

## Table of Contents

Methods		
<b>1</b>	System Preparation	S2
<b>2</b>	Details of MCPB	S2
<b>3</b>	MD Simulations	S3
<b>4</b>	QM cluster calculations	S3
<b>5</b>	QM/MM calculations using Gaussian 09	S4
<b>6</b>	QM/MM calculations using Chemshell	S4
List of Tables		
<b>Table S1.</b>	Hydrogen bonding interactions with respective occupancies, and distances for the PHF8: H3 substrate interactions	S5
<b>Table S2.</b>	Interactions of made the histone substrate with PHF8	S5
<b>Table S3.</b>	PHF8 QM/MM analysis of the distances between Fe(II) and its coordinating ligands	S6
<b>Table S4.</b>	Values of the dihedral angles of the two methyl groups (1 and 2) of the H3K9me2 substrate in PHF8	S7
<b>Table S5.</b>	QM/MM analysis of second sphere interactions within PHF8	S8
<b>Table S6.</b>	The CHO hydrogen bonding of the H3K9me2 substrate (also with H3K4me3) in the active site of PHF8	S9
List of Figures		
<b>Figure S1.</b>	Views from MD structures of (A) PHF8 and (B) KIAA1718	S10
<b>Figure S2.</b>	Distribution of the dihedral angles	S10
<b>Figure S3.</b>	2OG: Atom names and numbers used	S11
<b>Figure S4.</b>	Interactions of nearby lysine residues with the C5 carboxylate of 2OG	S11
<b>Figure S5.</b>	The RMSD profile of the KIAA1718	S12
<b>Figure S6.</b>	RMSD analyses of alpha C atoms for the linker regions	S13
<b>Figure S7.</b>	Radii of gyration of the linker regions of PHF8 and KIAA1718	S13
<b>Figure S8.</b>	Dynamic Cross Correlation Analysis (DCCA) of PHF8	S14
<b>Figure S9.</b>	DCCA of KIAA1718	S14
<b>Figure S10.</b>	DCCA of KIAA1718	S15
<b>Figure S11.</b>	Contributions of the residues to the principal component 1	S15
<b>Figure S12.</b>	RMSD of KIAA1718 and KIAA1718*	S16
<b>Figure S13.</b>	Radii of gyration of KIAA1718 and KIAA1718*	S16
<b>Figure S14.</b>	Distances between the centers of masses of the PHD-JmjC domains in KIAA1718 and KIAA1718*	S17
<b>Figure S15.</b>	Comparison of the QM region of PHF8 with the 300ns snapshot of the QM/MM optimized structure	S17
<b>Figure S16.</b>	Comparison of mechanical embedding to electronic embedding	S18
<b>Figure S17.</b>	Active site interactions in the QM/MM optimized geometry of the PHF8	S18
<b>Figure S18.</b>	Radius of gyration of the H3 histone substrate during MD simulations of PHF8	S19
<b>References</b>		S19

## Methods

### System Preparation

PHF8 (KDM7B) and KIAA1718 (KDM7A): A crystal structure of PHF8 (PDB: 3KV4) [1] in complex with a fragment of its histone (H3) substrate (H3K4me3 and H3K4me2) and close 2OG analogue N-oxalylglycine was used for initial modelling. The non-crystallographically observed linker region (residue 65-79) between the PHD and the JmjC domain was modelled using Modeller [2]. A crystal structure of KIAA1718 (PDB: 3KV6 chain D used here [1]) complexed with 2OG was used in further structure preparation for computational studies. The reported structure has Fe present, coordinated by 2OG and ligands. [1] The substrate histone H3 was modelled into the active site of KIAA1718 by superimposing the PHF8 crystal structure with that of KIAA1718 using Maestro (Schrodinger LLC, New York). To investigate how the linkers are crucial for regulating activity of both PHF8 and KIAA1718 we modelled the linker of PHF8 into KIAA1718 using Modeller [2]. The linker of PHF8 consists of residues 65-78 and KIAA1718 consists of residues 98-113.

The protonation states of ionisable sidechains were assessed using the H++ server [3]. Histidine residues coordinating with the metal were assigned protonation states based on visual inspection of their local environment. N-Oxalylglycine was modelled to 2OG by replacing its NH with a methylene using GaussView 5.0 [4]. Hydrogen atoms were added to 2OG using the REDUCE programme in AMBER 14 [5]. The Amber parameters for 2OG were developed using the general Amber force field (GAFF) [6] with Antechamber. The atomic charges of the cosubstrate were calculated based on the electrostatic potential from single point HF/6-31G\* calculations using Gaussian09 [7]. The restrained electrostatic potential (RESP) [8] method was used for charge fitting. The substrate histone H3 was modified by di- and tri-methylated lysine residues, denoted as H3K9me2 and H3K4me3 [1]. The Amber parameters for di-methylated lysine (M2L) and tri-methylated lysine (M3L) were generated using Antechamber. [9].

**MCPB:** The Amber parameters for the active site containing iron (Fe(II) high spin  $S=2$ ,  $M=5$ , ground state [10a,b,11] and the coordinating ligands (2OG, bidentate; His and Asp, both monodentate) were prepared using the Metal Centre Parameter Builder (MCPB) using MCPB.py v1.0 Beta2 [12] with a 5-coordinate distorted square pyramidal geometry. Binding of the substrate in close vicinity of 2OG results in the dissociation of the coordinated water and gives a five coordinate (5C) distorted square pyramidal geometry [10a,13]. The metal centre parameters were derived based on the bonded and electrostatic model approach as implemented in MCPB. Bond and the angle force constants were derived using the Seminario Method. Point charge parameters for the electrostatic potentials were obtained using the ChgModB method. Pabis et al.<sup>14</sup> have applied the MCPB tools for the description of the mononuclear non-heme iron centre and iron-sulfur Rieske cluster [14]. The molecular dynamics simulations run using these parameters have reproduced the geometry of metal-ligand complex successfully [14]. The description of the zinc ion and its coordinating ligands in the PHD domain were described using Zinc Amber Force Field (ZAFF) method. [15]

**MD Simulations:** Molecular dynamics simulations were performed using the GPU version [16] of the PMEMD engine integrated with Amber 14 [17]. The FF14SB [18] force field was used in all simulations; the Leap module was used to add missing hydrogen atoms and counter ions for neutralisation of the protein system. All systems were immersed into a truncated octahedral box with TIP3P water [19] molecules present such that no protein atom was within 10 Å of any box edge. Periodic boundary conditions were employed in all simulations. Long-range electrostatic interactions were calculated using the particle mesh Ewald (PME) method [20] with a direct space and vdW cut-off of 8 Å. The various systems were subjected to energy minimization using first steepest descent (5000 steps) followed by conjugate gradient (5000 steps). The solute molecules were restrained using a restrained potential of 100 kcal mol<sup>-1</sup> Å<sup>2</sup>; only solvent molecules and ions were allowed to minimize. This was followed by full minimization of the entire system with both steepest descent (5000 steps) and conjugate gradient (5000 steps) to relax the system prior to productive simulation. All the energy minimization, heating and equilibration steps during initial stages of MD were performed with the CPU version of PMEMD. The systems were then subjected to controlled heating from 0 to 300K at constant volume using Langevin thermostat [21] with a collision frequency of 1 ps<sup>-1</sup> using a canonical ensemble (NVT) MD simulation for 400 ps. Solute molecules were restrained using harmonic potential of 10 kcal mol<sup>-1</sup> Å<sup>2</sup> during the heating process. The SHAKE algorithm [22] was used to constrain bonds involving hydrogen. This was followed by equilibration at 300K in an NPT ensemble for 1 ns without restraints on solute molecules; the pressure was maintained at 1 bar using the Berendsen barostat [23]. A production MD run with explicit solvent for continuous 1μs was performed in a NPT ensemble with a target pressure set at 1 bar and constant pressure coupling of 2ps. Frames from the productive run were saved every 10 ps.

The trajectories were analysed using CPPTRAJ [24], VMD [25], UCSF Chimera [26] and R (Bio3D [27]). The Root Mean Square Deviation (RMSD) of Cα atoms of the protein with respect to minimized crystal structure, Root Mean Square Fluctuations (RMSF), Radius of gyration (Rg), Electrostatic interactions, hydrogen bonding, Solvent Accessible Surface Area (SASA), and cluster analyses were performed. The Bio3D package [27a] in R was used to produce PCA and domain cross correlation as described by Singh *et al.* [27b].

**QM cluster calculations:** Snapshots of structures were obtained from the equilibrated MD trajectory of the systems described above. The GaussView 5.0 [4] was used to set up QM calculations and Gaussian09 code [7] was used for all QM calculations. In all the calculations, iron (Fe (II) high spin S=2, M=5, ground state [10a,b,11] and the coordinating ligands (2OG (bidentate), with His and Asp as monodentate) was used. The ligating His and Asp residues were truncated and restrained at their Cβ positions and hydrogen atoms were added to saturate the bonds. Geometry optimization, frequency calculations and single point calculations were performed with Density Functional Theory (DFT) using unrestricted UBP86 functional with 10 % exact HF (Hartree Fork) exchange with 6-311G\* basis set of the Fe and its coordinating atoms (Oxygen and Nitrogen) from the ligand and rest of the atoms we employed 6-31G\* [28]. A conductor-like polarizable continuum model (CPCM) with ε=4.3 (diethyl ether as solvent) [29] was used in the QM calculations to mimic the hydrophobic active site in the protein [30-32].

**QM/MM calculations using Gaussian 09: ??CHECK** The QM/MM calculations were performed on minimized structures, corresponding to 300ns, 600ns, 700ns, 800ns, 900ns and 1000ns snapshots of PHF8 from the MD simulations. The enzyme residues and water molecules within 35 Å of iron were involved in the QM/MM optimization. These snapshots were first subjected to energy minimization for 10,000 steps by using both steepest descend (5000) and conjugate gradient (5000) algorithms using Amber14. Active site residues were restrained with restrained potential of 100 kcal mol<sup>-1</sup> Å<sup>2</sup> in the energy minimization in order to maintain the geometry of the active site. The energy minimized snapshots of all the enzymes were prepared using the Schlegel's toolkit TAO [33] for ONIOM [34-38] calculation in Gaussian09 [7]. The residues which are within 20Å of iron, including water molecules, were allowed to move freely during geometry optimization and rest of the system was frozen during geometry optimization in ONIOM. The QM/MM system was prepared using GaussView 5.0. All calculations were run using Gaussian09 code. Residues were assigned with the standard bonded and non-bonded terms available from the ff99SB force field in Gaussian09. The mechanical embedding scheme was used in the geometry optimization however we have also used electronic embedding scheme for the few snapshots. The non-bonded van der Waals parameters for iron were obtained from the method of Li *et al* [39]. The QM region in the QM/MM calculations is consistent with the QM calculations performed above; link atoms were used to saturate the 'dangling' bonds [40]. The Harmonic vibrational frequency calculations were performed at (UBP86/GEN: AMBER) level of theory and the basis set used here are consistent with the QM calculations performed.

**QM/MM calculations using Chemshell:** The QM/MM calculations were performed with the ChemShell software. [41] Turbomole [42] was used for the QM region and DL\_POLY [43] for the MM part. Amber force fields that generated for MD simulations were used for QM/MM calculations. The electronic embedding scheme is used to include the polarizing effect of the enzyme on the QM region. [44] Geometry is optimized at the B3LYP functional [45] with the def2-SVP basis set (B1). [46]

**Table S1.** Hydrogen bonding interactions with respective occupancies, and distances for the PHF8: H3 substrate interactions. Red and green colors denote PHD and JmjC domain residues, respectively; residues 452-465 belong to the H3 histone peptide. ARE BB / sc defined??

SYSTEM	Donor	Acceptor	% Occupancy	Distance (Å)
PHF8	H3A1(NH <sup>3+</sup> )	S286(sc)	28.0	3.5
PHF8	H3A1(NH <sup>3+</sup> )	D46(bb)	82.7	3.1
PHF8	H3A1(NH <sup>3+</sup> )	I45(bb)	82.7	3.1
PHF8	H3A1(NH <sup>3+</sup> )	A42(bb)	14.0	4.6
PHF8	H3R2(sc)	S284(bb)	53.0	3.5
PHF8	H3R2(bb)	S286(sc)	30.0	3.6
PHF8	E22(bb)	H3R2(bb)	89.0	3.1
PHF8	H3T3(sc)	H3A1(bb)	7.0	4.2
PHF8	H3T3(bb)	M20(bb)	82	3.1
PHF8	H3R8(bb)	H3T6(bb)	82.3	3.1
PHF8	H3S10(bb)	S192(sc)	79.6	3.3
PHF8	Y234(sc)	H3S10(bb)	99.0	2.8
PHF8	H3G12(bb)	N230(bb)	98.0	2.9

**Table S2.** Selected interactions made by the histone substrate with PHF8. The distance from the glutamyl carboxylate carbon or the nearest carbon of the phenyl ring in Phe (F) are considered.

Which atom on the histine?/

SYSTEM	Residue	Residue	Distance
PHF8	E22(sc)	H3R2 (sc)	5.1
PHF8	F19(sc)	H3T3	4.3
PHF8	E350(sc)	H3R8	4.5

**Table S3.** PHF8 QM/MM analysis of the distances between Fe(II) and its coordinating ligands in QM/MM optimized MD snapshots. Distances from the metal coordinating nitrogen of His and oxygen of Asp to Fe are given.

Name	Equatorial His 247	Axial His 319	ASP 249	2OG (C1)	2OG (C2)	H3K9me2 (Met2)	H3K9me2 (Met1)
QM minimized Cluster	2.12	2.16	2.05	2.12	2.12	5.15	5.23
QM/MM Minimized Crystal Structure	2.13	2.17	2.04	2.11	2.26	4.60	6.50
QM/MM 300ns	2.13	2.13	2.00	2.17	2.15	4.28	6.19
QM/MM 600ns	2.15	2.14	2.03	2.15	2.15	4.53	6.13
QM/MM 700ns	2.16	2.13	2.12	2.08	2.22	6.70	3.34
QM/MM 800ns	2.13	2.16	2.06	2.17	2.21	4.45	6.15
QM/MM 800ns (Electronic Embedding, ONIOM)	2.14	2.16	2.09	2.09	2.15	4.50	6.21
QM/MM 800ns (Electrostatic Embedding, ChemShell)	2.18	2.21	2.11	2.08	2.37	4.41	6.17
QM/MM 900ns	2.13	2.15	2.09	2.09	2.21	6.57	4.14
QM/MM 1000ns	2.11	2.22	1.98	2.19	2.17	4.57	6.34
<b>QM/MM Average</b>	<b>2.14</b>	<b>2.16</b>	<b>2.05</b>	<b>2.12</b>	<b>2.21</b>	<b>4.98</b>	<b>5.62</b>
<b>MD Average</b>	<b>2.30</b>	<b>2.15</b>	<b>2.022</b>	<b>1.98</b>	<b>2.11</b>	<b>4.64</b>	<b>6.00</b>

**Table S4.** Values of the dihedral angles of the two methyl groups (1 and 2) of the H3K9me2 substrate in PHF8 as observed in QM/MM optimizations of structures from MD simulation. Met1, Met2 and NZ are shown in Figure 2.

Name	CD:CE:NZ:C3 (Met2)	CD:CE:NZ:C2 (Met1)
------	--------------------	--------------------

QM/MM Minimized Crystal Structure	-61.07	65.17
QM/MM 300ns	174.9	-62.21
QM/MM 600ns	68.43	-168.87
QM/MM 700ns	64.0	-168.7
QM/MM 800ns	69.63	-168.04
QM/MM 800ns (Electronic Embedding, ONIOM)	68.09	-167.84
QM/MM 1000ns	72.15	-165.08
<b>QM/MM Average</b>	<b>65.16</b>	<b>-119.37</b>
<b>MD Average</b>	<b>67.17</b>	<b>-107.8</b>

**Table S5.** QM/MM analysis of second sphere interactions within PHF8.

Name	N189 (sc) O4 2OG	Y257 (sc) – O3 2OG	T244 (sc) – O4 2OG	K264 (sc) –C5 2OG	M2L (NZ-sc) – O1 2OG	N333 –O2 AP1
------	---------------------	-----------------------	-----------------------	----------------------	-------------------------	-----------------

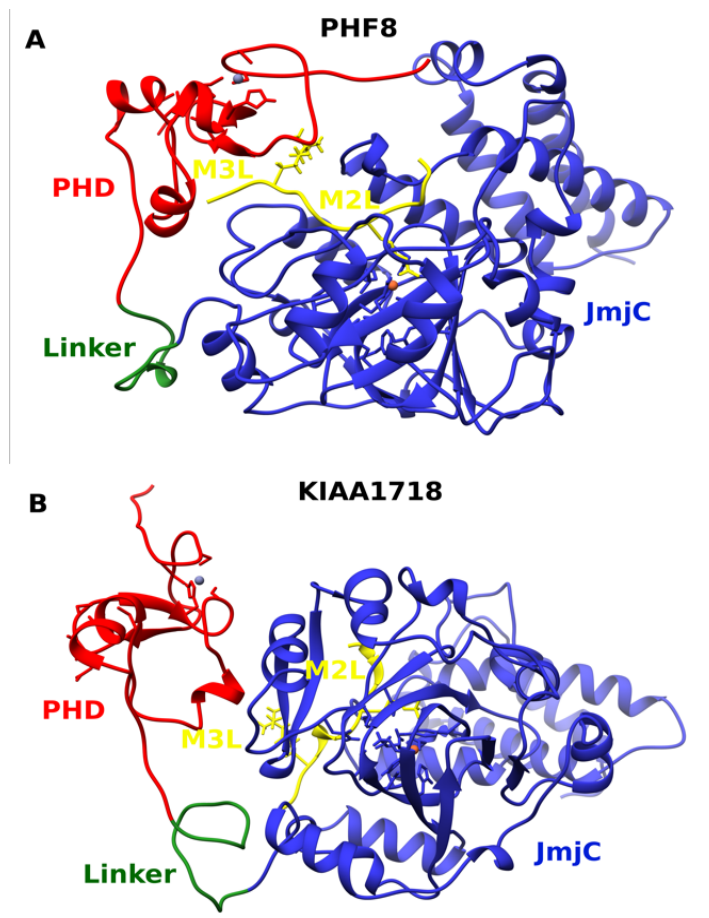


QM/MM Minimized Crystal Structure	2.93	4.41	2.63	3.56	2.64	3.02
QM/MM 300ns	4.39	2.64	2.66	3.50	2.62	5.76
QM/MM 600ns	4.04	2.72	4.33	3.19	2.63	3.53
QM/MM 700ns	3.17	4.32	2.59	3.48	4.07	3.36
QM/MM 800ns	5.11	2.71	4.41	3.19	2.64	3.73
QM/MM 800ns (Electronic Embedding, ONIOM)	5.31	2.68	4.49	3.20	2.78	3.62
QM/MM 900ns	4.76	2.69	4.29	3.16	3.93	3.31
QM/MM 1000ns	2.91	4.73	2.61	5.33	2.60	3.75
<b>QM/MM Average</b>	<b>4.07</b>	<b>3.36</b>	<b>3.50</b>	<b>3.58</b>	<b>2.99</b>	<b>3.76</b>
<b>MD Average</b>	<b>5.35</b>	<b>3.96</b>	<b>4.312</b>	<b>3.505</b>	<b>3.505</b>	<b>3.52</b>

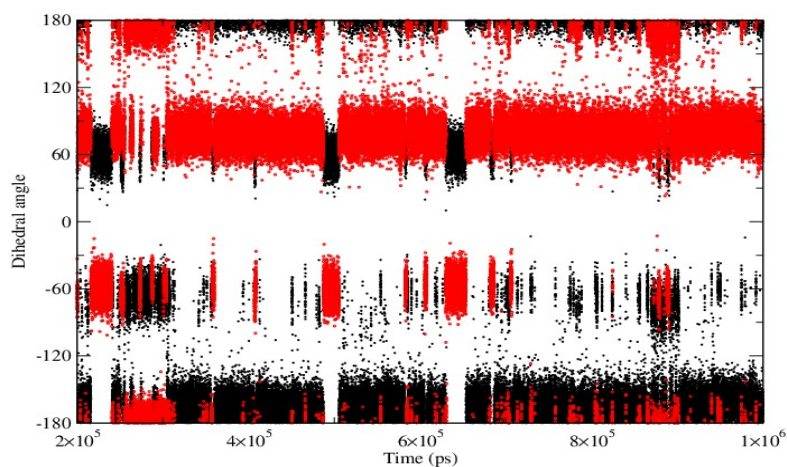
**Table S6.** The CHO hydrogen bonding of the H3K9me2 substrate (and H3K4me3) in the active site of PHF8.

Name	CH3-AP1 (Å)	CH3-AP1- angle	CH3-N33(Å)	CH3-N33 angle
------	-------------	-------------------	------------	------------------

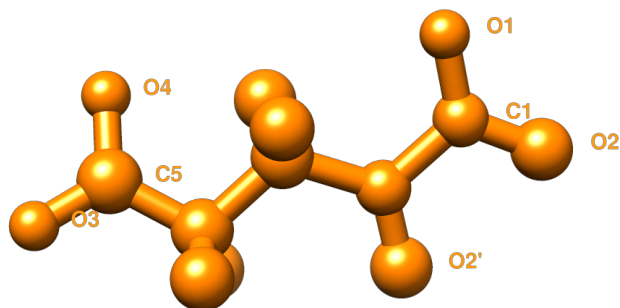
<b>QM/MM Minimized Crystal Structure</b>	3.3	112.4	3.0	146.02
QM/MM 300ns	3.4	149.1	3.1	108.8
QM/MM 600ns	3.5	161.7	3.3	150.4
QM/MM 700ns	3.3	153.4	3.0	111.8
QM/MM 800ns	3.7	156.3	3.2	173.9
QM/MM 800ns (Electronic Embedding, ONIOM)	3.5	152.7	3.2	172.0
QM/MM 1000ns	3.4	170.0	3.4	160.73
<b>QM/MM Average</b>	<b>3.4</b>	<b>150.8</b>	<b>3.17</b>	<b>146.2</b>



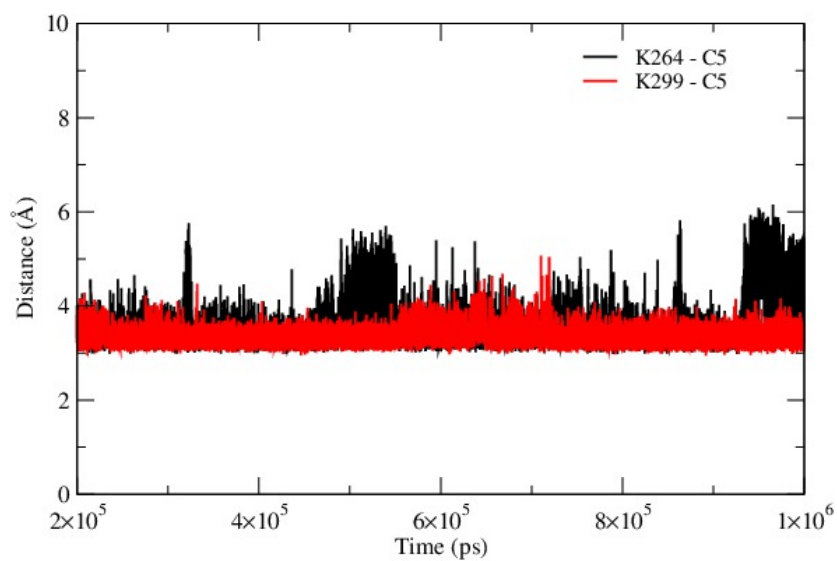
**Figure S1.** Views from MD structures of (A) PHF8 and (B) KIAA1718. Averaged structures from clustering are presented.



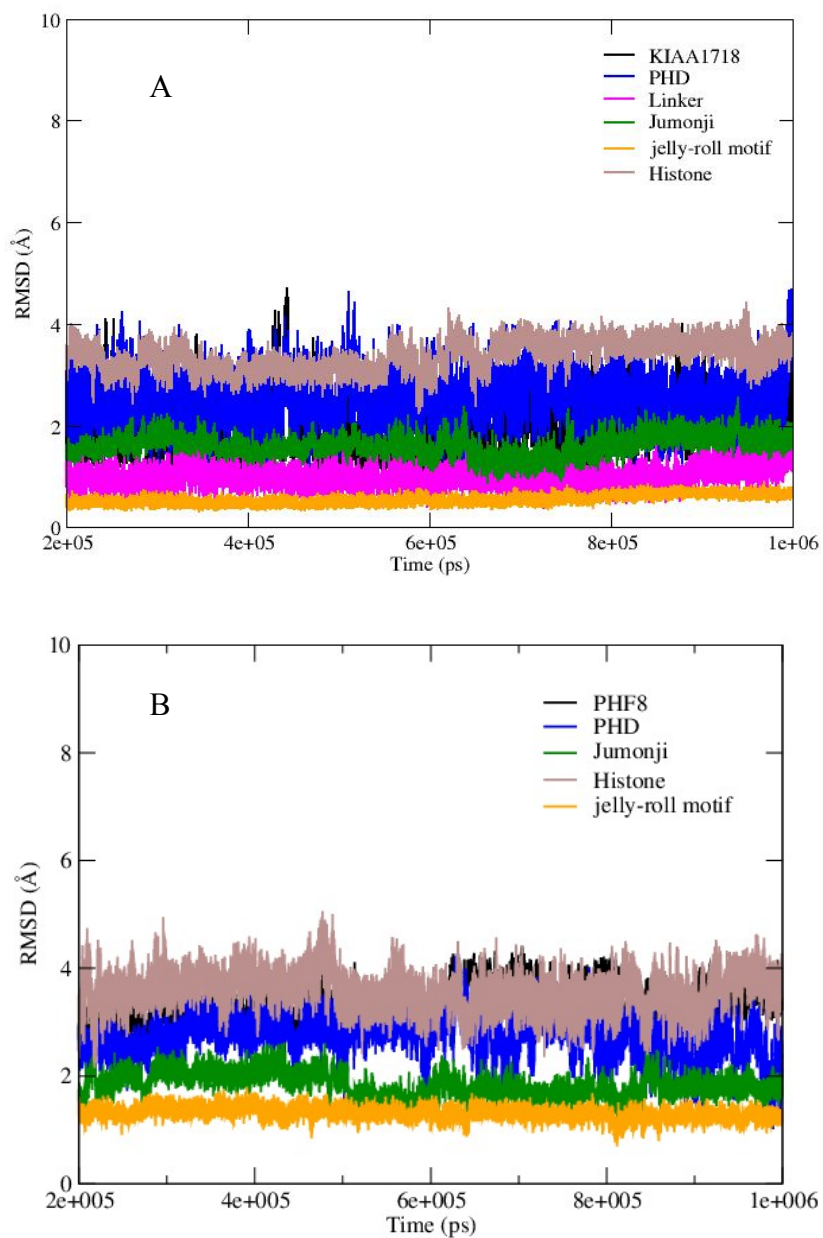
**Figure S2.** Distribution of the dihedral angles: CD: CE: NZ: Met1 (red colour) and CD: CE: NZ: Met2 (black colour) in PHF8 enzyme of the H3K9me2. The values indicate relatively rigid orientations for both Met1 and Met2 of H3K9me2.



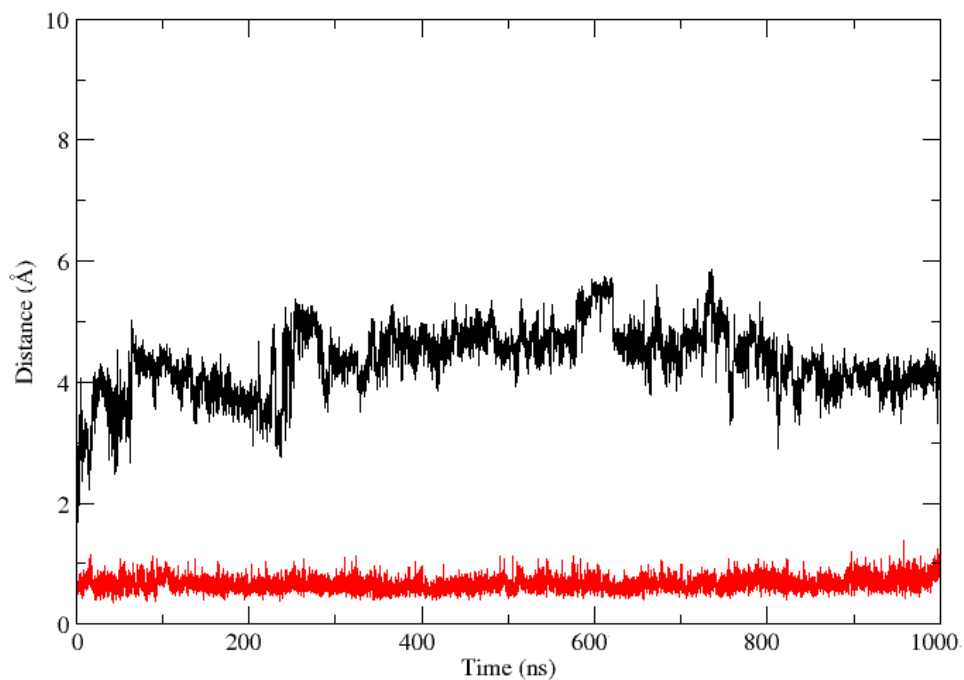
**Figure S3.** 2OG: Atom names and numbers used.



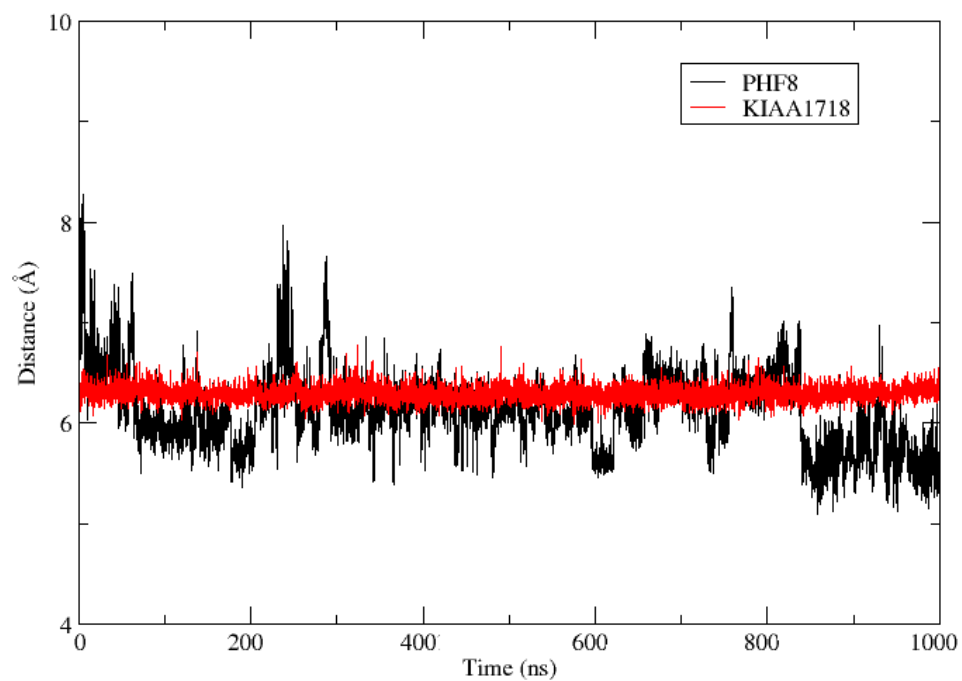
**Figure S4.** Interactions of nearby lysine residues with the C5 carboxylate of 2OG. K264 and K299 belong to PHF8 and KIAA1718. The stability of these interactions is important for the orientation of the 2OG in the active site.



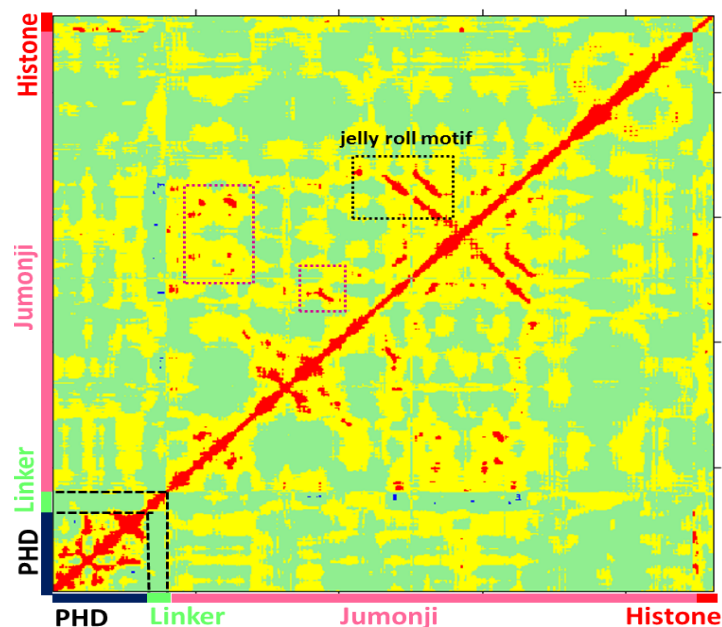
**Figure S5.** (A) The RMSD profile of the KIAA1718 with its individual domains/regions (PHD, Linker and JmjC) and jelly-roll motif, and histone substrate. (B) RMSD profile of PHF8 with its individual domains/regions (PHD, Linker, Jumonji), and histone substrate.



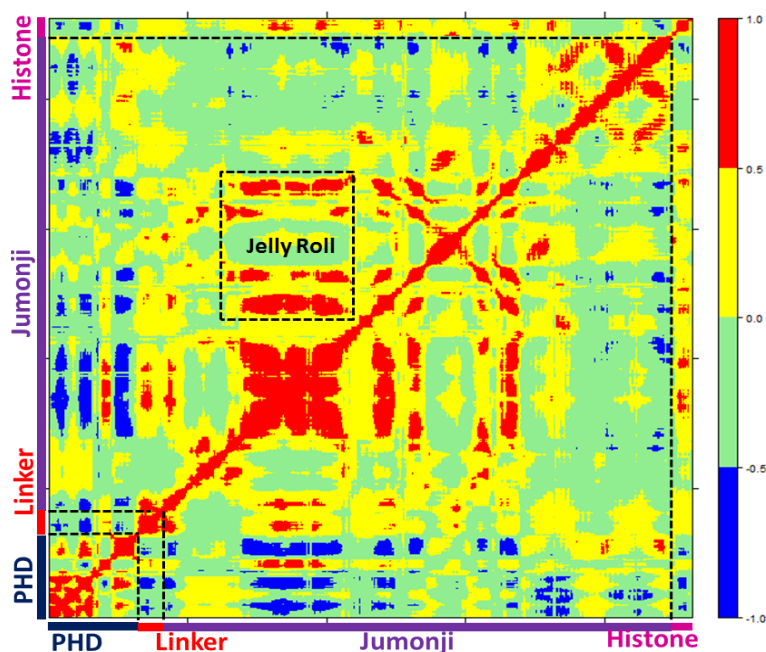
**Figure S6.** RMSD analyses of alpha C atoms for the linker regions of PHF8 (black) and KIAA1718 (red) imply the PHF8 linker is more flexible.



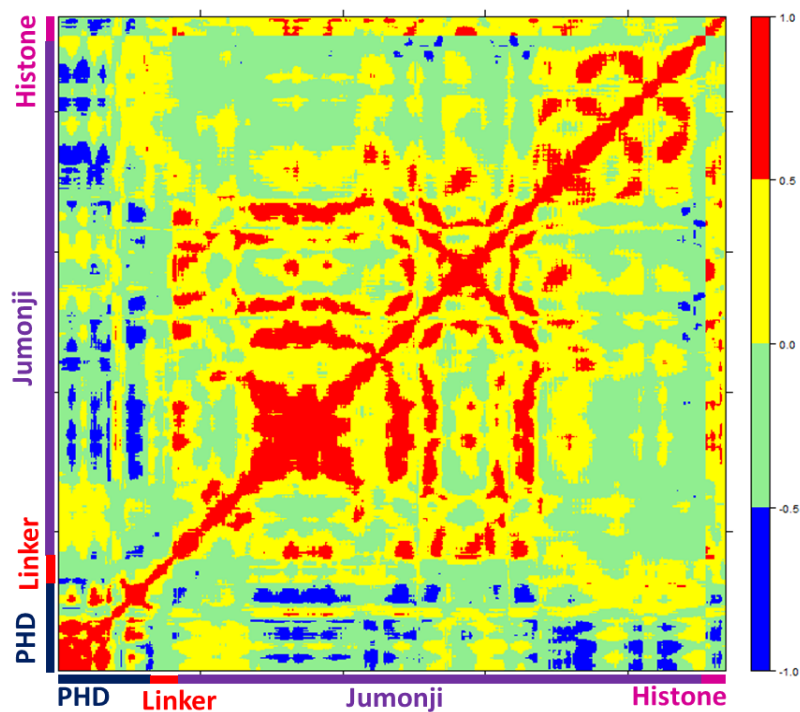
**Figure S7.** Radii of gyration of the linker regions of PHF8 (black) and KIAA1718 (red) reveal the more flexible nature of the PHF8 linker compared to the relatively rigid linker of KIAA1718.



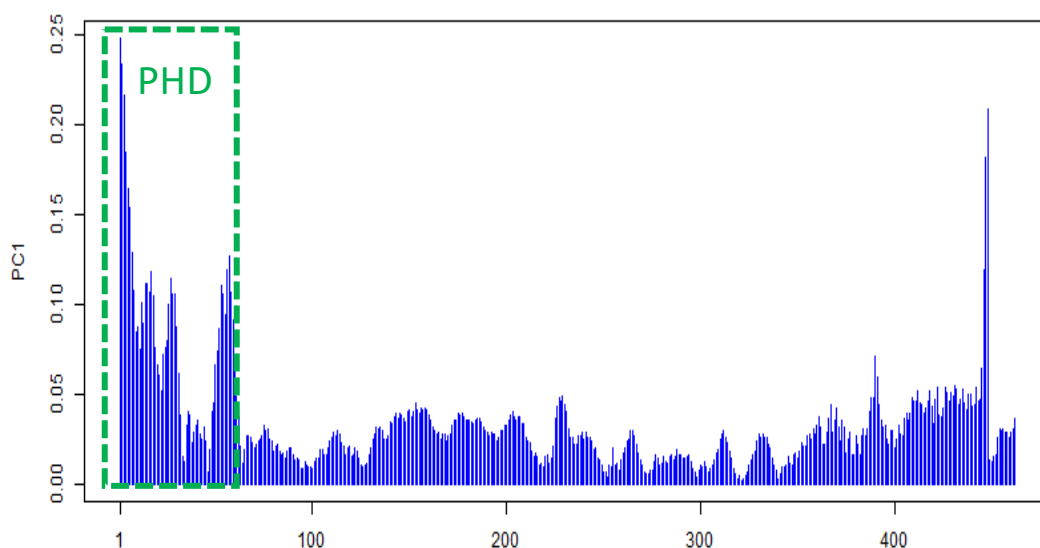
**Figure S8.** Dynamic Cross Correlation Analysis (DCCA) of PHF8. Strong positive correlations are shown in red and strong negative (anti-) correlations are shown in blue. The figure informs on collective motions of regions involving the protein and substrate.



**Figure S9.** DCCA of KIAA1718. Strong positive correlations are shown in red and strong negative (anti-) correlations - in blue.

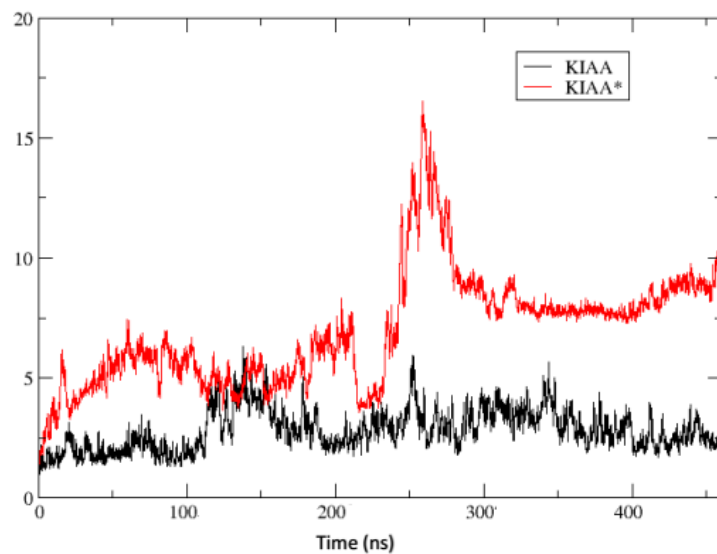


**Figure S10.** Dynamic Cross Correlation Analysis (DCCA) of KIAA1718\*. Strong positive correlations are shown in red and strong negative (anti-) correlations are shown in blue.

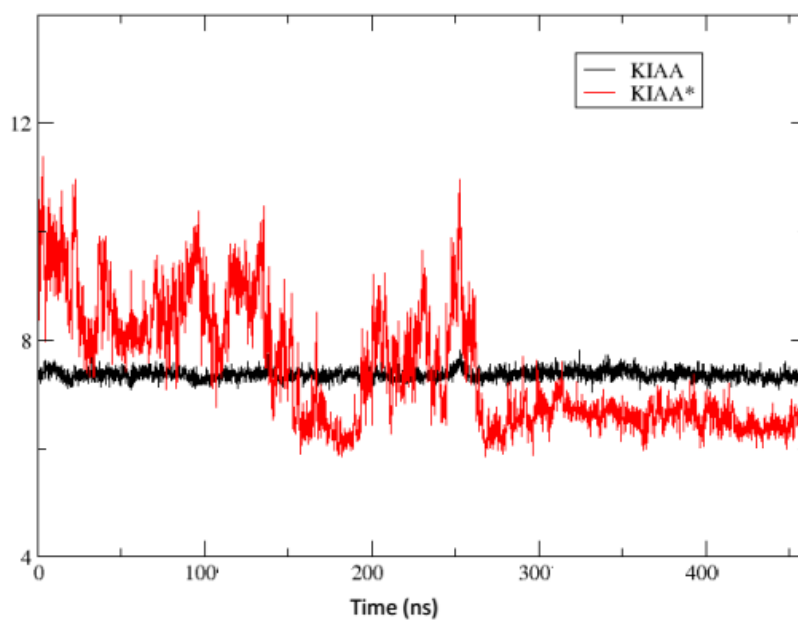


**Figure S11.** Contributions of the residues to the principal component 1 (PC1) in KIAA1718.

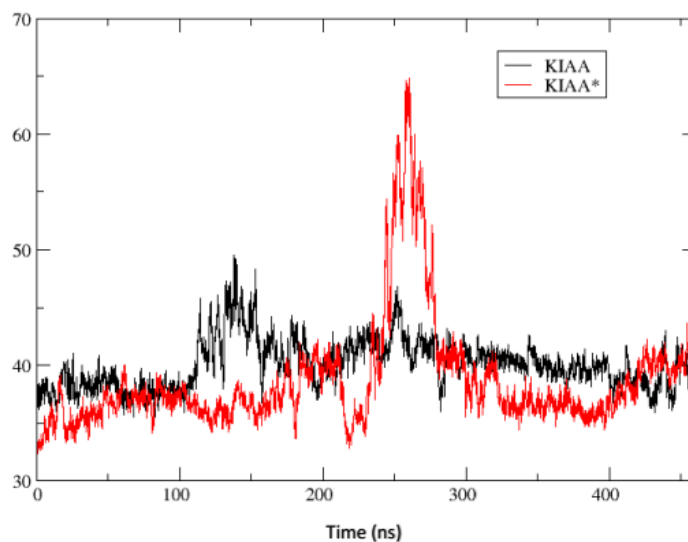




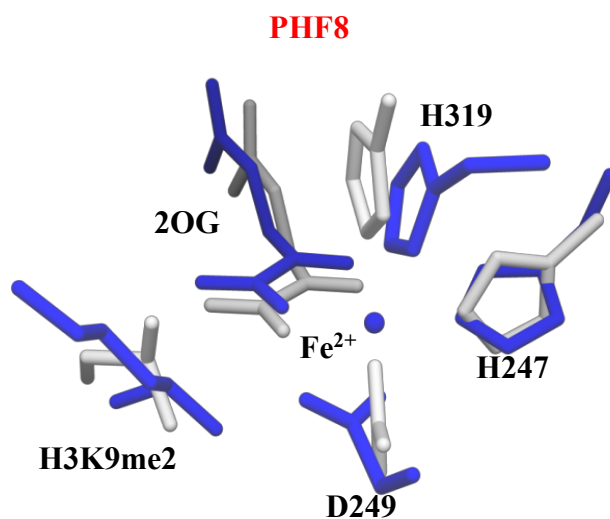
**Figure S12.** RMSD of KIAA1718 and KIAA1718\*. Exchange of the PHF8 linker into the latter leads to increased flexibility in KIAA1718\*.



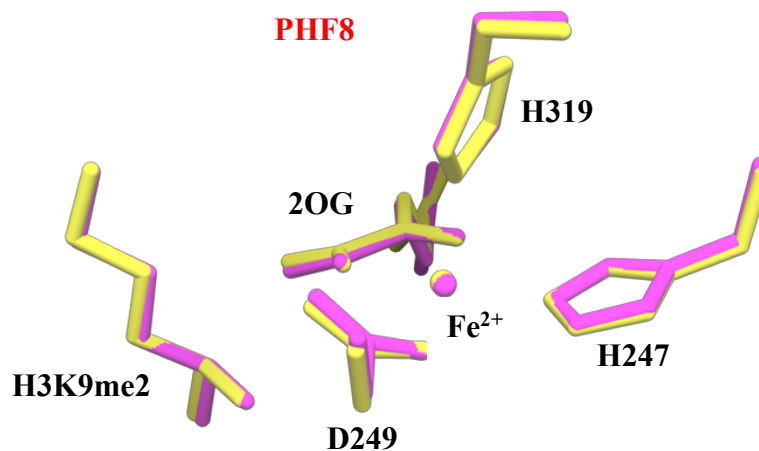
**Figure S13.** Radii of gyration of KIAA1718 and KIAA1718\*. The larger fluctuations observed for KIAA1718\* compared to KIAA1718 are due to the in silico introduced PHF8 linker of the latter.



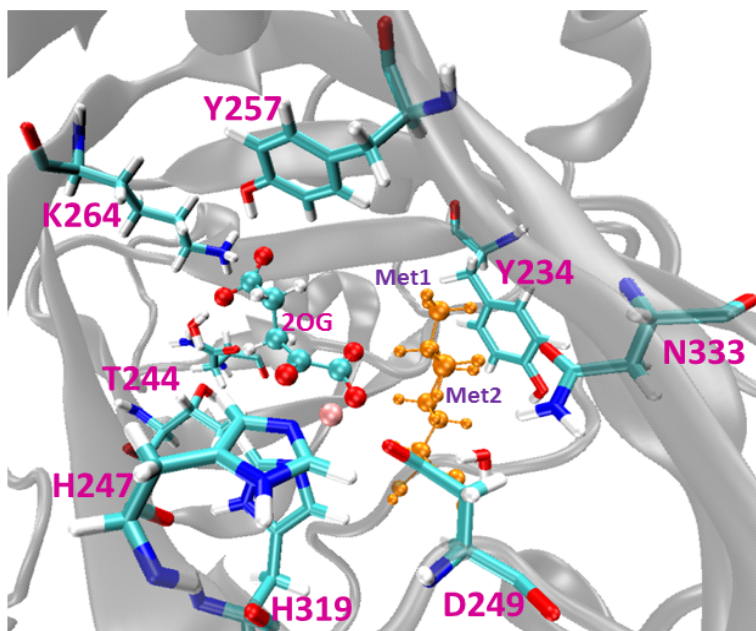
**Figure S14.** Distances between the centers of masses of the PHD-JmjC domains in KIAA1718 and KIAA1718\*. Introducing the PHF linker into KIAA1718\* leads to larger flexibility and a trend for closer positioning of PHD and JmjC domains in KIAA1718\* compared to KIAA1718.



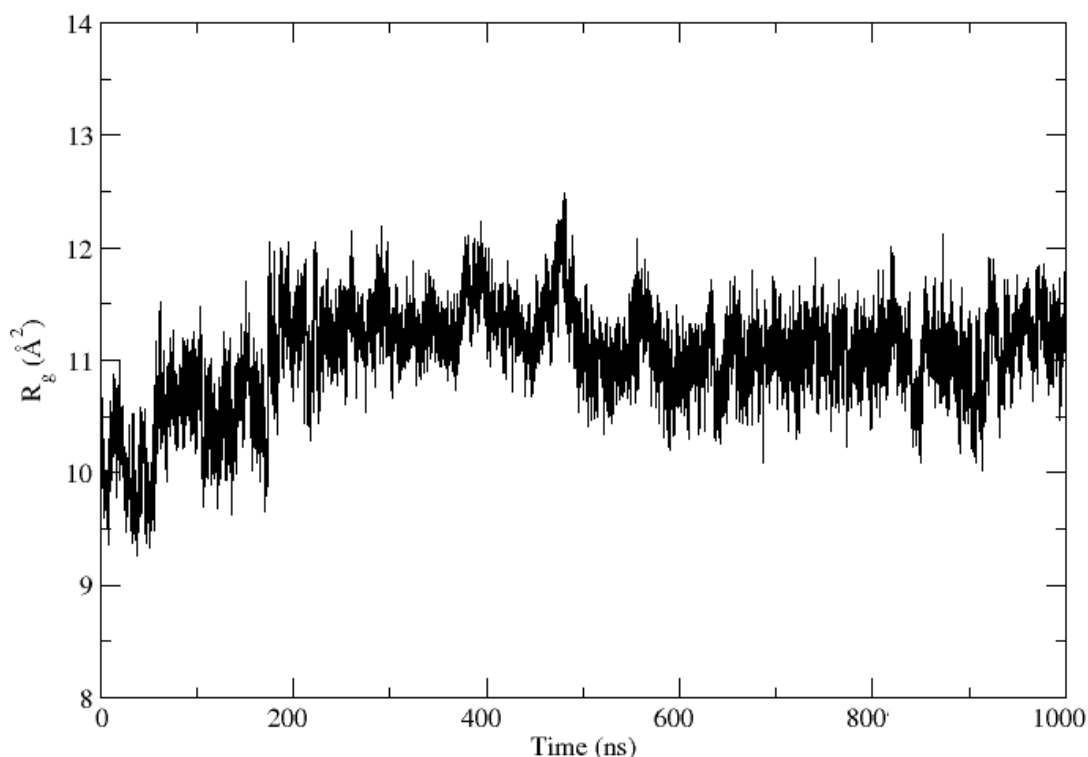
**Figure S15.** Comparison of the QM region of PHF8 with the 300ns snapshot of the QM/MM optimized structure. The ONIOM and QM optimized structures are shown in blue and silver; the differences reveal the importance of the protein environment.



**Figure S16.** Comparison of the QM/MM ONIOM optimizations with mechanical embedding to electronic embedding for the 800ns snapshot of PHF8. The yellow and pink liquorice representations correspond to electronic and mechanical embedding QM/MM minimization. The RMSD of the iron and the coordinating atoms for the two methods is 0.047Å, indicating consistency between the optimized structures obtained using different methods.



**Figure S17.** Active site interactions in the QM/MM optimized geometry of the PHF8:substrate complex at 800ns.



**Figure S18.** Radius of gyration of the H3 histone substrate during MD simulations of PHF8 indicate a more open conformation than observed by crystallography.

## References

- 1 J. R. Horton, A. K. Upadhyay, H. H. Qi, X. Zhang, Y. Shi and X. Cheng, *Nat. Struct. Mol. Biol.*, **2010**, *17*, 38–43.
- 2 A. Fiser and A. Šali, *Methods Enzymol.*, **2003**, *374*, 461–491
- 3 J. C. Gordon, J. B. Myers, T. Folta, V. Shoja, L. S. Heath and A. Onufriev, *Nucleic Acids Res.*, **2005**, *33*, W368-71.
- 4 R. Dennington, T. Keith, J. Millam, (**2009**) GaussView, Version 5. *J. Semichem Inc* Shawnee Mission KS
- 5 J. M. Word, S. C. Lovell, J. S. Richardson and D. C. Richardson, *J. Mol. Biol.*, **1999**, *285*, 1735–1747.
- 6 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, **2004**, *25*, 1157–1174.

- 7 Gaussian 09, Revision A.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian, Inc., Wallingford CT, **2016**.
- 8 W. D. Cornell, P. Cieplak, C. I. Bayly and P. A. Kollman, *J. Am. Chem. Soc.*, **1993**, *115*, 9620-9631.
- 9 Papageorgiou D (**2013**) AMBER parameters for modified amino acids.
- 10 a) E. I. Solomon, A. Decker and N. Lehnert, *Proc. Natl. Acad. Sci. U. S. A.*, **2003**, *100*, 3589–94. b) E. I. Solomon, S. Goudarzi and K. D. Sutherlin, *Biochemistry*, **2016**, *55*, 6363–6374. c) J. Zhou, M. Gunsior, B. O. Bachmann, C. A. Townsend and E. I. Solomon, *J. Am. Chem. Soc.*, **1998**, *120*, 13539–13540
- 11 B. Wang, Z. Cao, D. A. Sharon and S. Shaik, *ACS Catal.*, **2015**, *5*, 7077–7090.
- 12 P. Li and K. M. Merz, *J. Chem. Inf. Model.*, **2016**, *56*, 599–604.
- 13 E. I. Solomon, T. C. Brunold, M. I. Davis, J. N. Kemsley, S.-K. Lee, N. Lehnert, F. Neese, A. J. Skulan, Y.-S. Yang and J. Zhou, *Chem. Rev.*, **2000**, *100*, 235–350.
- 14 A. Pabis, I. Geronimo, D. M. York and P. Paneth, *J. Chem. Theory Comput.*, **2014**, *10*, 2246–2254.
- 15 M. B. Peters, Y. Yang, B. Wang, L. Füsti-Molnár, M. N. Weaver and K. M. Merz, *J. Chem. Theory Comput.*, **2010**, *6*, 2935–2947.
- 16 R. Salomon-Ferrer, A. W. Götz, D. Poole, S. Le Grand and R. C. Walker, *J. Chem. Theory Comput.*, **2013**, *9*, 3878–3888.

- 17 D.A. Case, V. Babin, J.T. Berryman, R.M. Betz, Q. Cai, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, H. Gohlke, A.W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T.S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K.M. Merz, F. Paesani, D.R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C.L. Simmerling, W. Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu and P.A. Kollman (**2014**), AMBER 14, University of California, San Francisco.
- 18 J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, *J. Chem. Theory Comput.*, **2015**, *11*, 3696–3713.
- 19 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.*, **1983**, *79*, 926-935.
- 20 T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, **1993**, *98*, 10089–10092.
- 21 R. L. Davidchack, R. Handel and M. V. Tretyakov, *J. Chem. Phys.*, **2009**, *130*, 234101.
- 22 J. P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comput. Phys.*, **1977**, *23*, 327–341.
- 23 H. J. C. Berendsen, J. P. M. Postma, W. F. Van Gunsteren, A. Dinola and J. R. Haak, *J. Chem. Phys.*, **1984**, *81*, 3684–3690.
- 24 D. R. Roe and T. E. Cheatham, *J. Chem. Theory Comput.*, **2013**, *9*, 3084–3095.
- 25 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graph.*, **1996**, *14*, 33–38.
- 26 E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J. Comput. Chem.*, **2004**, *25*, 1605–1612.
- 27 (a) B. J. Grant, A. P. C. Rodrigues, K. M. ElSawy, J. A. McCammon and L. S. D. Caves, *Bioinformatics*, **2006**, *22*, 2695–2696. (b) W. Singh, G. Fields, C. Christov, and T. G. Karabancheva-Christova, *RSC Advances*, **2016**, *6*, 23223-23232.
- 28 G. Schenk, M. Y. M. Pau and E. I. Solomon, *J. Am. Chem. Soc.*, **2004**, *126*, 505–515.
- 29 V. Barone and M. Cossi, *J. Phys. Chem. A*, **1998**, *102*, 1995–2001.
- 30 P. Walter, J. Metzger, C. Thiel and V. Helms, *PLoS One*, **2013**, *8*, e58583.
- 31 P. E. M. Siegbahn and T. Borowski, *Acc. Chem. Res.*, **2006**, *39*, 729–738.
- 32 P. E. M. Siegbahn and F. Himo, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, **2011**, *1*, 323–336.
- 33 P. Tao and H. Bernhard Schlegel, *J. Comput. Chem.*, **2010**, *31*, 2363–2369.

- 34 F. Maseras and K. Morokuma, *J. Comput. Chem.*, **1995**, *16*, 1170–1179.
- 35 M. Svensson, S. Humbel, K. Morokuma, *J. Chem. Phys.*, **1996**, *105*, 3654–3661.
- 36 S. Dapprich, I. Komáromi, K. S. Byun, K. Morokuma and M. J. Frisch, *J. Mol. Struct. THEOCHEM*, **1999**, *461–462*, 1–21.
- 37 T. Vreven, K. Morokuma, Ö. Farkas, H. B. Schlegel and M. J. Frisch, *J. Comput. Chem.*, **2003**, *24*, 760–769.
- 38 T. Vreven and K. Morokuma, *J. Comput. Chem.*, **2000**, *21*, 1419–1432.
- 39 P. Li, B. P. Roberts, D. K. Chakravorty and K. M. Merz, *J. Chem. Theory Comput.*, **2013**, *9*, 2733–2748.
- 40 H. M. Senn, and W. Thiel, *Angew. Chem. Int. Ed.* **2009**, *48*, 1198–1229.
- 41 (a) 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. (b) S. Metz, J. Kästner, A. Sokol, T. Keal, P. Sherwood, *WIREs 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**, *4*, 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) C. Gonzalez, H. B. Schlegel, *J. Chem. Phys.* **1989**, *90*, 2154. (b) C. Gonzalez, H. B. Schlegel, *J. Phys. Chem.* **1990**, *94*, 5523.