Manipulating Unconventional CH-based Hydrogen Bonding in a Methyltransferase via Non-Canonical Amino Acid Mutagenesis

Supporting Information

Scott Horowitz, Upendra Adhikari, Lynnette M.A. Dirk, Paul A. Del Rizzo, Ryan A. Mehl, Robert L. Houtz, Hashim M. Al-Hashimi, Steve Scheiner, and Raymond C. Trievel

Supplementary Text:

Protonation Status of Residue 335 in SET7/9

One consideration in substituting the active Tyr335 in SET7/9 with the non-canonical amino acid (ncAA) pAF is the possibility of introducing a positive charge into the enzyme's active site by the protonation of the aniline side chain of pAF. It is worth noting that the pK_A value of the aniline ammonium cation has a value of 4.6[1], implying that the aniline side chain of pAF would exist in the deprotonated, uncharged state at the neutral to basic pH values at which the SET7/9 Y335pAF mutant was characterized. Nonetheless, it is possible that the side chains of tyrosine and pAF differ in charge in the active site environment despite their isostericity and the relatively low pK_A value of aniline[2, 3]. To examine this possibility, we analyzed the thermal stability, structure, and substrate binding properties of the SET7/9 Y335pAF mutant. Measurement of the dissociation constant of the TAF10 peptide substrate for WT SET7/9 and the Y335pAF mutant by isothermal titration calorimetry demonstrated that the pAF substitution did not disrupt peptide substrate binding (Supplementary Figure 1A and Table 1). This observation is important because the ε-ammonium cation of the Lys189 substrate in the TAF10 peptide is positioned within ~4 Å of the Tyr335 hydroxyl group and Y335pAF amine group (Figure 1B). Consequently, if the pAF side chain was protonated, it would presumably result in strong electrostatic repulsion with the Lys189 ε-ammonium cation, impairing the binding of the TAF10 peptide to the Y335pAF mutant. In addition, the differential scanning calorimetry experiments demonstrated that pAF substitution did not alter the thermal stability of the Y335pAF mutant compared to WT SET7/9 (Supplementary Figure 1B), consistent with a conservation of the charge at residue 335 in the active site. Finally, the 1.6 Å resolution crystal structure of the SET7/9 Y335pAF mutant illustrates that its overall active site structure is essentially unchanged compared to the structure of WT SET7/9 (Figure 1A-B). Based on these findings, we conclude that the Y335pAF mutation does not alter the charge at this position in the active site of SET7/9.

QM Modeling of the SET7/9 Active Site and the Effect of the Dielectric Constant on Selectivity

Using a combination of QM modeling and experiment, we re-engineered the hydrogen bonding network within the active site of methyltransferase SET7/9. Of course, the reduction of the entire enzymatic system to a four-unit complex cannot be expected to recapitulate the subtleties of this entire system, nor the dynamic motions of groups within the enzyme. The calculations discussed in the main text also do not include the polarizing effects of surrounding enzymatic groups, nor of the solvent. Nonetheless, the trends in the QM calculations mirrored the experimental data, facilitating an understanding of the underlying causes of the selectivity of this system.

It is well established that the properties of a sulfonium group, and particularly its hydrogen bond donating capabilities, may in part depend on its solvation[4, 5], and thus we performed tests to ensure that the predicted trends are robust through different dielectric environments. As may be seen in Table 1, the selectivity for $MeS^+(Et)_2$ over $S(Et)_2$ is 11.0 kcal mol^{-1} for phenol, 6.0 for aniline, and 5.3 for benzene. In other words, phenol has a higher selectivity than the other two respective species by 5.0 and 5.7 kcal mol^{-1}. These calculations were repeated within the context of a polarizable medium, so as to simulate the effects of the surrounding protein. When the environment is modeled by a dielectric constant of 4, a value that is commonly taken for a protein interior[6-8], these values drop slightly to 4.5 and 1.9 kcal mol^{-1}, retaining the higher selectivity of phenol. Even in the extreme situation of a very high dielectric

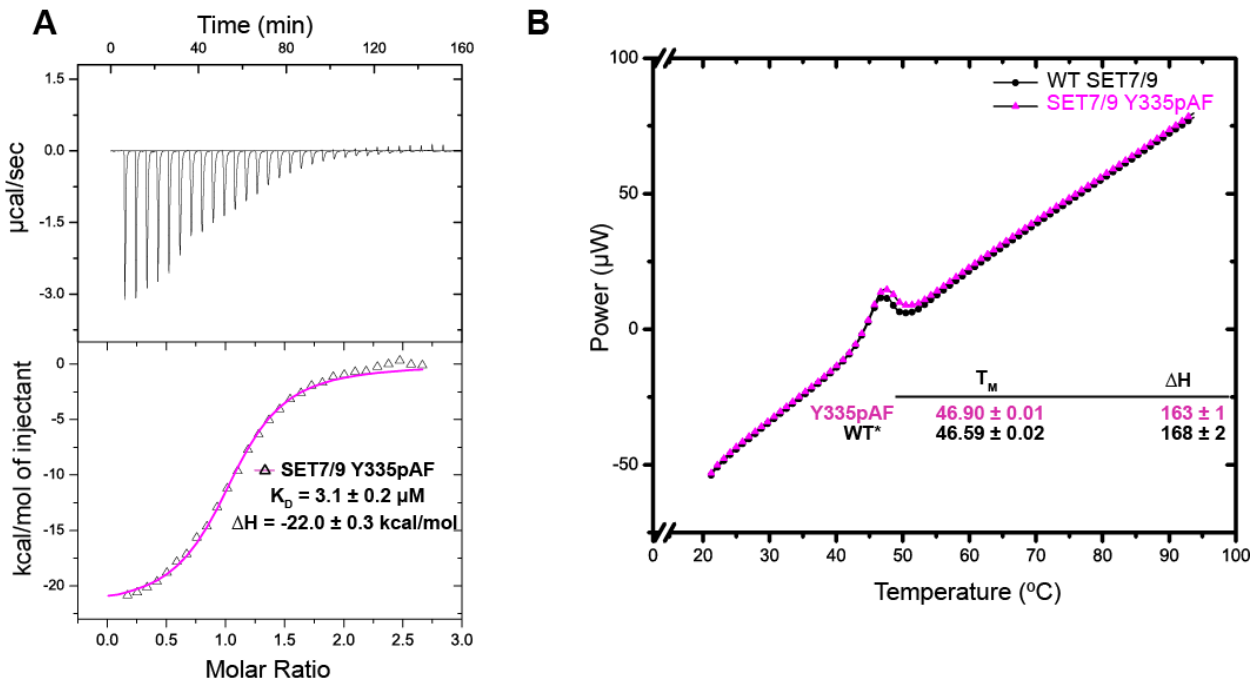
constant, viz. 80 as in water, the selectivity advantage of phenol over the other two species remains as high as 5.5 and 2.7 kcal mol^{-1}. It is concluded that the effects observed in vacuo are qualitatively unaffected by immersion of the system in an environment more akin to a protein interior.

Supplementary Table 1: Crystallographic and refinement statistics for the structure of SET7/9 Y335pAF•AdoHcy•TAF10 peptide complex

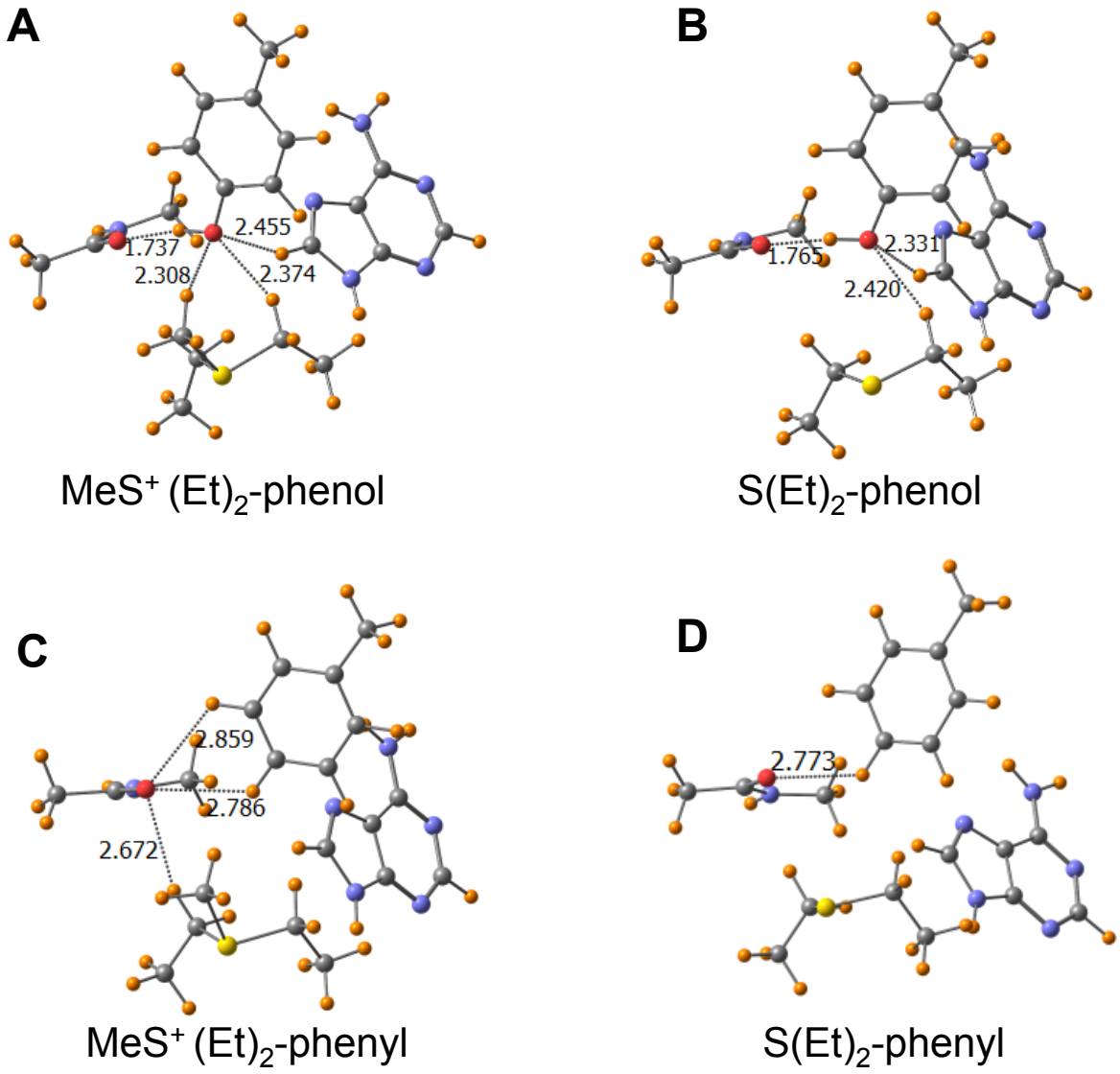
	SET7/9 Y335pAF• AdoHcy•TAF10 peptide
PDB Code	4J7F
Data collection	
Space group	$P3_221$
Cell dimensions	
a, *b*, *c* (Å)	83.0, 83.0, 95.5
α, *β*, *γ* (°)	90, 90, 120
Resolution (Å)	50.0-1.60 (1.64-1.60)
R_{merge} (%)	7.7 (44.2)
I/sI	26.7 (3.1)
Completeness (%)	99.4 (94.9)
Redundancy	10.7 (4.8)
Refinement	
Resolution (Å)	1.60
No. of Reflections	48,193
R_{work}/ R_{free}	0.20/0.22
No. of Atoms	1889
Protein	1941
Ligand/ion	75
Water	171
B-factors	23.9
Protein	23.4
Ligand/ion	34.9
Water	32
R.M.S Deviations	
Bond lengths (Å)	0.014
Bond angles (°)	1.51
MolProbity Score	
Percentile	84%
Resolution Range (Å)	1.59 ± 0.25
Ramachandran	
Favored (%)	95.4
Allowed (%)	4.6
Outliers (%)	0

Supplementary Table 2: Pairwise interaction energies (Y=phenol, aniline or benzene, X=$MeS^{+}(Et)_2$ or $S(Et)_2$) in kcal mol^{-1}

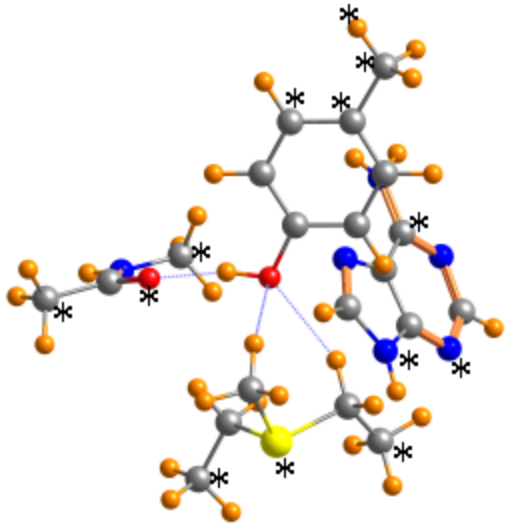
	$MeS^{+}(Et)_2$ complexes			$S(Et)_2$ complexes		
	Phenol	Aniline	Benzene	Phenol	Aniline	Benzene
Adenine---X	4.57	4.59	4.51	-0.50	-0.53	-0.54
Adenine---NMA	-1.44	-1.46	-1.45	-1.50	-1.50	-1.53
Adenine---Y	-2.11	-0.27	0.12	-2.11	-2.60	0.09
Y---NMA	-9.96	-2.08	-2.07	-10.47	-4.87	-2.43
Y---X	-9.36	-11.70	-2.76	-1.44	-2.02	-0.38
NMA---X	-10.15	-10.02	-9.94	-0.84	-0.78	-0.83
total	**-28.45**	**-20.94**	**-11.60**	**-16.86**	**-12.32**	**-5.62**
Minus ***opt*** monomers	**-27.93**	**-16.08**	**-10.83**	**-16.92**	**-10.05**	**-5.53**



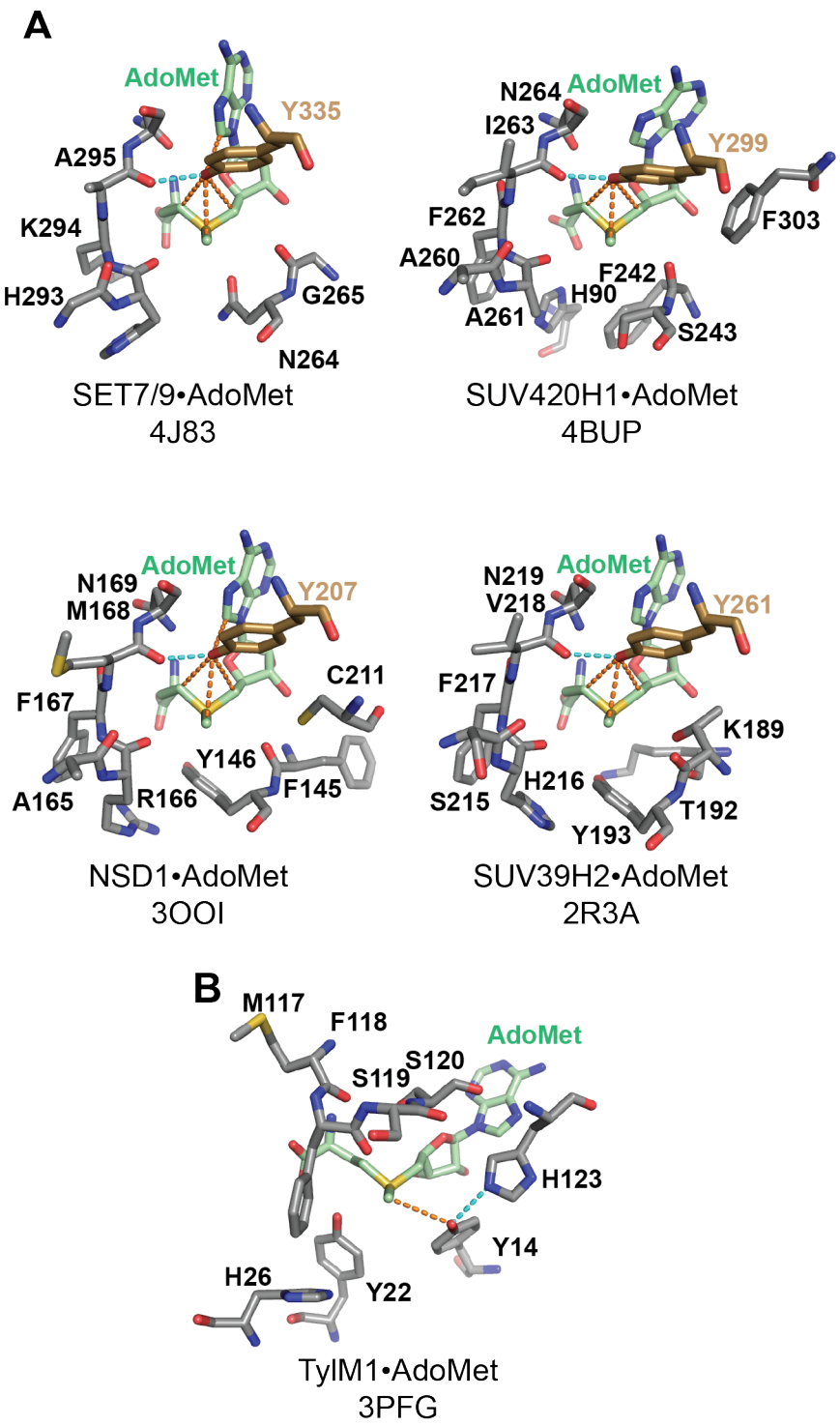
Supplementary Figure 1: Calorimetric analysis of the SET7/9 Y335pAF mutant. (A) Isothermal titration calorimetry of SET7/9 Y335pAF and the TAF10 peptide. (B) Differential scanning calorimetry melting curves of WT SET7/9 (black),*previously reported in[9] and Y335pAF mutant (magenta).



Supplementary Figure 2: QM optimized structures of the SET7/9 active site containing (A) $MeS^{+}(Et)_2$ and phenol, analogous to WT SET7/9 bound to AdoMet, (B) $S(Et)_2$ and phenol, analogous to WT SET7/9 bound to AdoHcy, (C) $MeS^{+}(Et)_2$ and phenyl, analogous to the SET7/9 Y335F mutant bound to AdoMet, and (D) $S(Et)_2$ and phenyl, analogous to the SET7/9 Y335F mutant bound to AdoHcy



Supplementary Figure 3: AdoMet complex where asterisks indicate atoms frozen into X-ray coordinates during optimization.



Supplementary Figure 4: Examples of tyrosine-mediated CH···O hydrogen bonding in methyltransferase active sites. (A) Position of the invariant tyrosine in the active sites of different SET domain lysine methyltransferases[9-12]. CH···O and NH···O hydrogen bonds formed by the tyrosine hydroxyl group are denoted by orange and cyan dashed lines, respectively. The invariant tyrosine is depicted in light brown. The PDB accession codes corresponding to each structure are shown parenthetically. (B) Hydrogen bonding by an active site tyrosine in TylM1, a class I (Rossmann fold) methyltransferase[13]. Hydrogen bonds are depicted as in panel A.

References

[1] Gross, K. C., and Seybold, P. G. (2000) Substituent effects on the physical properties and pKa of aniline, *Int J Quantum Chem 80*, 1107-1115.

[2] Gutbezahl, B., and Grunwald, E. (1953) The Effect of Solvent on Equilibrium and Rate Constants .2. The Measurement and Correlation of Acid Dissociation Constants of Anilinium and Ammonium Salts in the System Ethanol Water, *J Am Chem Soc 75*, 559-565.

[3] Bagno, A., and Terrier, F. (2001) Carbon and nitrogen basicity of aminothiophenes and anilines, *J Phys Chem A 105*, 6537-6542.

[4] Markham, G. D., and Bock, C. W. (1996) The interaction of water with sulfonium ions and the effects of hydration on the energetics of methyl group transfer: An ab initio molecular orbital study of the hydration of (CH3)(3) S+ and (CH3)(2) S+CH2CO2-, *Struct Chem 7*, 281-300.

[5] Adhikari, U., and Scheiner, S. (2013) Magnitude and mechanism of charge enhancement of CH..O hydrogen bonds, *J Phys Chem A 117*, 10551-10562.

[6] Wang, Z. X., and Duan, Y. (2004) Solvation effects on alanine dipeptide: A MP2/cc-pVTZ//MP2/6-31G** study of (Phi, Psi) energy maps and conformers in the gas phase, ether, and water, *J Comput Chem 25*, 1699-1716.

[7] Simonson, T., and Perahia, D. (1995) Internal and Interfacial Dielectric-Properties of Cytochrome-C from Molecular-Dynamics in Aqueous-Solution, *Proc Natl Acad Sci USA 92*, 1082-1086.

[8] Dwyer, J. J., Gittis, A. G., Karp, D. A., Lattman, E. E., Spencer, D. S., Stites, W. E., and Garcia-Moreno, B. (2000) High apparent dielectric constants in the interior of a protein reflect water penetration, *Biophys J 79*, 1610-1620.

[9] Horowitz, S., Dirk, L. M., Yesselman, J. D., Nimtz, J. S., Adhikari, U., Mehl, R. A., Scheiner, S., Houtz, R. L., Al-Hashimi, H. M., and Trievel, R. C. (2013) Conservation and functional importance of carbon-oxygen hydrogen bonding in AdoMet-dependent methyltransferases, *J Am Chem Soc 135*, 15536-15548.

[10] Southall, S. M., Cronin, N. B., and Wilson, J. R. (2014) A novel route to product specificity in the Suv4-20 family of histone H4K20 methyltransferases, *Nucleic Acids Res 42*, 661-671.

[11] Qiao, Q., Li, Y., Chen, Z., Wang, M., Reinberg, D., and Xu, R. M. (2011) The structure of NSD1 reveals an autoregulatory mechanism underlying histone H3K36 methylation, *J Biol Chem 286*, 8361-8368.

[12] Wu, H., Min, J., Lunin, V. V., Antoshenko, T., Dombrovski, L., Zeng, H., Allali-Hassani, A., Campagna-Slater, V., Vedadi, M., Arrowsmith, C. H., Plotnikov, A. N., and Schapira, M. (2010) Structural biology of human H3K9 methyltransferases, *Plos One 5*, e8570.

[13] Carney, A. E., and Holden, H. M. (2011) Molecular architecture of TylM1 from Streptomyces fradiae: an N,N-dimethyltransferase involved in the production of dTDP-D-mycaminose, *Biochemistry 50*, 780-787.