronic structure. The higher barrier for N-H bond activation over C-H bond activation is explained on the basis of the

higher bond dissociation energy (BDE) for the N-H bond, making it less reactive as it was demonstrated in the case of the KDM activity of KDM7B (PHF8).^[16] The presence of an alpha amino group significantly reduces the C-H bond dissociation energy. The C-H BDEs alpha to an amino group has a value around 91 kcal/mol while the N-H has ~100 kcal/mol.^[58]

The rebound steps in both the KDM4A KDM and KDM4E RDM reactions are lower in energy compared to the HAT step [Figures 4b and 4d]. The rebound reaction in the KDM4E RDM reaction has a lower barrier compared to the KDM4A KDM reaction. One of the oxygens of the non-coordinating carboxylate of the succinate in KDM4E makes a stabilizing interaction with the C-H hydrogen of the methylene radical at the rebound TS [FigureS11]. Such a stabilizing interaction is absent in the rebound TS in KDM4A KDM catalysis [Figure S12]. The optimized stationary point geometries along the reaction mechanism in KDM4E are presented in Figure S13.

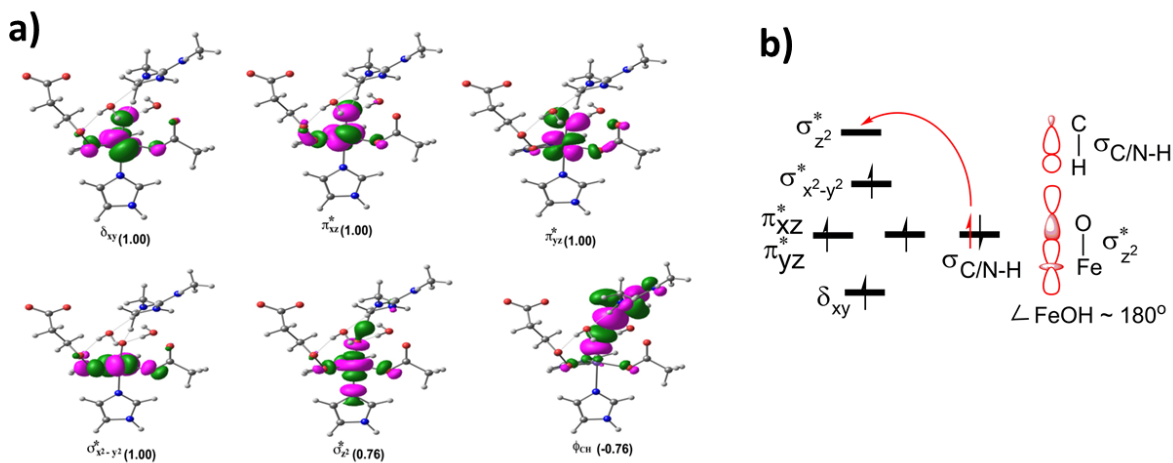


Figure 6. Electronic structure details of the HAT in KDM4E. **a)** Spin natural orbitals (SNO) with their respective orbital occupancies (in parentheses) for the CH hydrogen atom abstraction transition state mechanism. **b)** The electron shift diagram for a σ -trajectory of HAT.

Effect of an oriented External Electric field (EEF) on the C-H and N-H Hydrogen Abstraction Reactions During N-methyl Arginine Demethylation

To further investigate the possibility of hydrogen atom abstraction via C-H or N-H abstraction pathways for the RDM reaction we applied external electric fields (EEFs) to the KDM4E Fe(IV)=O complex. EEFs have the ability to control catalysis/inhibition of reactions and their selectivity.^[28b-d] We oriented EEFs both parallel and orthogonal to the Fe=O bond (i.e., reaction axis) in the Fe(IV)=O complex and then performed the HAT on the optimized ferryl complex geometries. The orthogonal direction is defined along the Fe-N^ε (His189) bond, which is perpendicular to Fe=O bond [Figure S6]. In the orthogonal direction, electric field strengths of ± 0.0025 au were applied [Table 1], resulting in barriers of 20.1 and 36.9 kcal/mol for HAT via C-H and N-H pathways, respectively at an electric field strength of -0.0025 au. When the electric field strength was changed to +0.0025 au, the obtained barriers are 17.4 and 40.8 kcal/mol for HAT via C-H and N-H, respectively. These barriers are higher than the 16.3 and 26.5 kcal/mol values calculated at B2+ZPE level of theory, obtained for HAT via C-H and N-H pathways, respectively, without an external electric field. Pronounced EEFs effects were observed for N-H activation as the barriers increased by 10.4 and 14.3 kcal/mol with an electric field strength of -0.0025 or +0.0025 au, respectively, compared with the values without EEFs. These results suggest that an applied external electric field with either a positive or negative field strengths in an orthogonal direction to the reaction axis inhibit (to a different extents) the rate of hydrogen atom abstraction in both C-H and N-H HAT pathways, thereby slowing the rate of demethylation. The EEF in an orthogonal direction to Fe=O bond, however, exercises a much stronger effect on the barrier of the N-H HAT pathway.

We then applied EEFs of strength ± 0.0025 au in the parallel direction to the Fe=O bond. The results [Figure 7, Table 2] show that at -0.0025 au field strength, the barriers increase for both the C-H and N-H HAT pathways, compared to the absence of an EEF. However, with a $+0.0025$ au field strength, a slightly lower barrier was observed than the one obtained without EEFs for HAT via C-H abstraction, while the barrier for N-H abstraction increases with almost similar magnitude as observed when EEFs were oriented orthogonal to the Fe=O bond.

Due to variations in the results, we doubled (± 0.0050 au) and tripled (± 0.0075 au) the magnitude of the EEF applied parallel to the Fe=O bond. The results [Figure 7, Table 2, at B2+ZPE level] reveal that for C-H activation with -0.0050 and -0.0075 au field strengths, the barrier increases by 5.5 and 5.6 kcal/mol, respectively; for N-H activation, the barrier increase by 11.1 and 7.6 kcal/mol, respectively, compared with the values without EEFs. However, with analogous positive field strengths, the barriers for C-H activation decrease by 2.5 and 4.2 kcal/mol for the $+0.0050$ and $+0.0075$ au field strengths, respectively; for the N-H activation, the barriers increase by 12.2 kcal/mol and 13.9 kcal/mol for the $+0.0050$ and $+0.0075$ au field strengths, respectively. The EEF results show that negative fields parallel to the reaction axis increase the barriers for both C-H and N-H HAT. However, the positive parallel fields have a significant barrier-lowering effect on C-H HAT, while for N-H HAT, they slow reaction. Overall, the results indicate that applying positive EEFs parallel to the reaction axis enhances the rate of C-H activation in a field strength dependent manner. By contrast, application of either positive or negative EEFs parallel to the reaction axis inhibits the rate of N-H activation, likely reflecting the preference for C-H activation over N-H activation.

To investigate the results, we analyzed the spin densities of the optimized Fe(IV)=O complexes. The results [Tables 1 and 2] reveal that in the C-H activation pathway where positive EEFs were

applied parallel to the reaction axis, the oxo group of the Fe(IV)=O is more polar than in the Fe(IV)=O complex without an EEF; increasing the magnitude of the applied EEFs increases such polarity, thereby increasing the rate of the HAT. However, in the other tested systems, the oxo group becomes less polar compared to the respective ferryl complex without an EEF, thus slowing HAT. The effects of the EEF are different for the C-H (decrease of the barrier for a positive EEF) and N-H (increase of the barrier in all cases) HAT processes with the effect being more profound for the former [Figure 4]. The observations indicate that after applying the EEF and optimizing the RCs structures, there are changes in the local interactions that are responsible for the observed differences.

With the EEFs, the HAT proceeds (for both C-H and N-H HAT) via an σ trajectory where an α electron is transferred from the substrate to the vacant d-orbital of the Fe in the HAT transition state in a similar manner as without EEFs, indicating that the electron transfer mechanism is preserved with or without EEFs.

Table 1. Energy barriers and spin densities for the C-H and N-H activations with EEFs oriented orthogonal to the Fe(IV)=O bond reaction axis, calculated at the B1 and B2+ZPE levels of theory.

Electric field strength (au)	Barrier for C-H (kcal/mol)	Barrier for N-H (kcal/mol)	Difference (kcal/mol)	Spin density of Fe of Fe(IV)=O in C-H	Spin density of O of Fe(IV)=O in C-H	Spin density of Fe of Fe(IV)=O in N-H	Spin density of O of Fe(IV)=O in N-H
-0.0025	21.7/20.1	40.3/36.9	18.6/16.8	3.145	0.551	3.149	0.603
0.0000	17.5/16.3	31.8/26.5	14.3/10.2	3.136	0.582	3.145	0.619
0.0025	19.9/17.4	46.1/40.8	26.2/23.4	3.139	0.575	3.153	0.592

Table 2. Energy barriers and spin densities for the C-H and N-H activations with EEFs oriented parallel to the Fe(IV)=O bond reaction axis, calculated at the B1 and B2+ZPE levels of theory.

Electric field strength (au)	Barrier for C-H (kcal/mol)	Barrier for N-H (kcal/mol)	Difference (kcal/mol)	Spin density of Fe of Fe(IV)=O in C-H	Spin density of O of Fe(IV)=O in C-H	Spin density of Fe of Fe(IV)=O in N-H	Spin density of O of Fe(IV)=O in N-H
-0.0075	24.2/21.9	42.2/34.1	18.0/12.2	3.181	0.543	3.149	0.605
-0.0050	23.1/21.8	44.1/37.6	21.0/15.8	3.164	0.557	3.156	0.597
-0.0025	18.6/18.8	44.8/39.1	26.2/20.3	3.152	0.567	3.169	0.580
0.0000	17.5/16.3	31.8/26.5	14.3/10.2	3.136	0.582	3.145	0.619
0.0025	16.8/16.0	43.9/36.4	27.1/20.4	3.126	0.590	3.152	0.601
0.0050	14.5/13.8	42.6/38.7	28.1/25.2	3.114	0.601	3.164	0.587
0.0075	12.4/12.1	45.6/40.4	33.2/28.3	3.100	0.615	3.173	0.572

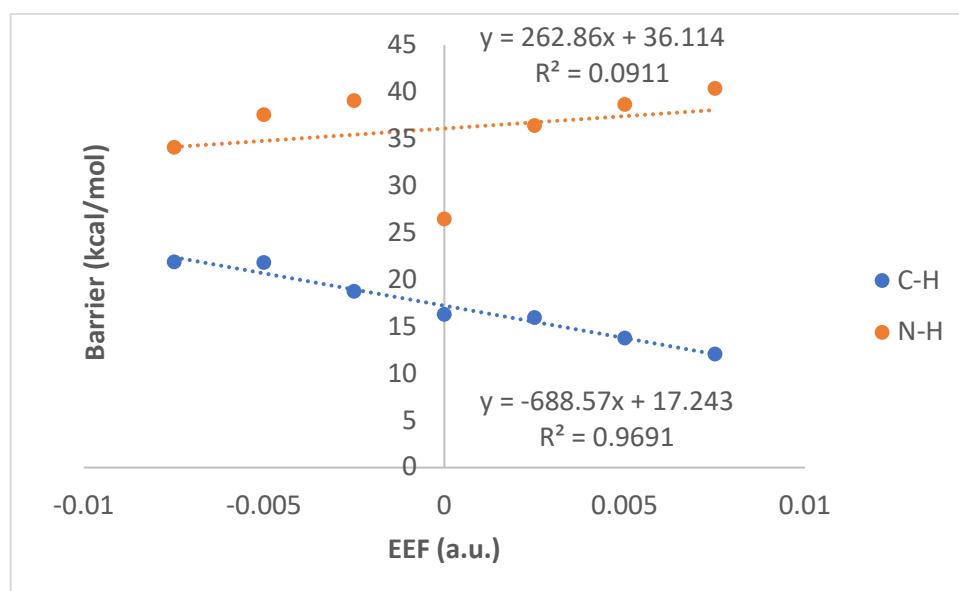


Figure 7. The relationship between the C-H and N-H pathways H-abstraction barriers and the EEF strength applied parallel to the Fe=O bond during KDM4E N-methyl arginine demethylation.

The presence of an EEF has been demonstrated to influence the spin densities and ultimately the rate of HAT reactions,^[28b,d-e,49] accelerate the rates of photochemical process in polar solvents,^[28f] and to change the reactivity of heme-containing enzymes.^[28b,e,49] There are no previous reports of studies of EEF effects on non-heme Fe(II) enzymes. We calculated the internal field from the protein scaffold without any applied EEF and used it as a background for the reaction path calculations with different values of the EEF. Our results provide evidence that altering the EEF increases the rate of hydrogen atom abstraction via C-H pathway, while inhibiting the rate of the N-H HAT pathway.

Conclusions

Our studies with MD and QM/MM calculations reveal the crucial role of a hydrogen-bonded scaffold at the active site of KDM4E for efficient RDM activity. Both interactions within and out of the active site are important in defining substrate selectivity and efficiency. Active site interactions of KDM4E by Glu191, Asn291 and Ser197 and an out of active site hydrogen-bond between Ser1(substrate) and Tyr178 collectively work to productively orientate the N-methylated arginine side chain in the active site. Despite the similarity of the different methylated substrates (histone tails with positively charged N-methylated arginine or N-methylated lysine-residues), there are differences in the interactions that bind the substrates, which stabilize the transition state, hence which might contribute to specificity and reaction rates. Importantly, such differences include not only the first sphere and the active site interactions, but also second sphere and more distant interactions.

We also show that an applied EEF can enhance the specificity of a non-heme oxygenase towards a particular reaction pathway. Thus, negative fields parallel to the reaction axis increase the barriers for both C-H and N-H activation processes during the KDM4E RDM reaction, implying that applying these fields in the stated direction inhibits both pathways. However, the positive fields have a significant barrier-lowering effect on C-H activation, while in N-H activation, they slow the reaction. The results suggest that further studies on the effects of EEFs on catalysis by non-heme Fe oxygenases are of interest.

Acknowledgments

C.Z.C. acknowledges NSF grant 1904215 (testing the methodology). C.Z.C and C.J.S. also thank NIH grant 1 R15 GM139118-01A1. This research was supported by Michigan Technological University Graduate Teaching Assistantship to S.O.W. C.J.S thanks Cancer Research UK for funding.

Conflicts of interest

The authors declare no competing financial interests.

References

- [1] A. T. Hauser, D. Robaa, M. Jung, *Curr. Opin. Chem. Biol.* **2018**, 45, 73–85.
- [2] J. R. Horton, M. Gale, Q. Yan, X. Cheng, *Cancer Drug Discov. Dev.* **2017**, 151–219.
- [3] Y. I. Tsukada, J. Fang, H. Erdjument-Bromage, M. E. Warren, C. H. Borchers, P. Tempst, Y. Zhang, *Nature* **2006**, 439, 811–816.
- [4] K. Yamane, C. Toumazou, Y-I. Tsukada, H. Erdjument-Bromage, P. Tempst, J. Wong, Y. Zhang, *Cell* **2006**, 125, 483–495.
- [5] J. R. Whetstine, A. Nottke, F. Lan, M. Huarte, S. Smolikov, Z. Chen, E. Spooner, E. Li, G. Zhang, M. Colaiacovo, Y. Shi, *Cell* **2006**, 125, 467–481.

- [6] R. P. Hausinger in *2-oxoglutarate-Dependent Oxygenases* (Eds.: C. J. Schofield, R. P. Hausinger), Royal Society of Chemistry, Cambridge, **2015**, pp. 1-58.
- [7] S. Martinez, R. P. Hausinger, *J. Biol. Chem.* **2015**, 290, 20702–20711.
- [8] L. J. Walport, R. J. Hopkinson, R. Chowdhury, R. Schiller, W. Ge, A. Kawamura, C. J. Schofield, *Nat. Commun.* **2016**, 7, 11974.
- [9] A. N. Iberg, A. Espejo, D. Cheng, D. Kim, J. Michaud-Levesque, S. Richard, M. T. Bedford, *J. Biol. Chem.* **2008**, 283, 3006-3010.
- [10] A. Di Lorenzo, M. T. Bedford, *FEBS Lett.* **2011**, 585, 2024-2031.
- [11] D. Hyllus, C. Stein, K. Schnabel, E. Schiltz, A. Imhof, Y. Dou, J. Hsieh, U. M. Bauer, *Genes Dev.* **2007**, 21, 3369-3380.
- [12] E. I. Solomon, S. Goudarzi, K. D. Sutherlin, *Biochem.* **2016**, 55, 6363–6374.
- [13] M. Costas, M. P. Mehn, M. P. Jensen, L. Que, *Chem. Rev.* **2004**, 104, 939–986.
- [14] M. R. A. Blomberg, T. Borowski, F. Himo, R. Z. Liao, P. E. M. Siegbahn, *Chem. Rev.* **2014**, 114, 3601–3658.
- [15] S. S. Chaturvedi, R. Ramanan, S. O. Waheed, J. Ainsley, M. P. Evison, J. M. Ames, C. J. Schofield, T. G. Karabancheva-Christova, C. Z. Christov, *Chem. Eur. J.* **2019**, 20, 5422-5426.
- [16] S. S. Chaturvedi, R. Ramanan, N. Lehnert, C. J. Schofield, T. G. Karabancheva-Christova, C. Z. Christov, *ACS Catal.* **2020**, 10, 1195-1209.
- [17] S. O. Waheed, R. Ramanan, S. S. Chaturvedi, N. Lehnert, C. J. Schofield, C. Z. Christov, T. G. Karabancheva-Christova, *ACS Cent. Sci.* **2020**, 6, 795–814.
- [18] R. Ramanan, S. S. Chaturvedi, N. Lehnert, C. J. Schofield, T. G. Karabancheva-Christova, C. Z. Christov, *Chem. Sci.* **2020**, 11, 9950-9961.
- [19] H. Torabifard, G. A. Cisneros, *Chem. Sci.* **2018**, 9, 8433–8445.
- [20] B. Wang, D. Usharani, C. Li, S. Shaik, *J. Am. Chem. Soc.* **2014**, 136, 13895-13901.

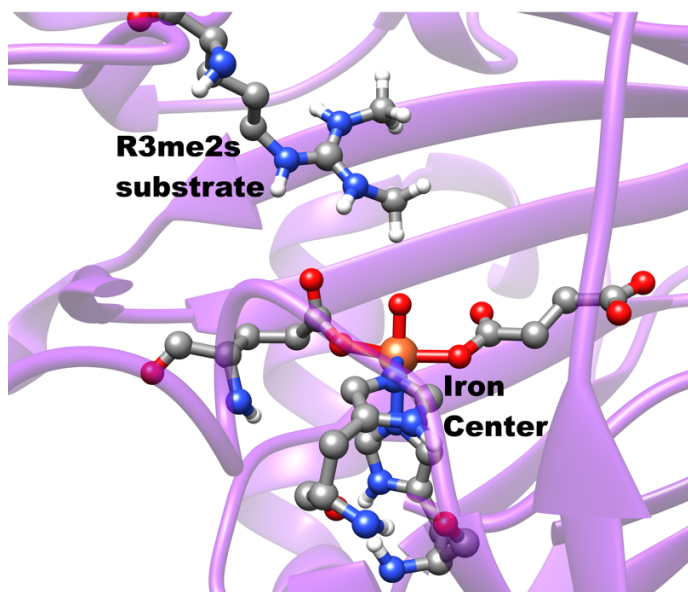
- [21] M. G. Quesne, R. Latifi, L. E. Gonzalez-Ovalle, D. Kumar, S. P. De Visser, *Chem. Eur. J.* **2014**, 20 (2), 435–446.
- [22] D. Fang, R. L. Lord, G. A. Cisneros, *J. Phys. Chem. B* **2013**, 117, 6410–6420.
- [23] J. Lu, L. Hu, J. Cheng, D. Fang, C. Wang, K. Yu, H. Jiang, Q. Cui, Y. Xu, C. A. Luo, *Phys. Chem. Chem. Phys.* **2016**, 18, 4728–4738.
- [24] S. Ye, C. Riplinger, A. Hansen, C. Krebs, J. M. Bollinger, F. Neese, *Chem. Eur. J.* **2012**, 18, 6555–6567.
- [25] (a) T. Borowski, A. Bassan, P. E. M. Siegbahn, *Chem. Eur. J.* **2004**, 10, 1031–1041; (b) N. Alberro, M. Torrent-Sucarrat, I. Arrastia, A. Arrieta, F. P. Cossio, *Chem. Eur. J.* **2017**, 23, 137–148.
- [26] (a) A. Wójcik, M. Radoń, T. Borowski, *J. Phys. Chem. A* **2016**, 120, 1261–1274; (b) S. O. Waheed, R. Ramanan, S. S. Chaturvedi, J. Ainsley, M. Evison, J. M. Ames, C. J. Schofield, C. Z. Christov, T. G. Karabancheva-Christova, *Org. Biomol. Chem.* **2019**, 17, 2223–2231; (c) S. S. Chaturvedi, R. Ramanan, S. O. Waheed, T. G. Karabancheva-Christova, C. Z. Christov, *Adv. Protein Chem. Struct. Biol.* **2019**, 117, 113–125.
- [27] D. Usharani, D. Janardanan, C. Li, S. Shaik, *Acc. Chem. Res.* **2013**, 46, 471–482.
- [28] (a) S. Shaik, H. Chen, D. Janardanan, *Nat. Chem.* **2011**, 3, 19–27; (b) S. Shaik, D. Mandal, R. Ramanan, *Nat. Chem.* **2016**, 8, 1091–1098; (c) R. Meir, H. Chen, W. Lai, S. Shaik, *ChemPhysChem* **2010**, 11, 301–310; (d) S. Shaik, R. Ramanan, D. Danovich, D. Mandal, *Chem. Soc. Rev.* **2018**, 47, 5125–5145; (e) T. Stuyver, R. Ramanan, D. Mallick, S. Shaik, *Angew. Chem. Int. Ed.* **2020**, 59, 7915–7920; *Angew. Chem.* **2020**, 132, 7989–7994; (f) N. S. Hill, M. L. Coote, *J. Am. Chem. Soc.* **2018**, 140, 17800–17804; (g) S. Shaik, D. David, J. Joy, Z. Wang, T. Stuyver, *J. Am. Chem. Soc.* **2020**, 142, 12551–12562.
- [29] L. Hillringhaus, W. W. Yue, N. R. Rose, S. S. Ng, C. Gileadi, C. Loenarz, S. H. Bello, J. E. Bray, C. J. Schofield, U. Oppermann, *J. Biol. Chem.* **2011**, 286, 41616–41625.

- [30] S. S. Ng, K. L. Kavanagh, M. A. McDonough, D. Butler, E. S. Pilka, B. M. R. Lienard, J. E. Bray, P. Savitsky, O. Gileadi, F. Von Delft, N. R. Rose, J. Offer, J. C. Scheinost, T. Borowski, M. Sundstrom, C. J. Schofield, U. Oppermann, *Nature* **2007**, 448, 87–91.
- [31] L. J. Walport, R. J. Hopkinson, R. Chowdhury, R. Schiller, W. Ge, A. Kawamura, C. J. Schofield, *Nat. Commun.* **2016**, 7, 11974.
- [32] R. Salomon-Ferrer, A. W. Götz, D. Poole, S. Le Grand, R. C. Walker, *J. Chem. Theory Comput.* **2013**, 9, 3878–3888.
- [33] (a) P. Li, K. M. Merz, *J. Chem. Inf. Model.* **2016**, 56, 599–604; (b) C. Krebs, D. G. Fujimori, C. T. Walsh, J. M. Bollinger, *Acc. Chem. Res.* **2007**, 40, 484–492; (c) E. I. Solomon, K. M. Light, L. V. Liu, M. Srnc, S. D. Wong, *Acc. Chem. Res.* **2013**, 56, 2725–2739; (d) X. Song, J. Lu, W. Lai, *Phys. Chem. Chem. Phys.* **2017**, 19, 20188–20197; (e) S. O. Waheed, S. S. Chaturvedi, T. G. Karabancheva-Christova, C. Z. Christov, *ACS Catal.* **2021**, 11, 3877–3890.
- [34] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. Simmerling, *J. Chem. Theory Comput.* **2015**, 11, 3696–3713.
- [35] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, 79, 926–935.
- [36] T. Darden, D. York, L. J. Pedersen, *Chem. Phys.* **1993**, 98, 10089–10092.
- [37] J. A. Lzaguirre, D. P. Catarello, J. M. Wozniak, R. D. Skeel, *J. Chem. Phys.* **2001**, 114, 2090–2098.
- [38] H. J. C. Berendsen, J. P. M. Postma, W. F. Van Gunsteren, A. Dinola, J. R. Haak, *J. Chem. Phys.* **1984**, 81, 3684–3690.
- [39] J. P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, *J. Comput. Phys.* **1977**, 23, 327–341.

- [40] P. Sherwood, A. H. De Vries, M. F. Guest, G. Schreckenbach, C. R. A. Catlow, S. A. French, A. A. Sokol, S. T. Bromley, W. Thiel, A. J. Turner, S. Billeter, F. Terstegen, S. Thiel, J. Kendrick, S. C. Rogers, J. Casci, M. Watson, F. King, E. Karlsen, M. Sjøvoll, A. Fahmi, A. Schäfer, C. Lennartz, *J. Mol. Struct. THEOCHEM* **2003**, 632, 1–28.
- [41] S. Metz, J. Kästner, A. A. Sokol, T. W. Keal, P. Sherwood, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, 4, 101–110.
- [42] R. Ahlrichs, M. Bär, M. Häser, H. Horn, C. Kölmel, *Chem. Phys. Lett.* **1989**, 162, 165–169.
- [43] W. Smith, T. R. Forester, *J. Mol. Graph.* **1996**, 14, 136–141.
- [44] D. Bakowies, W. Thiel, *J. Phys. Chem.* **1996**, 100, 10580–10594.
- [45] A. D. Becke, *J. Chem. Phys.* **1993**, 98, 5648–5652.
- [46] A. Schäfer, H. Horn, R. Ahlrichs, *J. Chem. Phys.* **1992**, 97, 2571–2577.
- [47] J. Kästner, J. M. Carr, T. W. Keal, W. Thiel, A. Wander, P. Sherwood, *J. Phys. Chem. A* **2009**, 113, 11856–11865.
- [48] S. R. Billeter, A. J. Turner, W. Thiel, *Phys. Chem. Chem. Phys.* **2000**, 2, 2177–2186.
- [49] (a) S. Grimme, J. Antony, S. Ehrlich, H. A. Krieg, *J. Chem. Phys.* **2010**, 132, 154104; (b) T. Stuyver, J. Huang, D. Mallick, D. Danovich, S. Shaik, *J. Comput. Chem.* **2020**, 41, 74–82.
- [50] C. J. Laconsay, K. Y. Tsui, D. J. Tantillo, *Chem. Sci.* **2020**, 11, 2231–2242.
- [51] K. D. Dubey, T. Stuyver, S. Kalita, S. Shaik, *J. Am. Chem. Soc.* **2020**, 142, 9955–9965.
- [52] J. R. Horton, A. K. Upadhyay, H. H. Qi, X. Zhang, Y. Shi, X. Cheng, *Nat. Struct. Mol. Biol.* **2010**, 17, 38–44.
- [53] C. Yi, G. Jia, G. Hou, Q. Dai, W. Zhang, G. Zheng, X. Jian, C. G. Yang, Q. Cui, C. He, *Nature* **2010**, 468, 330–333.
- [54] B. Yu, J. F. Hunt, *Proc. Natl. Acad. Sci.* **2009**, 106, 14315–14320.

- [55] C. Yi, B. Chen, B. Qi, W. Zhang, G. Jia, L. Zhang, C. J. Li, A. R. Dinner, C. G. Yang, C. He, *Nat. Struct. Mol. Biol.* **2012**, 19, 671-676.
- [56] C-G. Yang, C. Yi, E. M. Duguid, C. T. Sullivan, X. Jian, P. A. Rice, C. He, *Nature* **2008**, 452, 961-965.
- [57] R. J. Hopkinson, L. J. Walport, M. Münzel, N. R. Rose, T. J. Smart, A. Kawamura, T. D. W. Claridge, C. J. Schofield, *Angew. Chem. Int. Ed.* **2013**, 52, 7709 –7713; *Angew. Chem.* **2013**, 125, 7863-7867.
- [58] J. Lalevée, X. Allonas, J. Fouassier, *J. Am. Chem. Soc.* **2002**, 124, 9613-9621.

Entry for the Table of Contents:



The QM/MM calculations of the demethylation of symmetrical dimethylated arginine (R3me2s) substrate by KDM4E via C-H and N-H pathways show the preference for the former and application of external electric fields (EEFs) increases the rate of C-H activation, while inhibiting the rate of N-H activation.

Keywords

Histone demethylation, non-heme Iron enzymes, molecular dynamics, QM/MM calculations, JmjC demethylases (KDMs)

Institutional and/or researchers Twitter usernames:

@michigantech

@mturesearch

@MTU_chemistry

@SchofieldOxford

@OxfordChemistry

@sowaheed