Supporting Information

Exploration of the Substrate Preference of Lysine Methyltransferase SMYD3 by Molecular Dynamics Simulations

Jixue Sun[1], Fandi Shi[1] and Na Yang[1*]

[1]*State Key Laboratory of Medicinal Chemical Biology, College of Pharmacy and Key Laboratory of Medical Data Analysis and Statistical Research of Tianjin, Nankai University, Tianjin 300353, China*

*Corresponding author: Tel/Fax: + 8622 85358193.

E-mail: yangnaNKU@nankai.edu.cn (N. Yang).

SI Materials and methods

In vitro methyltransferase activity assay

In vitro methyltransferase activity of SMYD3 was measured using a continuous, enzyme-coupled microplate assay as described before[1]. In brief, a 100 μL assay mixture contains 200 μM SAM, 10 μM SMYD3, 10 μM SAH nucleosidase, 1 μM adenine deaminase was prepared in a buffer containing 20 mM Tris 8.0, 150 mM NaCl and 1mM $MnSO_4$. The reaction was initiated by adding 100 μM substrate peptide when absorbance at 265 nm reached equilibrium at 37 °C examined by a SpectraMax i3x multi-Mode Microplate Reader (Molecular Devices). The substrate consumption Δc was calculated by the change of absorbance at 265 nm based on the $\varepsilon 260_{nm}$ of SAH. K_{cat} value is calculated by Δc/t[e] (t is reaction time and [e] is the concentration of substrate). Three parallel experiments were performed and the K_{cat} of SMYD3 on WT or F258L of MAP3K2 peptide are listed below. From the results we can see the average K_{cat} of SMYD3 are similar though using different substrates.

Calculation of averaged RMSD value of key residues of SMYD3

The cpptraj module[2] of AmberTools18 was used to calculate averaged RMSD of the key residues both in the active site (Y124, N132, F183, Y239 and F259) and in the P-2 pocket (S101, L104, V178, I179, S182 and V195) of SMYD3 in each system during the whole MD simulation with a 10-ps interval, respectively.

Reference:

(1) Wayllace, N. Z.; Valdez, H. A.; Merás, A.; Ugalde, R. A.; Busi, M. V.; Gomez-Casati, D. F. An Enzyme-Coupled Continuous Spectrophotometric Assay for Glycogen Synthases. *Mol. Biol. Rep.* **2012**, *39* (1), 585–591.

(2) Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9* (7), 3084–3095.

SI figures and tables

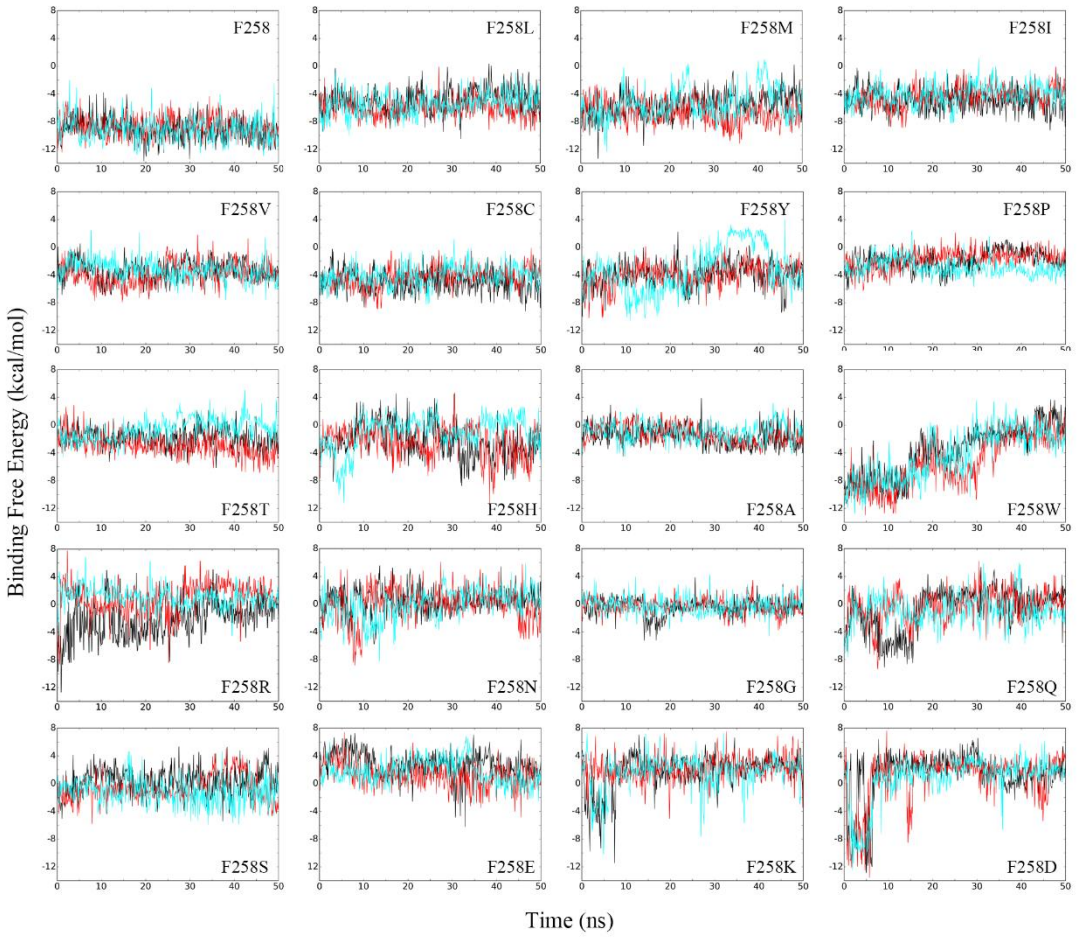

Figure S1. Time evolutions of binding free energy between SMYD3 and F258 or its substitutions during MD simulations in three repetitions (black, repetition 1; red repetition 2; cyan, repetition 3).

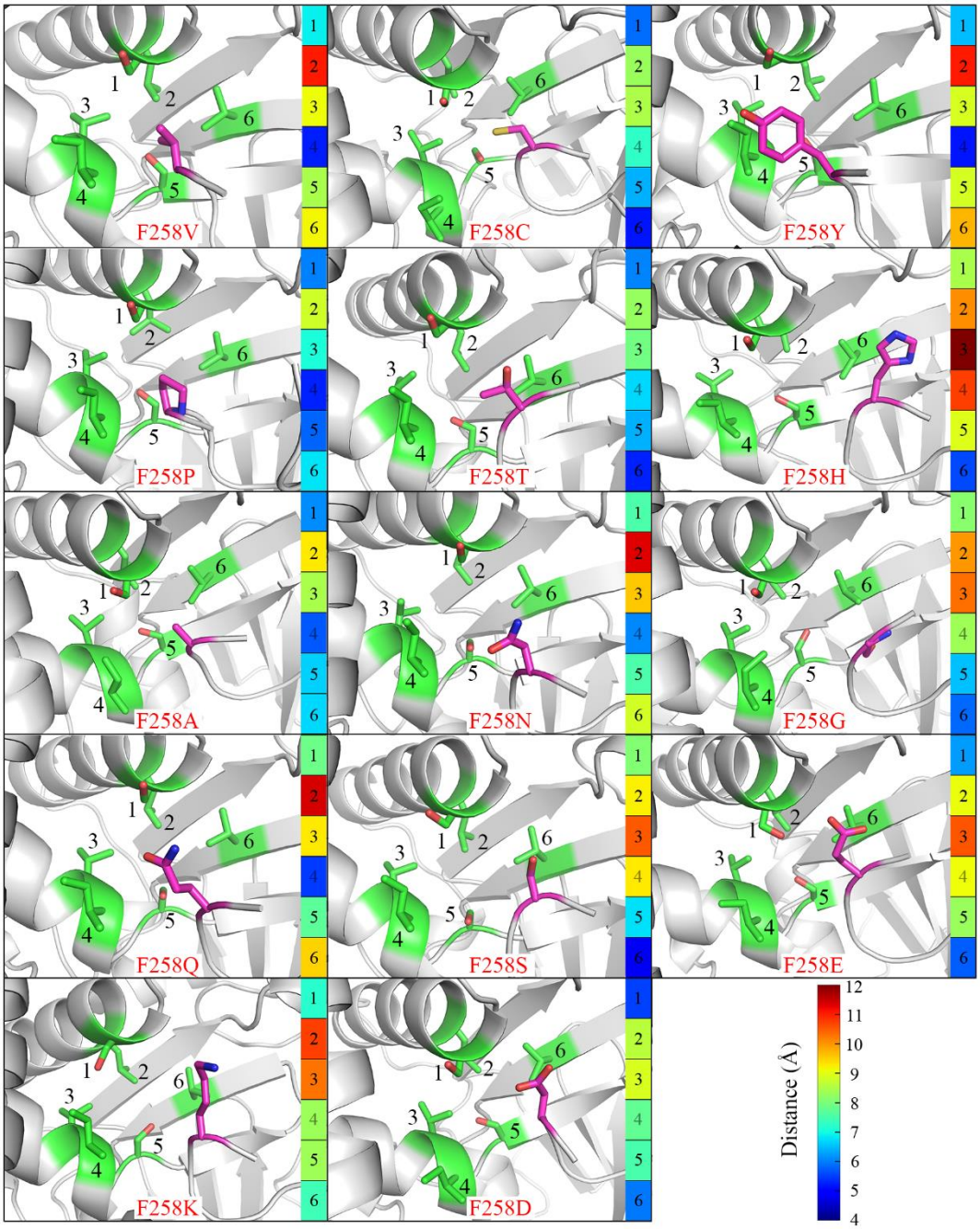

Figure S2. The conformations of other F258 substitutions after MD simulations. The distances between F258 mutations and the marked residues in SMYD3 are shown in the color map. The color scheme of residues is the same as Figure 3.

Figure S3. Time evolutions of binding free energy between SMYD3 and F258 modifications during MD simulations in three repetitions (black, repetition 1; red repetition 2; cyan, repetition 3).

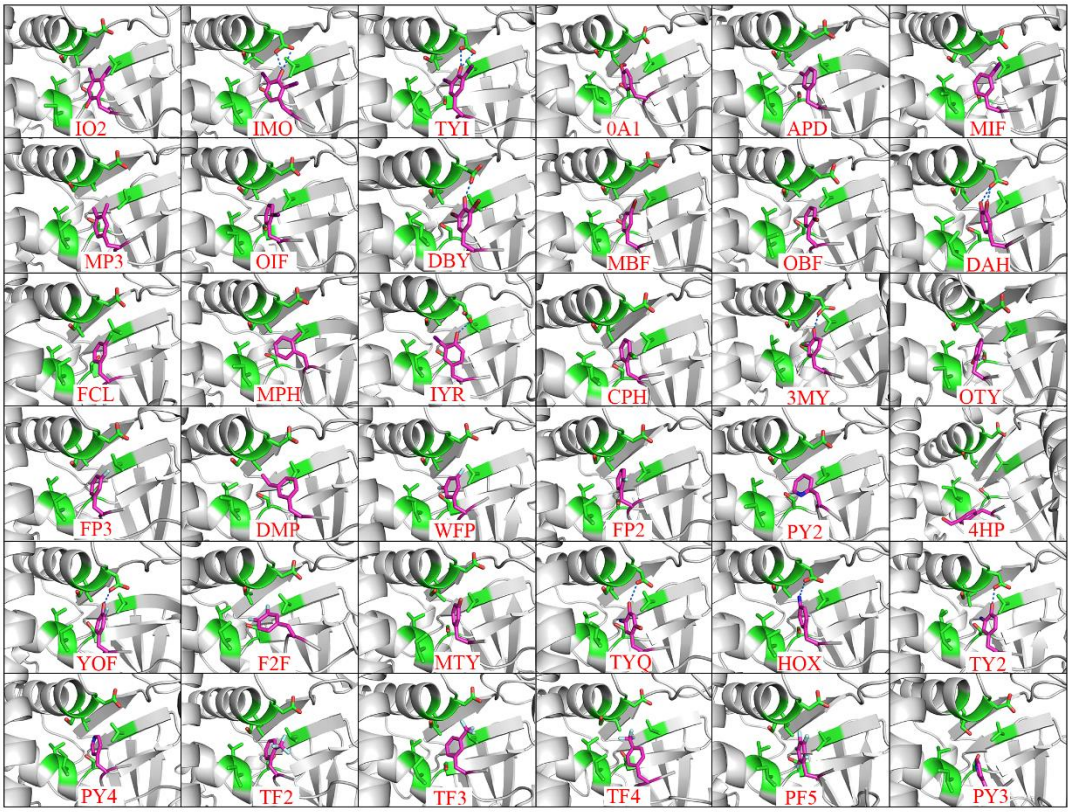

Figure S4. The conformations of other F258 modifications after MD simulations. The color scheme of residues is the same as Figure 4. Hydrogen bond is depicted as dash line if it exists.

Table S1. Averaged binding free energies and standard deviations between SMYD3 and wild-type L829, or L829F substitution of VEGFR1 peptide from three times-MD simulations, respectively. All binding free energies are in kcal/mol.

Residue type	Average	Std. Dev.	Repetition		
			1	2	3
LEU	-4.49	0.30	-4.37	-4.83	-4.28
PHE	-8.52	0.18	-8.60	-8.31	-8.65

Table S2. K_{cat} value of SMYD3 catalyzing on different substrates of MAP3K2 peptide or mutant. The values are in hr^{-1}.

Substrate	Average	Std. Dev.	Repetition		
			1	2	3
MAP3K2 F258	5.51	0.85	4.70	5.45	6.39
F258L mutant	5.13	0.58	4.55	5.13	5.71

Table S3. Averaged RMSD value of the key residues both in the active site (Y124, N132, F183, Y239 and F259) and in the P-2 pocket (S101, L104, V178, I179, S182 and V195) of SMYD3 in each system during MD simulation. Values in parentheses are standard deviations.

Residue type	RMSD of residues in the active site		
	1	2	3
PHE	0.69 (0.15)	0.79 (0.18)	0.59 (0.12)
LEU	0.72 (0.11)	0.85 (0.19)	0.61 (0.10)
MET	0.89 (0.14)	0.81 (0.28)	0.89 (0.17)
ILE	0.88 (0.36)	0.70 (0.11)	0.78 (0.12)
VAL	0.99 (0.21)	0.58 (0.11)	0.71 (0.12)
CYS	0.66 (0.10)	0.79 (0.21)	0.72 (0.15)
TYR	0.70 (0.15)	0.99 (0.27)	0.74 (0.11)
PRO	0.76 (0.11)	0.84 (0.11)	0.63 (0.10)
THR	0.94 (0.21)	0.69 (0.11)	0.71 (0.15)
HIS	0.69 (0.15)	0.66 (0.13)	0.64 (0.12)
ALA	0.66 (0.13)	0.73 (0.19)	0.70 (0.16)
TRP	0.98 (0.12)	0.70 (0.09)	0.92 (0.14)
ARG	0.62 (0.08)	0.72 (0.15)	0.71 (0.15)
ASN	0.60 (0.10)	0.62 (0.11)	1.03 (0.22)
GLY	0.98 (0.16)	0.68 (0.15)	0.76 (0.12)
GLN	0.76 (0.17)	0.80 (0.27)	0.65 (0.12)
SER	0.70 (0.12)	0.76 (0.12)	0.87 (0.20)
GLU	0.76 (0.12)	0.68 (0.12)	0.67 (0.15)
LYS	0.69 (0.17)	0.89 (0.19)	0.81 (0.12)
ASP	0.76 (0.17)	0.89 (0.19)	0.81 (0.12)
Residue type	RMSD of residues in the P-2 pocket		
	1	2	3
PHE	0.79 (0.12)	0.83 (0.13)	0.83 (0.12)
LEU	0.92 (0.15)	0.86 (0.19)	0.79 (0.11)
MET	0.85 (0.13)	0.97 (0.15)	0.86 (0.13)
ILE	0.81 (0.11)	0.78 (0.10)	0.88 (0.11)
VAL	0.95 (0.11)	0.94 (0.10)	0.73 (0.09)
CYS	0.82 (0.10)	0.89 (0.09)	0.90 (0.15)
TYR	0.86 (0.13)	0.84 (0.11)	0.90 (0.11)
PRO	0.95 (0.14)	0.86 (0.11)	0.75 (0.10)
THR	1.01 (0.27)	0.90 (0.08)	0.85 (0.09)
HIS	0.90 (0.14)	1.21 (0.12)	0.89 (0.12)
ALA	0.81 (0.10)	0.88 (0.15)	0.84 (0.12)
TRP	1.00 (0.12)	1.12 (0.10)	1.05 (0.17)
ARG	1.19 (0.13)	1.00 (0.12)	0.95 (0.12)
ASN	0.91 (0.22)	0.92 (0.12)	1.09 (0.18)
GLY	0.85 (0.11)	0.89 (0.18)	1.01 (0.10)
GLN	0.99 (0.15)	1.08 (0.19)	0.84 (0.13)

SER	1.00 (0.11)	0.88 (0.14)	0.80 (0.08)
GLU	0.96 (0.14)	0.85 (0.14)	0.85 (0.13)
LYS	0.94 (0.16)	0.94 (0.14)	0.94 (0.15)
ASP	1.02 (0.17)	1.16 (0.11)	1.05 (0.11)

Table S4. Averaged binding free energies and standard deviations between SMYD3 and F258 modification from three times-MD simulations, respectively. All binding free energies are in kcal/mol.

Name	Substituent group/Position					Avg.	Std. Dev.	Repetition		
	2	3	4	5	6			1	2	3
I34[a]	-	-I	-I	-	-	-12.79	0.68	-12.53	-13.57	-12.28
ALC[b]	cyclohexyl sidechain					-12.64	0.80	-11.93	-13.51	-12.48
IO2[a]	-OH	-	-I	-I	-	-12.31	0.27	-12.62	-12.15	-12.16
B34[a]	-	-Br	-Br	-	-	-12.29	1.37	-13.25	-10.73	-12.90
BOH[a]	-OH	-Br	-Br	-	-	-12.14	0.44	-11.70	-12.59	-12.12
CP3	-	-Cl	-Cl	-	-	-12.13	1.49	-10.43	-12.84	-13.13
BO2[a]	-OH	-	-Br	-Br	-	-11.49	1.18	-11.32	-12.74	-10.40
IMO[a]	-I	-OH	-I	-	-	-11.26	0.89	-11.07	-10.48	-12.24
CPA[b]	cyclopentyl sidechain					-10.82	0.79	-9.98	-11.56	-10.92
TYI	-	-I	-OH	-I	-	-10.73	1.08	-10.07	-11.99	-10.15
PHI	-	-	-I	-	-	-10.53	0.50	-11.02	-10.02	-10.54
4BF	-	-	-Br	-	-	-10.50	0.88	-10.32	-9.72	-11.45
TBP	-	-	$-C(CH_3)_3$	-	-	-10.46	0.43	-10.72	-10.68	-9.97
BMO[a]	-Br	-OH	-Br	-	-	-10.40	0.16	-10.57	-10.38	-10.24
4PH	-	-	$-CH_3$	-	-	-10.17	0.81	-11.03	-10.03	-9.44
0A1	-	-	$-OCH_3$	-	-	-10.07	0.54	-9.96	-10.65	-9.59
F20	-	-	-Cl	-	-	-10.04	1.67	-11.87	-8.58	-9.66
CP2	-Cl	-	-Cl	-	-	-9.97	0.56	-10.59	-9.51	-9.82
APD	-	$-CH_3$	-	-	-	-9.88	0.37	-9.48	-10.21	-9.96
MIF[a]	-	-I	-	-	-	-9.82	0.15	-9.80	-9.98	-9.68
MP3	-	$-CH_3$	$-CH_3$	-	-	-9.80	1.46	-9.24	-8.70	-11.45
OIF[a]	-I	-	-	-	-	-9.69	0.60	-9.20	-10.36	-9.50

DBY	-	-Br	-OH	-Br	-	-9.67	0.16	-9.60	-9.85	-9.56
MBF[a]	-	-Br	-	-	-	-9.66	1.14	-9.09	-10.97	-8.94
PHE	-	-	-	-	-	-9.64	0.13	-9.64	-9.52	-9.77
OBF[a]	-Br	-	-	-	-	-9.52	0.74	-8.67	-10.06	-9.81
PFF	-	-	-F	-	-	-9.51	1.95	-7.65	-9.35	-11.54
DAH	-	-OH	-OH	-	-	-9.47	0.35	-9.17	-9.86	-9.38
FCL	-	-Cl	-	-	-	-9.46	0.92	-9.07	-8.80	-10.51
MPH	$-CH_3$	-	-	-	-	-9.45	0.62	-9.95	-8.76	-9.63
IYR	-	-I	-OH	-	-	-9.35	0.94	-9.96	-9.83	-8.27
CPH	-Cl	-	-	-	-	-8.96	0.69	-8.17	-9.46	-9.24
3MY	-	-Cl	-OH	-	-	-8.92	0.41	-9.38	-8.72	-8.65
OTY	-OH	-	-	-	-	-8.91	0.41	-8.45	-9.04	-9.24
FP3	-	-F	-	-	-	-8.65	1.26	-10.09	-7.79	-8.06
DMP	-	$-C_2H_5$	-	-	-	-8.32	1.10	-7.27	-9.47	-8.23
WFP	-	-F	-	-F	-	-7.72	0.77	-8.51	-6.98	-7.66
FP2	-F	-	-	-	-	-7.63	1.44	-8.47	-5.97	-8.45
PY2[c]	-N-	-	-	-	-	-7.43	0.61	-7.15	-8.14	-7.01
4HP	-	-	$-CH_2OH$	-	-	-7.24	0.81	-7.59	-6.32	-7.82
YOF	-	-F	-OH	-	-	-7.22	0.64	-6.53	-7.79	-7.33
F2F	-	-F	-F	-	-	-6.93	1.11	-5.71	-7.87	-7.20
MTY	-	-OH	-	-	-	-6.70	1.77	-7.14	-4.75	-8.21
TYQ	-OH	-	-OH	$-NH_2$	-	-6.32	0.53	-5.78	-6.37	-6.83
HOX	-	-	$-NH_2$	-	-	-6.23	0.11	-6.28	-6.31	-6.11
TY2	-	$-NH_2$	-OH	-	-	-5.56	0.57	-4.97	-5.59	-6.11
PY4[c]	-	-	-N-	-	-	-5.23	1.16	-6.07	-3.90	-5.71
TF2	$-CF_3$	-	-	-	-	-5.19	0.24	-4.99	-5.13	-5.45

TF3	-	$-CF_3$	-	-	-	-4.34	0.76	-3.50	-4.97	-4.54
TF4	-	-	$-CF_3$	-	-	-4.18	0.84	-3.21	-4.76	-4.55
PF5	-F	-F	-F	-F	-F	-3.63	1.08	-4.74	-2.59	-3.56
PY3[c]	-	-N-	-	-	-	-2.61	1.30	-1.36	-3.96	-2.51

[a] These phenylalanine derivatives were modified from 4BF by changing substituted groups.

[b] Two alanine derivatives in the database were selected with cyclohexyl and cyclopentyl sidechains.

[c] The substituted group is pyridyl group.