Insight into Wild–Type and T1372E TET2–mediated 5hmC oxidation using ab initio QM/MM calculations

Hedieh Torabifard^{\dagger} and G. Andrés Cisneros^{*,‡}

†Department of Chemistry, Wayne State University, Detroit, MI 48202

‡Department of Chemistry, University of North Texas, Denton, TX 76203

E-mail: andres@unt.edu

Biological names	Acronyms
Ten-eleven translocation 2	TET2
α -ketoglutarate	α –KG
5-methylcytosine	$5\mathrm{mC}$
5-hydroxy-methylcytosine	$5 \mathrm{hmC}$
5-formylcytosine	$5 \mathrm{fC}$
5-carboxylcytosine	$5 \mathrm{caC}$
Thymine DNA glycosylase	TDG
Base excision repair	BER
1-methyladenine	$1 \mathrm{meA}$

Table S1: Biological acronyms used in the paper.

Mulliken Spin densitysnapshot number 1snapshot number 3Fe (III)3.284.32Oxyl O0.50-0.69

Electron configuration

 $^{IS}Fe(III) - O_F$

 $^{HS}Fe(III) - O_{AF}$

Table S2: Mulliken spin densities and electron configurations for the reactant of the snapshot number 1 and 3 in the T1372E mutant.

Table S3: Comparison of the reaction and the first barrier energies for TET2, AlkB and ABH2.

	ΔE barrier (kcal/mol)	Δ E reaction (kcal/mol)
AlkB	23.2	-3.7
AlkBH2	25.7	-3.5
TET2	20.1	5.4

Table S4: Comparison of Mulliken spin densities of important atoms involved in the oxidation in TET2, AlkB, ABH2.

Mullikon anin donaitioa	Fo	O(Orrul)	C(1meA)/	H(1meA)/	
Mulliken spin densities	ге	O(OXyI)	O(hmC)	H(hmC)	
AlkB reactant	3.26	0.54	0.00	0.00	
AlkBH2 reactant	3.20	0.61	0.00	0.00	
TET2 reactant	3.25	0.55	0.00	0.00	
AlkB I1	4.35	0.26	-0.92	0.005	
AlkBH2 I1	4.36	0.26	-0.92	0.004	
TET2 I1	4.35	0.25	-0.85	0.01	

Table S5: Mulliken spin densities for critical points of wild-type.

Mulliken Spin density	Reactant	TS1	I1	TS2	I2	TS3	Product
Fe (III)	3.25	4.13	4.35	4.35	4.36	4.37	3.83
Oxyl O	0.54	0.01	0.25	0.26	0.25	0.18	0.02
Hydroxyl O	0.00	-0.43	-0.86	-0.87	-0.85	-0.74	0.00
Hydroxyl H	0.00	0.04	0.01	0.01	0.01	0.01	0.00
Methylene C	0.00	0.03	0.05	0.06	0.05	0.01	0.00
Methylene H1	0.00	-0.04	-0.06	-0.08	-0.11	-0.14	0.00
Methylene H2	0.00	-0.04	-0.10	-0.09	-0.06	-0.05	0.00

	Volume	Population		Volume	Population		Volume	Population
I^{IS} Fe ^{III} –O _F reactant			H_{s}^{*}	$^{S}\mathrm{Fe}^{III}\mathrm{-O}_{A}$	$_{AF}$ TS1	Н	$^{HS}\mathrm{Fe}^{III}-\mathrm{O}_{AF}$ I1	
C(Fe)	0.29	10.07	C(Fe)	0.30	10.08	C(Fe)	0.27	9.92
C(Fe)	6.25	3.50	C(Fe)	3.13	1.88	C(Fe)	3.46	2.01
C(Fe)	4.39	2.76	C(Fe)	4.57	2.51	C(Fe)	4.67	2.63
C(Fe)	3.21	2.38	C(Fe)	4.63	2.78	C(Fe)	5.42	2.95
C(Fe)	4.86	3.36	C(Fe)	3.29	2.27	C(Fe)	3.19	2.21
C(Fe)	3.92	2.45	C(Fe)	4.16	2.70	C(Fe)	11.23	8.10
C(Fe)	4.35	3.30	C(Fe)	3.85	2.51	C(O)	0.22	2.00
C(O)	0.22	2.02	C(Fe)	4.53	3.13	V(O)	41.66	3.60
V(O)	67.54	6.65	C(Fe)	43.29	4.00	V(O)	30.20	3.22
V(O)	11.53	1.23	C(O)	0.22	2.01	V(H,O)	24.27	1.18
			V(O)	34.20	3.46			
			V(O)	1.27	0.23			

Table S6: ELF basins of the iron-oxyl moiety in the reactant, TS1 and I1.



Figure S1: Relative QM/MM energies of the selected snapshots from the MD simulations in the wild–type and the T1372E mutant.

וי ת	$\Delta\Delta E$	$\Delta \Delta E$	$\Delta\Delta E$	$\Delta \Delta E$	וי ת ו	$\Delta \Delta E$	$\Delta\Delta E$	$\Delta\Delta E$	$\Delta\Delta E$
Residue	(TS1)	(TS2)	(TS3)	(pro)	Residue	(TS1)	(TS2)	(TS3)	(pro)
K1142	_	_	-1.1	_	E1318	-1.0	_	_	_
E1144	_	1.1	_	_	K1321	_	1.5	_	_
K1173	_	_	1.1	1.2	E11323	_	_	_	-1.2
E1186	—	1.2	—	—	N1347	—	1.5	—	1.4
K1188	-1.1	—	-1.2	—	C1374	—	1.4	—	1.9
C1193	—	1.0	—	—	D1376	—	—	-1.1	—
K1208	—	—	—	1.1	R1383	-1.3	—	-1.4	—
D1242	—	—	-1.0	—	G1391	1.0	1.1	—	—
E1247	—	-1.1	-1.5	—	T1393	1.3	1.6	1.8	—
E1250	—	—	-1.3	—	K1409	—	—	—	1.8
R1253	—	—	1.4	—	E1411	—	—	1.4	
T1257	—	—	-1.0	-1.0	K1439	—	1.2	—	—
T1259	—	—	1.7	-1.9	R1455	1.7	1.1	1.4	1.8
N1260	-1.2	-1.2	-1.4	-1.1	K1462	-1.1	-1.2	—	—
R1262	-1.9	—	—	1.2	G^{**}	-1.3	-1.5	—	-1.6
C1263	—	-1.3	—	—	E1874	—	-1.4	—	-1.7
E1267	—	1.9	1.3	—	K1877	—	-1.1	—	—
E1268	1.2	1.0	—	—	R1878	—	-1.1	—	—
R1269	-1.4	1.1	—	—	E1879	1.4	1.2	1.1	—
C1271	—	1.2	—	—	R1896	1.5	1.2	—	—
S1290	—	—	—	1.2	V1900	—	1.0	—	1.6
K1299	_	_	-1.0	_	Y1902	1.2	1.3	_	1.3
R1302	—	—	1.8	1.2	K1905	—	-1.0	-1.1	—
K1308	_	_	_	1.2	K1920	_	_	_	1.2
K1310	-1.3	-1.1	—	—	E1923	—	—	-2.2	—

Table S7: Residues with more than 1 kcal/mol difference by EDA for TS1/2/3/product with respect to the reactant. All values are in kcal/mol. G^{**} is one of the linkers.



Figure S2: Electron configurations for the iron-oxyl moiety in the quintet state.



Figure S3: Optimized geometries of critical structures for the T1372E mutant. The black circle shows the hydrogen bond in the reactant and product.



Figure S4: Difference of total, Coulomb and vdW energies of a) I1, b) I2, c) product and reactant.

					K1142	
SCORE=56			CD 106N021157F72	1114		1196
* BAD AVG GOOD			sp Q4JK59 mTET2	1037	DTPGE-Q-SONGKCEGCNPDKDEAPYYTHLGAGPDVAAIRLIMEERYGEKGKAIRIEKVIIIGAEGKSSQGCPIA	1109
*			sp Q8NFU7 hTET1	1405	NPTKNLV <mark>SITKDSEL</mark> PTCSCLDRVIQKDKGPYYTHLGAGPSVAAVREIMENRYGQKGNAIRIEIVVYTGKEGKSSHGCPIA	1485
sp Q6N021 hTET2	÷ .	48	sp Q3URK3 mTET1 sp Q43151 bTET3	1354	KPAKNLIAGLKEQEAAPCDC-DGGTQKEKGPYYTHLGAGPSVAAVKELMETRFGQKGKAIRIEKIVFTGKEGKSSQCCPVA Tptkslintpakbaoefptcncvcoivekdegyvythlgagptvastbelimeerygkgkairiekuvytgkegksgccpia	1433
sp 04JK59 mTET2 sp 08NFU7 hTET1		50	sp Q8BG87 mTET3	681	TPTKSLLDTPAKKAQ2EFPTCDCVEQIVEKDEGPYYTHLGSGPTVASIRELMEDRYGEKGKAIRIEKVIYTGKEGKSSRGCPIA	764
sp Q3URK3 mTET1	:	56			R1261	
sp 043151 hTET3	5	62 62	cons	1933	: * :*:*****	2016
cons	÷.	5			R1253 T1259 R1262 E1279	
			spl06N021 hTET2	1197	KWVVRRSSSEEKLLCLVRERAGHTCEAAVIVILILVWEGIPLSLADKLYSELTETLKY-GTLTNRRCALNEERTCACOGLDPE	1279
			sp Q4JK59 mTET2	1110	KWVYRRSSEEEKLLCLVRVRPNHTCETAVMVIAIMLWDGIPKLLASELYSELTDILGKC-GICTNRRCSQNETRNCCCQGENPE	1192
			sp Q8NFU7 hTET1	1486	KWVLRRSSDEEKVLCLVRQRTGHHCPTAVMVVLIMVWDGIPLPMADRLYTELTENLKSYNGHPTDRRCTLNENRTCTCQGIDPE	1569
			sp 043151 hTET3	757	KWVIRRSGFERLICLVRERVDHCSIAVIVIIILWEGIFRSLGDELYOELTDELKSISGFIDKCLINKKICLOGUDFR	839
			sp Q8BG87 mTET3	765	KWVIRRHTLEEKLLCLVRHRAGHHCQNAVIVILILAWEGIPRSLGDTLYQELTDTLRKY-GNPTSRRCGLNDDRTCACQGKDPN	847
				0017		0100
			cons	2017		2100
					K1299 K1321	
			sp Q6N021 hTET2	1280	TCGASFSFGCSWSMYYNGCKFARSKIPRKFKLLGDDPKEEEKLESHLQNLSTLMAPTYKKLAPDAYNNQIEYEHRAPECRLGLK	1363
			sp Q4JK59 mTET2	1193	${\tt TCGASFSFGCSWSMYYNGCKFARSKKPRKFRLHGAEPKEEERLGSHLQNLATVIAPIYKKLAPDAYNNQVEFEHQAPDCCLGLK}$	1276
			sp Q8NFU7 hTET1 sp O3URK3 mTET1	1570	TCGASFSFGCSWSMYENGCKFGRSEDFRFRIDFSSPLHEKNLEDNLØSLATKLAFIYKØYAFVAYØNÖVEYEVARECRLGSK TCGASFSFGCSWSMYENGCKFGRSENPEKFRLADNYPLHEKOLEKNLØELATVLAPLYKØMAPVAYØNÖVEYEEVAGDCRLGNE	1653
			sp 043151 hTET3	840	TCGASFSFGCSWSMYFNGCKYARSKTPRKFRLAGDNPKEEEVLRKSFQDLATEVAPLYKRLAPQAYQNQVTNEEIAIDCRLGLK	923
			sp Q8BG87 mTET3	848	TCGASFSFGCSWSMYFNGCKYARSKTPRKFRLTGDNPKEEEVLRNSFQDLATEVAPLYKRLAPQAYQNQVTNEDVAIDCRLGLK	931
			CODE	2101	H1386	2184
			cons	2101		2104
					K1383 N1387 N1403 K1409	
			sp Q6N021 hTET2	1364	EGRPFSGVTACLDFCAHAHRDLHNMQNGSTLVCTLTREDNREFGGKPEDEQLHVLPLYKVSDVDEFGSVEAQEEKKRSGAIQVL	1447
			sp Q4JK59 mTET2 sp O8NFU7 bTET1	1277	EGRPFSGVTACLDFSAHSHRDOONMPNGSTVVVTLNREDNREVGAKPEDEOFHVLPMYIIAPEDEFGSTEGOEKKIRMGSIEVL EGRPFSGVTACLDFCAHDHBNIHNMNNGSTVVVTLNREDNSIGUTDODEOLHVLDIVKISDTDEFGSKEGGREAKIKSGIEVL	1360 1737
			sp Q3URK3 mTET1	1602	EGRPFSGVTCCMDFCAHSHKDIHNMHNGSTVVCTLIRADGRDTN-CPEDEQLHVLPLYRLADTDEFGSVEGMKAKIKSGAIQVN	1684
			sp 043151 hTET3	924	EGR DFGAUTACMDFCAHAHKDOHNLYNGCTVVCTLTKEDNRCVGKI PEDEOLHULPLYKMANTDEFGSEENONAKVGSGAIOVL	1007
			sp Qobdo / ImiEis	332	BOKELSOVIACHDECHDAUKDUNDUNDUNDCIVVCIDIKEDHACVOQIEEDEQUNVLEDIKMASIDEEOSEENQMAKVSSOATQVI	1015
			cons	2185	*****:***.*:**.*:*:*:*:*:*:*:*:*:: : *:*:*:: : *:**:*::*:	2268
				1440		1 5 1 0
			sp Q4JK59 mTET2	1361	OSFRARYIRIGEDK-SCHKA-EPKKAKTKAARKSSLENCSSRFEKGKSSHTKMENASHKOMTAOPOLSG	1435
			sp Q8NFU7 hTET1	1738	APRRKK-RTCFTQPVPRSGKKRAAMMTEVLAHKIRAVEKKP-IP-RIKRKNNSTTTNNSKPSSLPTL	1801
			sp Q3URK3 mTET1 sp Q43151 hTET3	1685	GPTRKR-RLRFTEPVP-RCGKRAKMKONHNKSGSHNTKSFSSASSTSHLVKDESTDFCPLOASSAET TAFDRE-VRRLPEDAK-SCROBOLEARKAAEKKKICKEKLSTERIKOEALELAGITSDELSLKG	1749
			sp Q8BG87 mTET3	1016	TAFPRE-VRRLPEPAK-SCRORQLEARKAAAEKKKLOKEKLSTPEKIKOEALELAGVTTDPGLSLKG	1080
			cons	2269		2352
					E1874	
			sp Q6N021 hTET2	1819	LSDANGQEKQPLALVQGVASGAEDNDEVWSDSEOSFLDPDIGGVAVAPTHGSILIECAKRELHATTPLK	1887
			sp Q4JK59 mTET2	1736	LTG-SSQEKQPEVSGQDAAAVQEIEYWSDSEHNFQDPCIGGVAIAPTHGSILIECAKCEVHATTKVN	1801
			sp Q8NFU7 hTET1	1962		2034
			sp 043151 hTET3	1464	EEKLWDPFSLEEGP-AEEPPSKGAVKEEKGGGGAEEEEELWSDSEHNFLDENIGGVAVAPAHGSILIECARRELHATTPLK	1544
			sp Q8BG87 mTET3	1472	EEKLWDPFSLEEGT-AEEPPSKGVVKEEKSGPTVEEDEEELWSDSEHNFLDENIGGVAVAPAHCSILIECARRELHATTPLK	1552
			CODS	2773	* ***** * * ***** * * *****	2856
				2,75		2000
					R1896 E1923 🕌 K1924	
			sp Q6N021 hTET2	1888	NPNRNHPTRISLVFYQHKSMNEPKHGLALWEAKMA <mark>EKAREKE</mark> <mark>EECEKYGPDYVPQ</mark> KSHGKKVKREPAEPHETS	1960
			sp Q4JK59 mTET2 sp O8NFU7 hTET1	1802	DPDRNHPTRISLVLYRHKNLFLPKHCLALWEAKMAEKAR-KEEECGKNGSDHVSOKNHGKQEKREPTGPQ HPNRNHPTRLSLVFYOHKNLNKPOHGFELNKIKFEAKEAKNKKMKASEOKDOAANEGPEOSSE	1870
			sp Q3URK3 mTET1	1914	SPKRGVPFRVSLVFYQHKSLNKPNHGFDINKIKCKCKKVTKKKPADRECPDVSPE	1968
			sp 043151 hTET3	1545	KPNRCHPTRISLVFYOHKNLNOPNHGLALWEAKMKOLAERARARO	1617
			spigones/imter3	1003	VENNOREINISTALIÄUVNINÄLNUGTATMEAVMVÄTAEVANÄKÄREVAKTOTOÄÄEVYTIOKKKKMOGAMAV	1023
			CODS	2857	* * * * * * * * * * * * * * * * * * * *	2940

Figure S5: Protein sequence alignment of h/m-TET1-3 created by T-coffee (J. Mol. Biol., 13, 205) The residues with significant change in stabilizing energy ($|\Delta\Delta E| \ge 2 \text{ kcal/mol}$) are shown with red arrow. Our results show these residues are important for the catalytic function of this family of enzymes and could be interesting targets for experimental mutagenesis studies.