Supporting Information for

High-Throughput Computational Investigation of Protein Electrostatics and Cavity for SAM-Dependent Methyltransferases

Christopher Jurich,[1] and Zhongyue J. Yang[1-5,*]

[1]*Department of Chemistry, Vanderbilt University, Nashville, Tennessee 37235, United States* [2]*Center for Structural Biology, Vanderbilt University, Nashville, Tennessee 37235, United States* [3]*Vanderbilt Institute of Chemical Biology, Vanderbilt University, Nashville, Tennessee 37235, United States* [4]*Data Science Institute, Vanderbilt University, Nashville, Tennessee, 37235, United States* [5]*Department of Chemical and Biomolecular Engineering, Vanderbilt University, Nashville, Tennessee 37235, United States*

Contents

Table S1. Activation free energy barriers for various EC 2.1.1.X enzymes. The values were converted from turnover numbers extracted from IntEnzyDB.

EC Number	Activation Free Energy (kcal/mol)
2.1.1.1	14.9
2.1.1.103	11.3
2.1.1.104	13.8
2.1.1.141	15.8
2.1.1.15	13.5
2.1.1.164	12.6
2.1.1.184	10.5
2.1.1.2	12.3
2.1.1.228	12.6
2.1.1.235	13.1
2.1.1.293	8.8
2.1.1.303	15.1
2.1.1.41	14.0
2.1.1.53	12.1
2.1.1.72	10.2

Table S2. List of PDB ids

The below table contains all of the 175 PDB codes gathered by the query. The used column indicates if they were included in the final dataset of 91.

pdb	used	pdb	used	pdb	used	pdb	used
1eiz	TRUE	3elu	TRUE	5e9q	FALSE	5zzd	FALSE
1ej0	FALSE	3elw	FALSE	5ec8	FALSE	6a5k	TRUE
1fpq	TRUE	3fzg	TRUE	5ehi	FALSE	6bm5	FALSE
1g60	TRUE	3g88	TRUE	5eif	FALSE	6h2u	TRUE
1i9g	TRUE	3g89	FALSE	5eiw	FALSE	6i3c	FALSE
1jg4	TRUE	3gcz	FALSE	5ekx	FALSE	6ift	TRUE
1msk	FALSE	3lcv	TRUE	5fhq	FALSE	6inw	FALSE
1n2x	TRUE	3m6v	TRUE	5fhr	TRUE	6ix5	FALSE
1n6a	TRUE	3m6w	TRUE	5h5f	TRUE	6ix8	TRUE
1rjd	TRUE	3mb5	TRUE	5il1	TRUE	6ix9	FALSE
1sqf	TRUE	3mte	TRUE	5jhn	FALSE	6j1o	FALSE
1v2x	TRUE	3o7w	TRUE	5jin	FALSE	6kmr	TRUE
1vid	FALSE	3ooi	TRUE	5jiy	FALSE	6kqq	FALSE
1vpt	TRUE	3p97	FALSE	5jj0	FALSE	6lfe	FALSE
1wg8	TRUE	3pfg	TRUE	5k7u	TRUE	6nw6	TRUE

1zq9	TRUE	3qwp	FALSE	5kjk	FALSE	6p3n	TRUE
2avd	TRUE	3r24	FALSE	5kqr	FALSE	6p6k	FALSE
2b9e	TRUE	3rq4	TRUE	5kqs	FALSE	6p7z	TRUE
2cl5	FALSE	3s68	FALSE	5l6e	TRUE	6paf	FALSE
2dpm	FALSE	3s8p	TRUE	5lsa	FALSE	6w4h	FALSE
2egv	TRUE	3tg4	TRUE	5nfj	TRUE	6w61	FALSE
2fk8	TRUE	3tm4	TRUE	5njv	FALSE	6w75	FALSE
2g72	TRUE	4blv	TRUE	5t0k	TRUE	6wvn	FALSE
2nxe	TRUE	4ctj	FALSE	5t0m	FALSE	6wzw	TRUE
2nyu	TRUE	4df3	TRUE	5teg	TRUE	6x0p	FALSE
2plw	TRUE	4e47	FALSE	5ttf	TRUE	6yuh	FALSE
2qmm	TRUE	4fak	TRUE	5ttg	TRUE	6zrb	FALSE
2r3a	TRUE	4fzv	TRUE	5tuz	FALSE	7bj1	FALSE
2uyq	TRUE	4htf	TRUE	5ubb	TRUE	7btv	FALSE
2v3k	TRUE	4ij8	TRUE	5upd	TRUE	7dcf	FALSE
2yqz	FALSE	4iv8	FALSE	5v37	FALSE	7edc	TRUE
3ai9	TRUE	4j83	FALSE	5v9i	FALSE	7lzd	TRUE
3aia	TRUE	4jlg	FALSE	5v9j	TRUE	7m6b	TRUE
3bwm	FALSE	4lg1	TRUE	5vsc	TRUE	7o2a	FALSE
3bwy	TRUE	4n49	TRUE	5vsd	FALSE	7o2c	FALSE
3bxo	TRUE	4qdj	TRUE	5vse	TRUE	7o4n	TRUE
3ckk	TRUE	4xuc	FALSE	5vsf	TRUE	7phd	TRUE
3dmf	TRUE	4yvg	TRUE	5wy0	TRUE	7r7o	TRUE
3dmh	FALSE	5arf	FALSE	5x7f	FALSE	7se7	TRUE
3dou	TRUE	5arg	FALSE	5xxd	FALSE	7t7l	FALSE
3dxy	TRUE	5ccl	FALSE	5xxj	FALSE	7tw7	FALSE

Query: (Data Collection Resolution **<=** 2 **AND**
Enzyme Classification Number = 2.1.1 **AND**
Polymer Entity Type **NOT** = "DNA" **AND**
Polymer Entity Type **NOT** = "RNA" **AND**
Polymer Entity Type **NOT** = "NA-hybrid")
AND Chemical ID(s) = SAM

Figure S1: PDB query used

The above query was used on 2022-10-03 to collect the initial list of 175 PDB ids.

Table S3. Enzyme quaternary state of the dataset. Biological assemblies from the PDB were used.

Number of Units	Percentage	PDB IDs
1	80.22%	1eiz,1fpq,1i9g,1jg4,1n2x,1n6a,1rjd,1sqf,1v2x,1vpt,1wg8,2b9e,2fk8,2g72, 2nxe,2nyu,2plw,2r3a,2uyq,2v3k,3ai9,3bwy,3ckk,3dmf,3dou,3dxy,3elu,3fzg, 3g88,3lcv,3m6v,3m6w,3mb5,3mte,3o7w,3ooi,3pfg,3rq4,3s8p,3tg4,3tm4, 4blv,4df3,4fak,4fzv,4n49,4qdj,4yvg,5h5f,5il1,5k7u,5l6e,5nfj,5t0k,5teg,5ttg, 5ubb,5upd,5vse,5wy0,6a5k,6h2u,6ift,6kmr,6p3n,6p7z,6wzw,7edc,7lzd, 7m6b,7o4n,7r7o,7se7
2	18.68%	1g60,1zq9,2avd,2egv,2qmm,3aia,3bxo,4htf,4ij8,5fhr,5ttf,5v9j,5vsc,5vsf, 6ix8,6nw6,7phd
3	1.10%	4lg1

Text S1: Sequence clustering

Sequences were taken from the FASTA files associated with each structure and clustered through using a 95% similarity cutoff and 5% distance cutoff. For each sequence, it would be compared to all the sequences in existing clusters and would be included if it was the edit distance between the sequences and the length difference between the two is less than 5% the average length of the sequence being compared to and the candidate sequence. In this version of edit distance, all edits are given a weighting of 1.

Table S4: Enzymatic function data annotations

Below are the annotations and reaction classifications for all entries included in the final analysis. Reaction categories were derived from EC numbers when available or publication data when not.

pdb	EC	protein reaction	RNA/DNA reaction	histone reaction	small molecule reaction	SAM reaction
1eiz	2.1.1.166	0	1	0	0	1
1fpq	2.1.1.154	0	0	0	1	1
1g60	2.1.1.72	0	1	0	0	1
1i9g	2.1.1.220	0	1	0	0	1
1jg4	2.1.1.77	1	0	0	0	1
1n2x	2.1.1.199	0	1	0	0	1
1n6a	2.1.1.364	0	0	1	0	1
1rjd	2.1.1.233	1	0	0	0	1
1sqf	2.1.1.176	0	1	0	0	1
1v2x	2.1.1.34	0	1	0	0	1

1vpt	2.1.1.57	0	1	0	0	1
1wg8	2.1.1.199	0	1	0	0	1
1zq9	2.1.1.183	0	1	0	0	1
2avd	2.1.1.6	0	0	0	1	1
2b9e	2.1.1.-	0	1	0	0	1
2egv	2.1.1.193	0	1	0	0	1
2fk8	2.1.1.-	0	0	0	1	1
2g72	2.1.1.28	0	0	0	1	1
2nxe	2.1.1.244	1	0	0	0	1
2nyu	2.1.1.-	0	1	0	0	1
2plw	2.1.1.56	0	1	0	0	1
2qmm	2.1.1.257	0	1	0	0	1
2r3a	2.1.1.355	0	0	1	0	1
2uyq	2.1.1.-	0	0	0	0	1
2v3k	2.1.1.260	0	1	0	0	1
3ai9	2.1.1.257	0	1	0	0	1
3aia	2.1.1.257	0	1	0	0	1
3bwy	2.1.1.6	0	0	0	1	1
3bxo	2.1.1.234	0	0	0	1	1
3ckk	2.1.1.33	0	1	0	0	1
3dmf		0	1	0	0	1
3dou	2.1.1.166	0	1	0	0	1
3dxy	2.1.1.33	0	1	0	0	1
3elu	2.1.1.57	0	0	0	0	1
3fzg	2.1.1.179	0	1	0	0	1
3g88	2.1.1.170	0	1	0	0	1
3lcv	2.1.1.179	0	1	0	0	1
3m6v	2.1.1.-	0	1	0	0	1
3m6w	2.1.1.-	0	1	0	0	1
3mb5	2.1.1.219	0	1	0	0	1
3mte	2.1.1.180	0	1	0	0	1
3o7w	2.1.1.-	0	1	0	0	1
3ooi	2.1.1.357	0	0	1	0	1
3pfg	2.1.1.235	0	0	0	1	1
3rq4	2.1.1.362	0	0	1	0	1
3s8p	2.1.1.362	0	0	1	0	1
3tg4	2.1.1.354	0	0	1	0	1
3tm4	2.1.1.256	0	1	0	0	1
4blv	2.1.1.266	0	1	0	0	1
4df3	2.1.1.-	0	1	0	0	1
4fak	2.1.1.177	0	1	0	0	1
4fzv	2.1.1.-	0	1	0	0	1
4htf	2.1.1.-	0	1	0	0	1

4ij8	2.1.1.-	0	0	1	0	1
4lg1	2.1.1.-	1	0	0	0	1
4n49	2.1.1.57	0	1	0	0	1
4qdj	2.1.1.11	0	0	0	1	1
4yvg	2.1.1.228	0	1	0	0	1
5fhr	2.1.1.6	0	0	0	1	1
5h5f	2.1.1.-	0	0	0	0	1
5il1	2.1.1.348	0	1	0	0	1
5k7u		0	1	0	0	1
5l6e	2.1.1.348	0	1	0	0	1
5nfj	2.1.1.218	0	1	0	0	1
5t0k	2.1.1.367	0	0	1	0	1
5teg	2.1.1.367	0	0	1	0	1
5ttf	2.1.1.367	0	0	1	0	1
5ttg	2.1.1.367	0	0	1	0	1
5ubb	2.1.1.299	1	0	0	0	1
5upd	2.1.1.370	0	0	1	0	1
5v9j	2.1.1.367	0	0	1	0	1
5vsc	2.1.1.367	0	0	1	0	1
5vse	2.1.1.367	0	0	1	0	1
5vsf	2.1.1.367	0	0	1	0	1
5wy0	2.1.1.n8	0	1	0	0	1
6a5k	2.1.1.43	0	0	1	0	1
6h2u	2.1.1.-	0	1	0	0	1
6ift	2.1.1.182	0	1	0	0	1
6ix8	2.1.1.-	0	0	0	0	1
6kmr	2.1.1.72	0	0	1	0	1
6nw6	2.1.1.228	0	1	0	0	1
6p3n	2.1.1.122	0	0	0	0	1
6p7z	2.1.1.354	0	0	1	0	1
6wzw	2.1.1.359	0	0	1	0	1
7edc	2.1.1.34	0	1	0	0	1
7lzd	2.1.1.359	0	0	1	0	1
7m6b	2.1.1.72	0	1	0	0	1
7o4n	2.1.1.217	0	1	0	0	1
7phd						
7r7o						
7se7		0	0	1	0	1

The below content was used in the min.in file supplied to Amber19's sander application:

```
Minimize
 &cntrl
  imin=1,
  ntx=1,
  irest=0,
  maxcyc=20000,
  ncyc=10000,
  ntpr=4000,
  ntwx=0,
  cut=8.0,
 /
```

Figure S2: AMBER MD minimization settings

Table S5. Dipole Moments and ΔG_{elec} Values. In the case that multiple active sites exist, they are denoted as <code>.<n>.

Active Site	Dipole Moment Magnitude (Debye)	ΔG_{elec} (kcal/mol)
1eiz	0.764	19.619
1g60.1	0.802	-0.656
1g60.2	0.840	4.422
1i9g	0.784	6.079
1jg4	0.772	-0.180
1n2x	0.748	12.611
1n6a	0.831	4.889
1sqf	0.719	2.756
2avd.1	0.833	8.320
2avd.2	0.820	1.658
2r3a	0.854	5.160
3ooi	0.744	-8.742

Table S6. B-Factors By Enzyme Location. Raw B values were taken from each respective PDB entry. B_{norm} was calculated using Eq 1. Active site is defined as all atoms within 3 Å of the SAM ligand.

	SAM Ligand		Active Site		Rest of Enzyme	
Active Site Location	$\overline{B}_{norm}$	$\sigma_{B_{norm}}$	$\overline{B}_{norm}$	$\sigma_{B_{norm}}$	$\overline{B}_{norm}$	$\sigma_{B_{norm}}$
1eiz.A.301.SAM	-0.408	0.277	-0.400	0.665	0.088	1.043
1fpq.A.1699.SAM	0.800	0.819	-0.601	0.377	0.038	1.017
1g60.A.500.SAM	0.425	0.355	-0.063	0.626	-0.025	1.028
1g60.B.501.SAM	0.695	0.163	0.265	0.929	-0.025	1.028
1i9g.A.301.SAM	-0.441	0.109	-0.243	0.466	0.039	1.053
1jg4.A.500.SAM	-0.100	0.380	-0.451	0.644	0.080	1.037
1n2x.A.401.SAM	-0.405	0.228	-0.164	0.923	0.029	1.014
1n6a.A.402.SAM	-0.167	0.276	-0.159	0.792	0.024	1.030
1rjd.A.801.SAM	-0.760	0.160	-0.427	0.792	0.059	1.012
1sqf.A.430.SAM	1.746	0.545	-0.431	0.826	0.011	0.994
1v2x.A.400.SAM	0.366	0.901	-0.370	0.518	0.049	1.044
1vpt.A.400.SAM	-0.722	0.317	-0.571	0.826	0.068	1.000
1wg8.A.3142.SAM	-0.564	0.266	-0.371	0.707	0.059	1.027
1zq9.A.4000.SAM	3.124	1.778	0.152	1.041	-0.058	0.929
1zq9.B.4001.SAM	1.722	1.964	0.294	1.015	-0.058	0.929
2avd.A.501.SAM	-0.467	0.197	-0.488	0.568	0.047	1.050
2avd.B.601.SAM	-0.050	0.166	-0.012	0.570	0.047	1.050
2b9e.A.1201.SAM	-0.117	0.155	-0.128	0.671	0.016	1.036
2egv.A.1300.SAM	-0.133	0.865	-0.480	0.547	0.059	1.027
2egv.B.1400.SAM	-0.231	0.873	-0.380	0.695	0.059	1.027
2fk8.A.302.SAM	-0.130	0.377	-0.117	0.752	0.015	1.029
2g72.A.2001.SAM	0.241	0.600	-0.365	0.844	0.049	1.014
2nxe.A.302.SAM	-0.513	0.556	-0.742	0.478	0.112	1.012
2nyu.A.201.SAM	-0.634	0.203	-0.439	0.658	0.095	1.035
2plw.A.203.SAM	-0.066	0.471	-0.314	0.750	0.058	1.037
2qmm.A.301.SAM	1.448	0.179	0.275	1.026	-0.032	1.012
2qmm.B.301.SAM	0.691	0.147	-0.036	0.658	-0.032	1.012
2r3a.A.304.SAM	-0.499	0.351	-0.437	0.704	0.057	1.020
2uyq.A.1311.SAM	1.150	0.276	-0.179	0.458	0.004	1.029
2v3k.A.1254.SAM	-0.467	0.668	-0.532	0.756	0.086	1.010
3ai9.X.501.SAM	2.323	0.401	0.329	0.969	-0.078	0.962
3aia.A.206.SAM	0.304	0.105	-0.078	0.708	-0.085	0.967
3aia.B.206.SAM	2.032	0.189	1.069	1.008	-0.085	0.967
3bwy.A.301.SAM	-0.491	0.253	-0.527	0.374	0.106	1.053
3bxo.A.238.SAM	0.201	0.951	-0.029	0.794	0.021	1.036
3bxo.B.238.SAM	-0.153	0.335	-0.272	0.594	0.021	1.036
3ckk.A.301.SAM	-0.322	0.181	-0.124	1.016	0.026	1.005

3dmf.A.388.SAM	-1.364	0.549	-0.388	0.566	0.052	1.018
3dou.A.1.SAM	-0.331	0.378	-0.364	0.475	0.067	1.057
3dxy.A.1.SAM	-0.076	0.179	-0.553	0.918	0.074	0.996
3elu.A.4633.SAM	-0.337	0.479	-0.217	0.709	0.029	1.029
3fzg.A.300.SAM	-0.671	0.233	-0.656	0.434	0.118	1.028
3g88.A.303.SAM	-0.516	0.284	-0.486	0.639	0.057	1.021
3lcv.B.301.SAM	-0.631	0.350	-0.628	0.530	0.115	1.024
3m6v.A.465.SAM	0.069	0.287	-0.405	0.354	0.025	1.025
3m6w.A.465.SAM	0.170	0.351	-0.350	0.422	0.021	1.025
3mb5.A.301.SAM	-0.582	0.184	-0.263	0.670	0.041	1.033
3mte.A.220.SAM	-0.441	0.337	-0.745	0.365	0.121	1.022
3o7w.A.801.SAM	-0.344	0.207	-0.512	0.384	0.057	1.033
3ooi.A.237.SAM	-0.691	0.260	-0.565	0.582	0.095	1.024
3pfg.A.264.SAM	-0.752	0.127	-0.490	0.450	0.109	1.051
3rq4.A.500.SAM	-0.826	0.271	-0.669	0.514	0.109	1.017
3s8p.A.500.SAM	-0.866	0.192	-0.675	0.483	0.128	1.020
3tg4.A.434.SAM	-0.506	0.114	-0.179	0.505	0.018	1.030
3tm4.A.401.SAM	-0.685	0.197	-0.550	0.603	0.054	1.015
4blv.A.1281.SAM	-0.164	0.204	-0.433	0.351	0.068	1.055
4df3.A.301.SAM	-0.175	0.179	-0.306	0.757	0.044	1.029
4fak.A.201.SAM	-0.159	0.627	-0.373	0.989	0.064	0.995
4fzv.A.401.SAM	-0.652	0.216	-0.672	0.399	0.072	1.019
4htf.A.301.SAM	-0.229	0.169	-0.247	0.560	0.006	1.040
4htf.B.301.SAM	-0.028	0.146	0.191	0.737	0.006	1.040
4ij8.A.501.SAM	-0.953	0.305	-0.830	0.289	0.219	1.002
4ij8.B.501.SAM	-0.932	0.330	-0.857	0.308	0.219	1.002
4lg1.A.301.SAM	-0.177	0.153	-0.392	0.697	0.058	1.026
4lg1.B.301.SAM	-0.658	0.172	-0.755	0.440	0.058	1.026
4lg1.C.301.SAM	0.230	0.109	0.245	0.935	0.058	1.026
4n49.A.601.SAM	-0.422	0.247	-0.455	0.417	0.035	1.024
4qdj.A.301.SAM	-0.134	0.603	-0.335	0.627	0.067	1.050
4yvg.A.301.SAM	-0.140	0.964	-0.618	0.492	0.076	1.020
5fhr.A.303.SAM	-0.571	0.245	-0.513	0.492	0.118	1.045
5fhr.B.303.SAM	-0.619	0.301	-0.542	0.520	0.118	1.045
5h5f.A.301.SAM	-0.878	0.378	-0.368	0.899	0.101	1.002
5il1.A.601.SAM	-0.485	0.277	-0.129	0.866	0.031	1.025
5k7u.A.601.SAM	-0.335	0.212	-0.246	0.674	0.047	1.046
5l6e.A.601.SAM	-0.117	0.574	-0.296	0.629	0.058	1.052
5nfj.A.501.SAM	-0.463	0.348	-0.398	0.485	0.091	1.063
5t0k.A.1304.SAM	-0.503	0.263	-0.331	0.655	0.053	1.035
5teg.A.401.SAM	-0.446	0.232	-0.621	0.257	0.122	1.050

5ttf.A.1505.SAM	-0.334	0.249	-0.295	0.662	0.046	1.033
5ttf.B.1505.SAM	-0.304	0.220	-0.361	0.678	0.046	1.033
5ttg.A.1301.SAM	0.162	0.244	0.023	0.818	-0.005	1.026
5ubb.A.301.SAM	-0.331	0.352	-0.522	0.477	0.097	1.045
5upd.A.1301.SAM	-0.709	0.222	-0.534	0.678	0.102	1.020
5v9j.A.1301.SAM	0.127	0.203	0.128	0.832	0.000	1.024
5v9j.B.1301.SAM	-0.246	0.136	-0.110	0.833	0.000	1.024
5vsc.A.1505.SAM	-0.543	0.150	-0.208	0.493	0.052	1.039
5vsc.B.1505.SAM	-0.714	0.144	-0.524	0.438	0.052	1.039
5vse.A.1505.SAM	-0.685	0.149	-0.408	0.452	0.058	1.039
5vsf.A.3001.SAM	-0.261	0.254	-0.118	0.750	0.017	1.032
5vsf.B.3001.SAM	-0.094	0.260	-0.113	0.746	0.017	1.032
5wy0.A.800.SAM	0.108	0.467	0.044	0.671	-0.009	1.049
6a5k.A.805.SAM	-0.732	0.512	-0.542	0.619	0.047	1.012
6h2u.A.301.SAM	-0.552	0.279	-0.411	0.529	0.083	1.050
6ift.A.301.SAM	-0.745	0.196	-0.520	0.492	0.083	1.033
6ix8.A.401.SAM	0.058	0.471	-0.352	0.610	0.030	1.021
6ix8.B.401.SAM	-0.418	0.447	-0.397	0.608	0.030	1.021
6nw6.A.301.SAM	0.290	1.652	-0.400	0.736	0.066	1.006
6nw6.B.301.SAM	0.354	1.878	-0.444	0.640	0.066	1.006
6p3n.A.401.SAM	-0.342	0.313	-0.462	0.355	0.041	1.030
6p7z.A.506.SAM	-0.923	0.253	-0.397	1.003	0.042	0.992
6wzw.A.2302.SAM	-0.623	0.280	-0.267	0.803	0.056	1.027
7edc.A.300.SAM	2.595	0.906	-0.014	1.237	-0.042	0.902
7lzd.A.1804.SAM	-0.846	0.107	-0.661	0.352	0.111	1.028
7m6b.A.301.SAM	-0.642	0.299	-0.554	0.756	0.093	1.009
7o4n.A.301.SAM	0.626	0.764	-0.218	0.750	0.014	1.022
7phd.A.405.SAM	-0.771	0.663	-0.340	1.078	0.041	0.960
7phd.B.404.SAM	-0.705	0.654	-0.147	1.374	0.041	0.960
7r7o.A.301.SAM	-0.620	0.254	-0.649	0.372	0.117	1.033
7se7.A.401.SAM	-0.386	0.320	-0.407	0.417	0.073	1.056

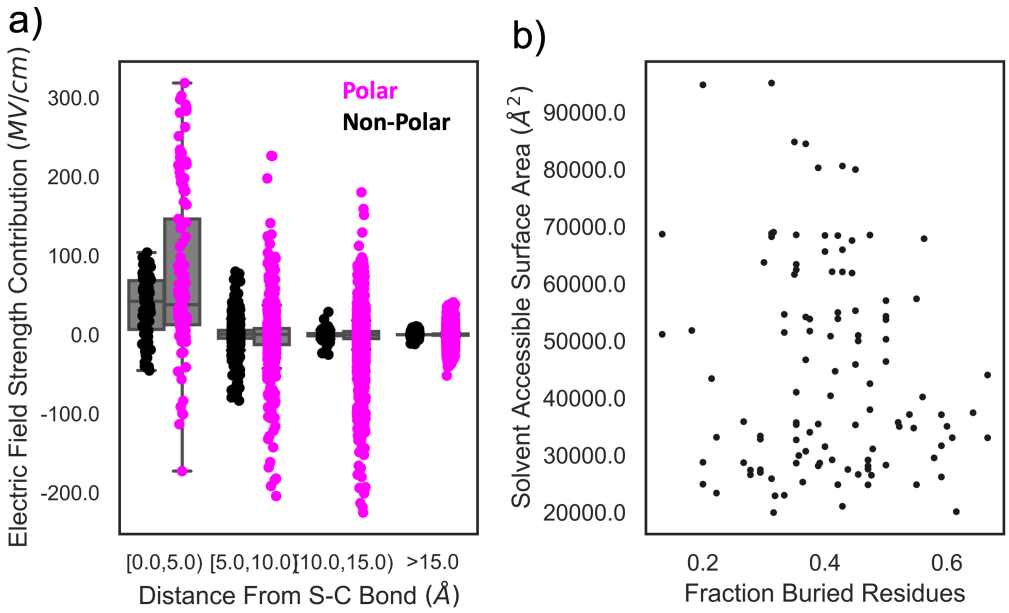

Figure S3 EF strength contribution stratified by residue characteristics. The left (a) shows distributions of the EF strength contribution by residue in each of the 91enzymes as a function of distance from the S—C bond in their respective enzyme. Boxplots are in grey for both polar and non-polar residues whereas stripplots are color coded by polarity with magenta and black representing polar and non-polar residues, respectively. The right shows the surface accessible surface area (SASA) of the active site and percent of buried residues in the active site, respectively. A residue was considered to be in the active site if it had an atom with 3 Å of SAM. A buried residue is defined as one which has less than 10% of its surface area exposed to solvent. All values were calculated using pymol.

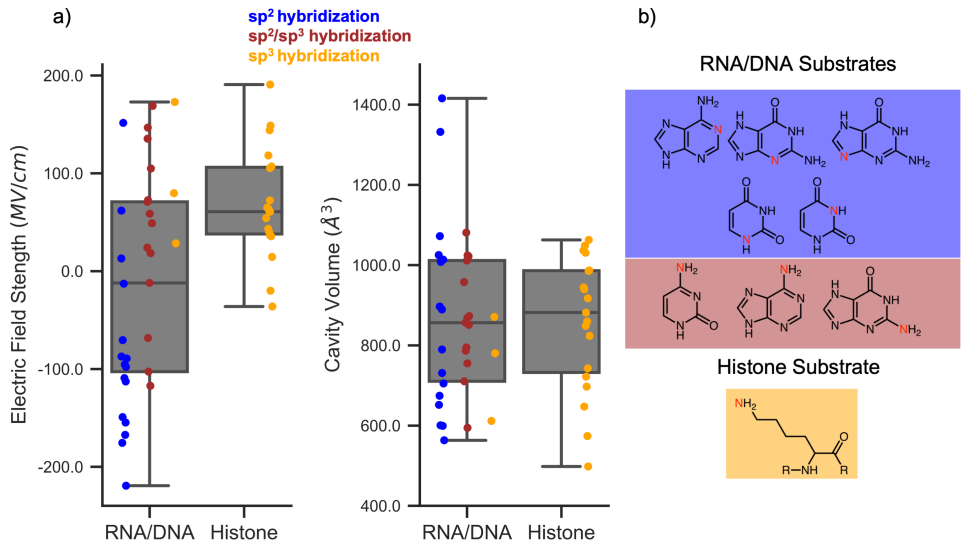

Figure S4 *N*-targeting MTases stratified by substrate and hybridization. Below is a variation of Figure 4 from the main text where exocylic -NH_2 groups are placed into their own sp^2/sp^3 group, colored in brown.