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

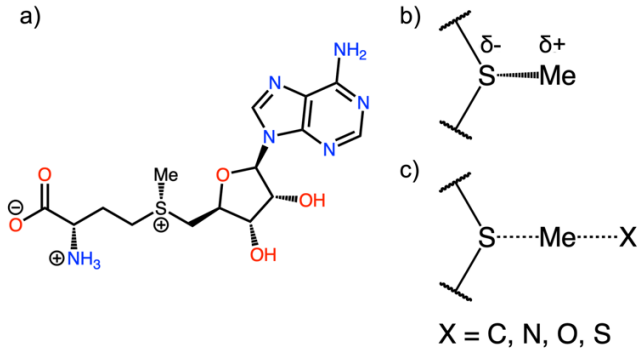
Christopher Jurich,[1] and Zhongyue J. Yang[1-4,*]

[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*

ABSTRACT: *S*-adenosyl methionine (SAM) -dependent methyl transferases (MTases) are a ubiquitous class of enzymes catalyzing dozens of essential life processes. Despite targeting a large space of substrates with diverse intrinsic reactivity, SAM MTases have similar catalytic efficiency. While understanding of MTase mechanism has grown tremendously through integration of structural characterization, kinetic assays, and multiscale simulations, it remains elusive regarding how these enzymes have evolved to fit the diverse chemical needs of their respective substrates. In this work, we performed a high-throughput computational analysis of 91 SAM MTases to better understand how the properties (i.e., electric field strength and active site volumes) of SAM MTases have been evolved to achieve similar catalytic efficiency towards substrates of different reactivity. We found that electric field strengths have largely evolved to make the target atom a better methyl acceptor. For MTases that target RNA/DNA- and histone protein, our results suggest that electric field strength accommodates hybridization state and variation in cavity volume trends with diversity of substrate classes. Metal ions in SAM MTases contribute negatively to electric field strength for methyl donation and enzyme scaffolds have evolved to offset these contributions.

1. Introduction

S-adenosyl methionine methyltransferases (SAM MTases) are a ubiquitous class of enzymes. SAM MTases are observed in a wide range of organisms including bacteria, fungus, plants, and humans.[1-4] SAM MTases catalyze S_N2 methylation, using SAM as a methyl donor. Their functions are involved in many essential life processes, including gene expression,[5-7] protein modification,[8-10] neurotransmitter degradation,[11-13] and natural product synthesis.[14-16] SAM MTases target various types of atoms, including carbon, nitrogen, oxygen and sulfur (Scheme 1). Enabled by advances in structural determination, kinetic studies, and multiscale molecular modeling,[17-20] the mechanistic detail of SAM MTases has been unveiled.[21-23] Combination of binding isotope effect experiment and large-scale quantum mechanical calculations shows the dependence of ground state donor-acceptor distance on the catalytic efficiency of catechol-*O*-methyltransferase.[18, 24, 25] Investigations into C–H hydrogen bonding have elucidated the role of these non-bonding interactions in ensuring catalytic efficiency.[26-28] Analysis of charge transfer and electrostatics on four Class I SAM MTases has illustrated the ability of these enzymes to customize their electrostatic potentials to the intrinsic reactivity of their target substrates.[29]



Scheme 1. Chemical structure of *S*-adenosyl methionine. The left side (a) shows the full molecule in the protonation state seen at a pH of 7.0. Top right (b) shows the relevant S–C bond broken during methyl donation by SAM. The negative partial charge of sulfur and positive partial charge

of the methyl group are shown. Bottom right (c) shows a hypothetical S_N2 transition state where the methyl group is being transferred to either a carbon, nitrogen, oxygen, or sulfur.

Although SAM MTases involve diverse substrate scope with a wide range of intrinsic methyl-accepting capability, the kinetic properties of MTase-catalyzed reactions are largely consistent. A survey of 15 unique MTases from IntEnzyDB shows an average activation barrier of 12.7 kcal/mol with a standard deviation of 1.9 kcal/mol (Supporting Information, Table S1).[30, 31] Similarity of turnover number across diverse substrates is not unique to SAM MTases and is observed in many classes of enzymes.[32] The combination of diverse substrates and functional roles with consistent kinetic output suggests that SAM MTases have evolved to address the specific characteristics of each catalyzed reaction. However, the molecular origins behind the substrate kinetic homogeneity remain largely unknown.

A high throughput analysis of SAM MTase electrostatic and topological properties is needed to rationalize how this diverse class of enzymes has evolved to achieve a narrow distribution of kinetics albeit substrate diversity. Existing studies have used different types of MTases but remained relatively low throughput, usually using one or two structures.[17, 19, 23, 33] A notable exception is the structural survey conducted by the Trievel lab, which involves 46 different SAM MTases.[26, 29] Combined growth in computational resources and protein databank (PDB) entries makes a strong case for higher throughput studies that have greater potential to uncover more universal trends.[34-36]

Here, we present a computational analysis of 91 high-quality SAM MTase structures that examines both enzyme interior electric field (EF) strength and cavity volume using EnzyHTP, a high-throughput enzyme modeling software developed by our lab.[34] Existing computational and experimental studies of SAM MTases[29] and other types of enzymes have elucidated interior

enzyme electrostatics to be among the determining factors in mediating catalytic efficiency.[37-40] Volume of active site cavity represents a topological factor that informs the capability of SAM MTases in substrate binding. Measuring both an electrostatic and topological value provides a holistic view of how each MTase has evolved to the specific characteristics of their respective substrates to achieve strong catalytic efficiency. Our analyses of 91 unique SAM MTases have enabled us to evaluate how SAM MTases have evolved to specific characteristics of their substrates' target atoms, the diversity of their class of target substrates, and the presence of metal ions in their structures.

2. Computational Methods

Data Curation. Structures were curated from the Protein Databank (PDB) on October 3rd, 2022.[35] Filtration criteria are: (1) resolution under 2.0 Å, (2) enzyme commission (EC) number of 2.1.1.X (i.e., SAM-dependent methyltransferase), (3) inclusion of *S*-adenosyl methionine (SAM), and (4) no RNA or DNA fragments. These criteria yielded 175 PDB entries. The PDB codes and corresponding search query are listed in the Supporting Information (Table S2 and Figure S1).

The biological assembly and FASTA sequence for each entry was downloaded from the PDB website. When alternative locations or ANISOU records were available, the first coordinates were used. Co-crystallizing ligands and ions were removed manually except Zn and Mg ions. Mg ions were only kept for catechol *O*-methyltransferase (COMT) structures. Missing residues were added to each structure using the Modeller python package[41] by treating the incomplete sequence as a template and aligning it to the full sequence provided by the FASTA. All structures were protonated at a pH of 7.0 using the EnzyHTP[34] package and ligands were protonated using the Molecular Operating Environment (MOE) software package.[42] Ligands missing heavy atoms were replaced with idealized models found in the PDB's chemical library.

Each entry was clustered into one of 104 groups using edit distance with a cutoff of 95% sequence similarity and a maximum length difference of 5% (Supporting Information, Text S1). Clustering was performed by iterating through the list of sequences and iteratively checking if the sequence has satisfied the criteria listed above. If it satisfied these criteria, it would be added. If not, a new cluster containing just that sequence would be created. A total of 21 clusters were identified to contain more than one sequence. The sequence from each cluster with the best resolution in angstroms (i.e. the smallest value) was selected by default. If a cluster contains a native version of the enzyme and versions with inhibitors, the native version was selected as the representative member from the cluster. PDB entries 3m6v and 3m6w were included due to being in different space groups as specified by their authors.

EC numbers were sourced from PDB entries when available, and from the UniProtKB database when not listed in the PDB.[43] Information on substrate type and target heteroatom were derived from the BRENDA database entry corresponding to EC number.[44] Substrate and heteroatom information was also derived from PDB entries when no EC number could be derived. We found 11 entries that have no corresponding EC number and removed them from the dataset for analyses that focus on substrate and target atom specificity. We also removed 21 entries that have EC numbers other than 2.1.1.X. Complete enzymatic function data annotations are listed in the Supporting Information (Table S3).

Molecular Mechanics Minimization. The SAM-methyltransferase complexes were parameterized using the antechamber, parmchk2 and tleap utilities from AMBER19.[45] Non-SAM ligands were removed from all structures. Each structure was minimized using Amber19's sander application. Minimization was run for a maximum of 20,000 steps and the remaining settings are listed in the Supporting Information (Figure S2).

Electric Field Calculation. For each SAM-methyltransferase complex, the enzyme interior electric field (EF) strength was calculated along the S–C bond of SAM using the RESP charges for each atom as well as the coordinates of the minimized structure. The below equation was used to calculate the electric field strength at the mid-point of the S–C bond by summing over all atoms:

$$Electric\ Field\ Strength = \sum \frac{kq}{|\vec{r}|^2}(\hat{r} \cdot \hat{d})$$

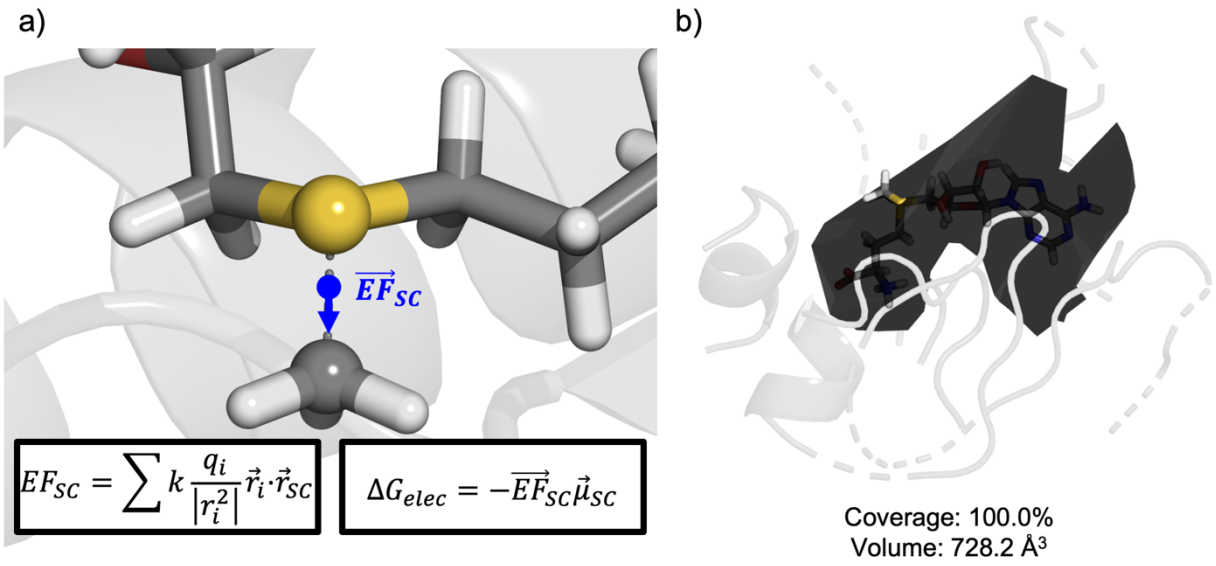
Here $\vec{r}$ is the vector from each atom to the S-C bond mid-point, $\hat{d}$ is a unit vector pointing along the direction of the S-C bond from the center of the bond, q is the charge of the atom in internal Amber units and k is a conversion constant with a value of 332.4 kcal Å/ mol e^2. The resulting units of EF strength are MV/cm.

Cavity Volume Calculation. We used the Mole2.0 software package to calculate the cavity volume of SAM-binding pocket.[46] A truncated version of each cavity was created by selecting residues containing at least one heavy atom that is within 8 Å of a heavy atom in SAM. Volume was calculated a total of 378 times for each cavity with the outer probe radius ranging from 3.0 Å to 6.50 Å by 0.25 Å, inclusive, and with inner probe values of 1.0 Å to 2.0 Å by 0.05 Å, inclusive. Coverage of SAM by each parameter set was calculated by exporting the cavity to a mesh and then counting the number of heavy atoms whose centers are inside the mesh using the pyvista python module. Each cavity was assigned the lowest volume with coverage greater than 0.95 or with the max coverage if none were greater than 0.95.

3. Results and Discussion

We curated 91 MTases for analyses of enzyme interior electric field and cavity size of active site. As shown in Scheme 2, electric field strength (EF) for each enzyme was measured

along the sulfur-carbon bond (S–C) at the middle point between geometric centers of the two atoms with the charges derived from the AMBER force field. Positive EF strength values indicate a stronger propensity of SAM to donate the methyl group. Cavity volumes were calculated using mesh objects generated from the Mole2.0 software package. The protocols for structural curation and property analysis are detailed in the Computational Methods section.



Scheme 2: Measuring electric field (EF) strength and cavity volume. Visualized by the blue arrow in (a), electric field is measured along the S-C bond axis at the center point of the bond. Charges and atom locations for the EF calculation are derived from the minimized AMBER structure. Inlay (b) shows a sample mesh from Mole2.0 used to calculate the cavity volume. Coverage denotes the percentage of SAM's heavy atoms whose centers are inside the mesh.

The curated MTases involves a diverse range of substrate specificity (Figure 1a) with nucleic acids and histones being the most abundant, representing 55.0% and 23.1% of the dataset,

respectively. MTases catalyzing small molecules (i.e., molecular weight ⩽582.9 g/mol), proteins and catechol each contribute 11.0%, 7.7% and 3.3% of the structures, respectively. We isolated catechol from the category of small molecules because of its high abundance and well-established existing studies elucidating its mechanism.[11, 17-19, 22, 29, 47, 48] The distributions of cavity volumes and electric field values are displayed in Figure 1b. Cavity volumes range from 498.0 $Å^3$ to 1494.7 $Å^3$, with a median value of 859.6 $Å^3$. Kernel density estimation (KDE) of cavity volume generates a distribution with a single peak around 806.3 $Å^3$ and a small shoulder around 1400.0 $Å^3$. SAM has a volume of 295.2 $Å^3$, meaning that the median cavity volume has space of approximately twice the size of SAM to accommodate methyl-acceptor substrate.

Electric field strengths range from -306.2 MV/cm to 190.7 MV/cm, with a median of 13.8 MV/cm. This result indicates that a large number of SAM MTases have positive electric field strength although a substantial number have negative strength values. KDE estimation of electric field generated a distribution with two peaks, one around -89.5 MV/cm and another around 93.0 MV/cm. While the median electric field is on average slightly positive, KDE analysis indicates that the wide distribution of EF strength in specific MTases. Cavity volume and EF strength have low correlation with an R^2 of only 0.01. The lack of relationship implies the two properties are independent and EF strength does directly depend on binding cavity size.

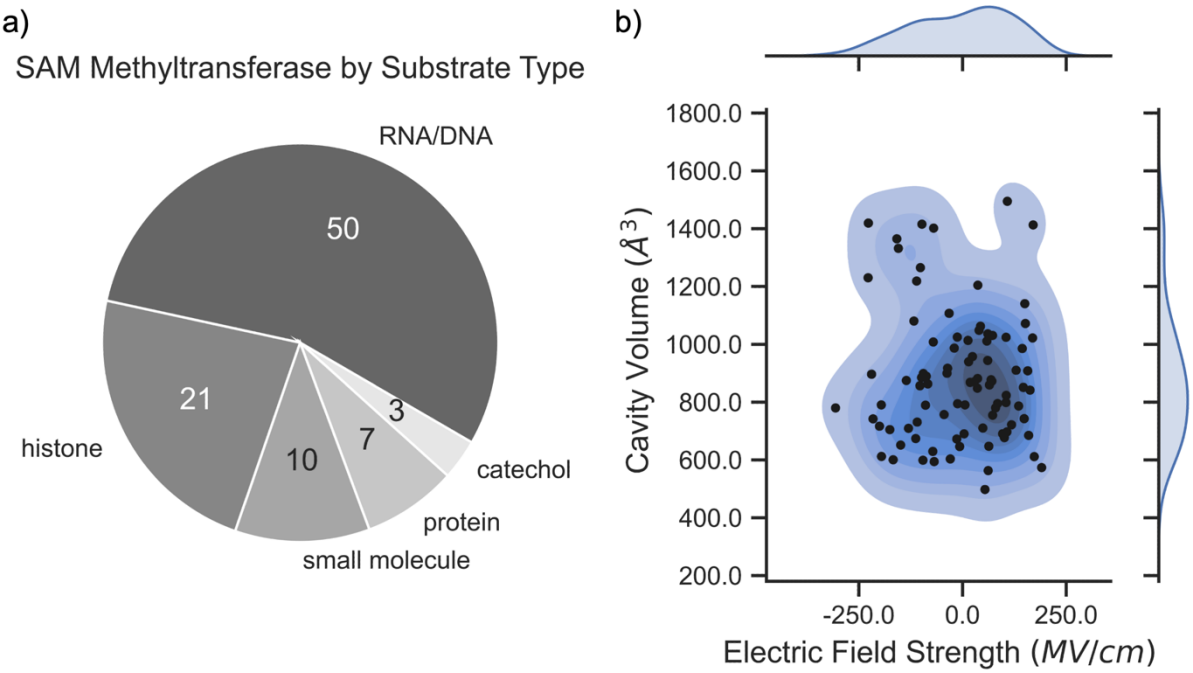


Figure 1. A survey of the compiled SAM MTase dataset. The pie chart (a) on the left shows the functional breakdown of the MTases. Classifications are based on the EC number when available and are derived from the original publication if no EC number was supplied. The heatmap (b) on the right shows the distributions of cavity volume and electric field for all 91 MTases. The center of (b) shows a topographical view of the values in two dimensions and the margins show one dimensional KDE plots of each respective dimension.

Next, we investigated the electrostatic mediation of MTases on SAM co-factor with substrates containing different methyl-accepting polar atoms. We observe statistically significant differences in the distribution of electric field strengths for *O*- and *N*-targeting MTases (Figure 2, *left*). The EF strengths of *O*-targeting MTases are on average 136.0 MV/cm more negative than those targeting nitrogen. Using student's t-test, the electric field distributions are statistically significant with a p-value of 1.95×10^{-6}. The average electric field of 20.0 MV/cm for *N*-targeting MTases is 62.1% more positive than the 13.8 MV/cm average field for the broader dataset. The

results show that the protein scaffolds of *N*-targeting MTases have evolved more positive electric fields to overcome the lower electronegativity of nitrogen, in turn making it a better acceptor of SAM's methyl group. Oxygen's stronger electronegativity makes it a naturally stronger methyl group acceptor, and therefore, MTases take less efforts to evolve their interior electric field. Notably, these analyses only reflect a general trend because the specific functional groups under each heteroatom-targeting category involve large disparities of methyl-accepting ability (e.g., hydroxyl versus carboxyl).

In contrast, differences in cavity volumes for *O*- and *N*-targeting MTases are not statistically significant. Cavity volumes only differ by 70.8 $Å^3$ on average (Figure 2b), which is considerably smaller than the standard deviations of 262.0 $Å^3$ and 185.0 $Å^3$ for *O*- and *N*-targeting targeting MTases, respectively. Using student's t-test, the cavity volume distributions are not considered to be statistically significant with a p-value of 0.20. The upper and lower limits of cavity volumes for both heteroatom targeting enzymes are also comparable (i.e., ~550 $Å^3$ to ~1425 $Å^3$), further suggesting that the enzymes do not directly adjust their volumes to accommodate different heteroatoms.

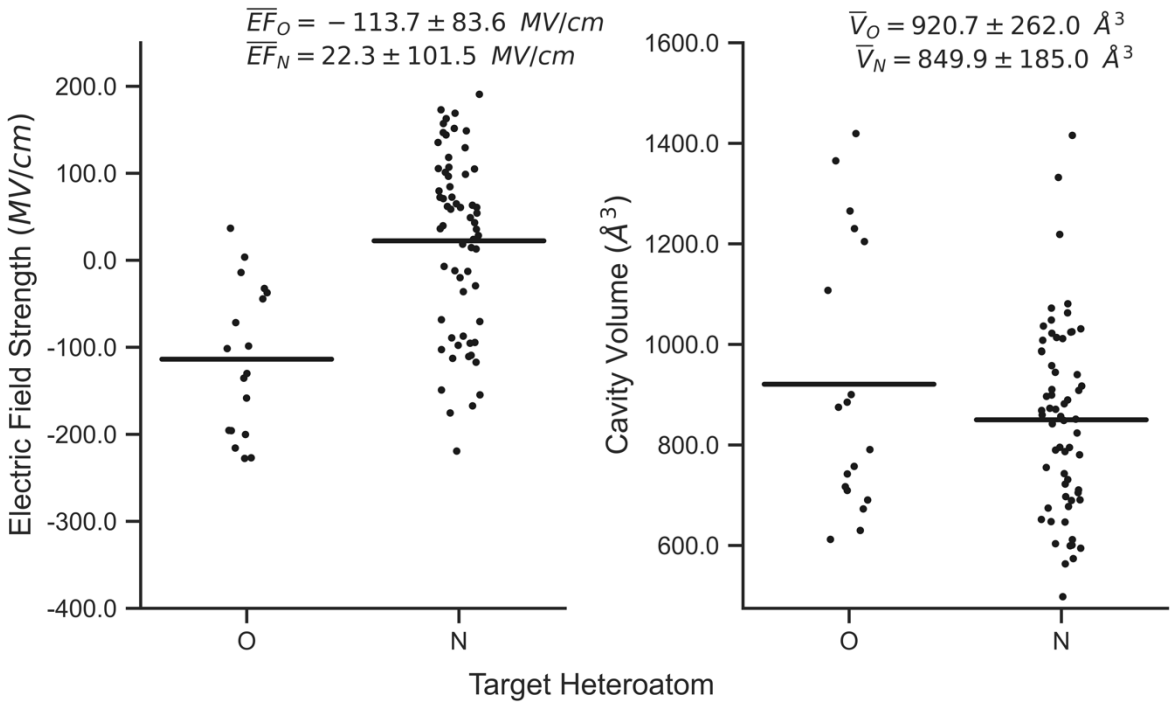


Figure 2. Distributions of measured electric field strength and cavity volume as a function of target heteroatom. The left and right plots show strip plots with the electric field and cavity volume measurements for oxygen (O) and nitrogen (N), respectively. The bars for each data series represent the mean value for that subset. Average values and standard deviations are shown at the top of each subplot. Carbon-targeting MTases do exist in the dataset but have been omitted from this figure as only five data points exist.

The inverse relationship between electric field and electronegativity of target atom is conserved when looking only at MTases targeting RNA/DNA (Figure 3). The MTases targeting carbon atoms are also included in this comparison. *O*-targeting MTases have the most negative electric field strength with an average value of -166.5 MV/cm. *N*-targeting MTases still have negative electric field strength with an average of -11.4 MV/cm, but it is much less negative than those of *O*-targeting counterparts. *C*-targeting is the least electronegative of the three elements and has the most positive electric field strength with an average of 83.2 MV/cm, which is almost four

times higher than the median seen across the larger MTase dataset. This observation is consistent with our hypothesis that MTases with targeting atoms of a weaker electronegativity experience a stronger evolutionary pressure for accelerating the methyl transfer reactions. As such, they evolve to produce more positive electric fields in order to enhance the cleavage of C–S bond for methyl donation.

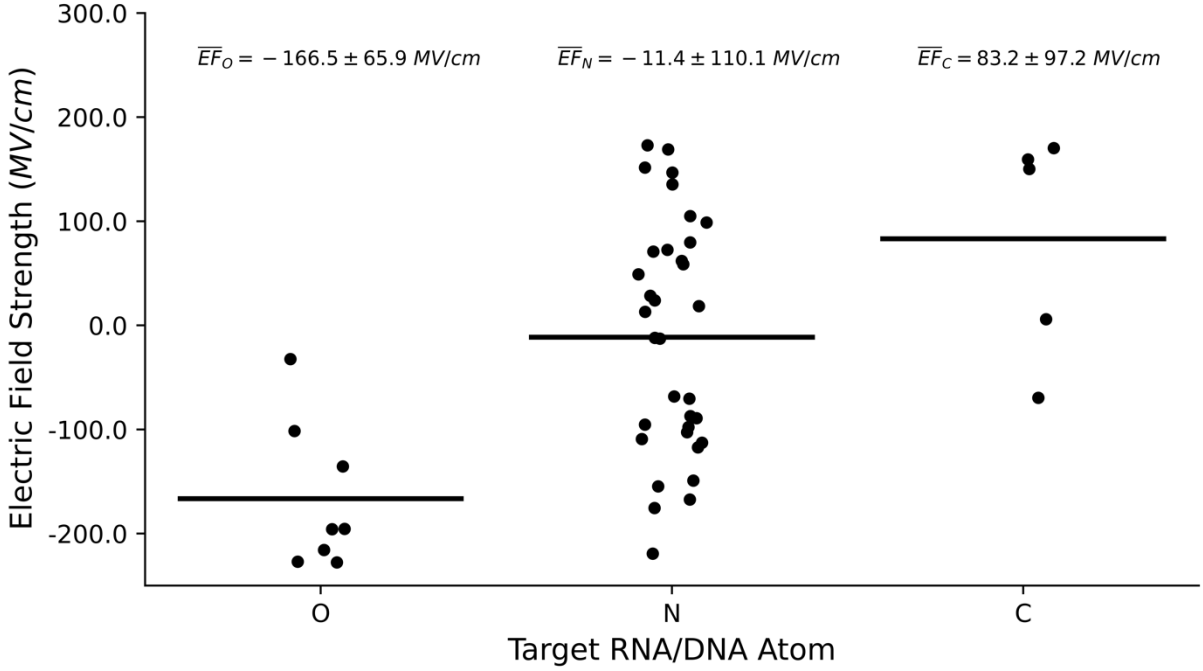


Figure 3. Distributions of measured electric field strength for MTases targeting RNA or DNA substrates. Values are grouped by target atom with oxygen, nitrogen, and carbon being represented by O, N, and C, respectively. The black line for each category represents the average electric field and the average for each atom type plus or minus the standard deviation. Averages and standard deviations for each atom are listed above their respective distribution plots.

Furthermore, we investigated how interior electric field and cavity adapt to substrates with different hybridization state of nitrogen atoms. We emphasized the *N*-targeting MTases for RNA/DNA and histones because they are the two largest substrate types in the dataset with 50 and

21 enzymes catalyzing for these reactants, respectively. Nucleic acids are more diverse, containing nitrogen in both sp^2 and sp^3 hybridization states (Figure 4b). Histone MTases target nitrogen on lysine residues in sp^3 hybridization only. Notably, the amine group with sp^3 N can also have varying protonation states depending on local chemical environment. Wide distribution of electric field strengths along the S–C bonds mirror this diversity, with RNA/DNA and histone MTases having standard deviations of 121.4 MV/cm and 55.3 MV/cm, respectively (Figure 4a). RNA/DNA and histone targeting MTases have average electric fields of -31.7 MV/cm and 68.6 MV/cm, respectively. The stronger electric field for histone MTases may alternatively be due to the protonation state of the terminal nitrogen. In biological pH ranges, it will be in an NH_3^+ meaning the methyl-accepting lone pair on the nitrogen is occupied. The stronger EF strength may partially work to deprotonate the nitrogen and allow methylation to occur. This notion is supported by the differences in EF strength for sp^2 and sp^3-targeting RNA/DNA MTases which have averages of -65.6 MV/cm and 52.9 MV/cm, respectively. This result indicates MTases have directly evolved their EF strengths to adjust to hybridization state of target atom.

Cavity volume spreads repeat the trend seen with electric field spreads with RNA/DNA and histone MTases having standard deviations of 231.0 $Å^3$ and 162.9 $Å^3$, respectively. We believe the increased variation in volume for RNA/DNA targeting enzymes reflects the geometric diversity seen in nucleic acid substrates. Target atoms in nucleic acid substrates have diverse locations within their respective residues and have both sp^2 and sp^3 hybridization whereas all histone targets have the same sp^3 hybridization.

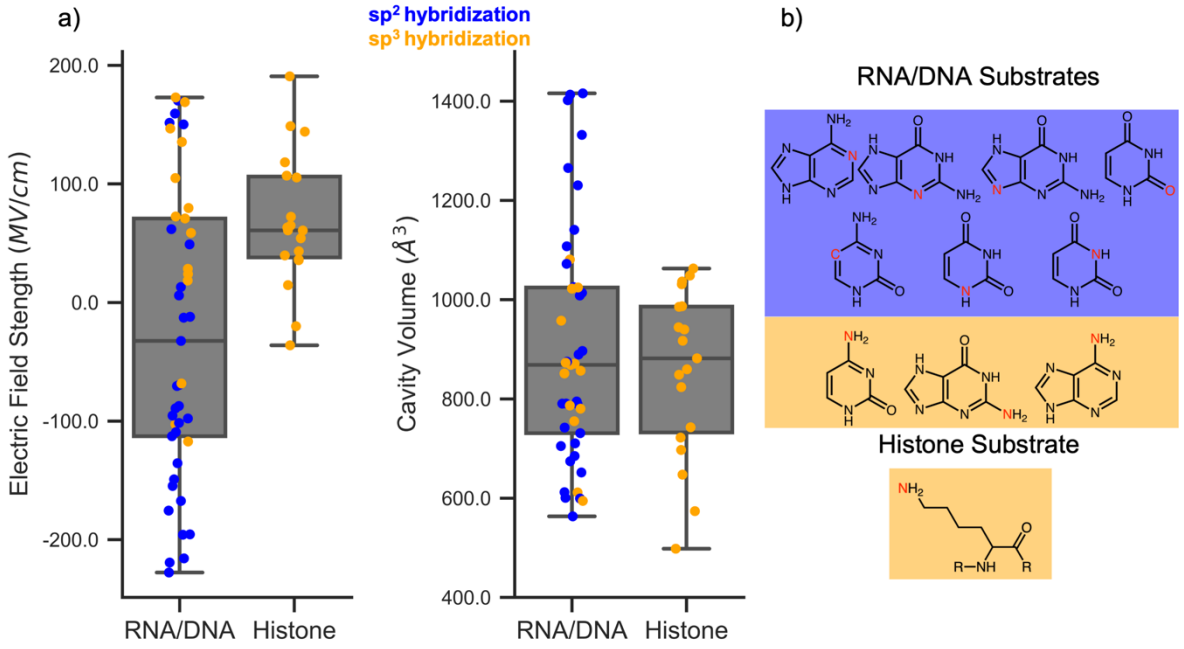


Figure 4. Electric field strength and cavity volume data stratified by target substrate type. Target substrates are determined by EC number and the subplots in (a) show actual data points overlaid onto boxplots. They are color coded with blue corresponding to substrates with the target atom in a sp^2 hybridization state and orange corresponding to when the target atom is in a sp^3 hybridization state. On the right (b) the substrates with target atom highlighted in red. Substrates with sp^2 and sp^3 target atoms have blue and orange backgrounds, respectively.

Finally, we investigated how the presence of metal ions mediate interior electric field of MTases (Figure 5). A total of 19 enzymes have a metal ion present, 2 having magnesium, and 17 having zinc. For 17 of the 19, the metal ion makes the electric field along SAM's S–C bond more negative by an average of -55.5 MV/cm. Metal ions do not typically enhance the electric field strength for the purpose of methyl donation, but their corresponding host protein scaffolds have evolved to offset these effects. MTases with metal ions have an average electric field of 58.8 MV/cm versus the average of -20.1 MV/cm for those without metals (Figure 5a). Given the largely

negative contributions of metal ions, MTases containing metals feature extremely positive contributions from protein scaffold alone. Instead of directly aiding the transfer of SAM's methyl, these binding metal ions likely contribute to stabilize the enzyme structure or mediate protein dynamics. Many of the zinc-containing enzymes have zinc fingertips which stabilize distal regions of enzyme structure.[49, 50] In catechol *O*-methyl transferases (COMT), the magnesium ion stabilizes the catecholate intermediate although it directly worsens the EF strength. Similar to heteroatom-based stratifications, there is an observed upper limit of electric field near ~200 MV/cm for both metal and non-metal MTases. The lower bound for non-metal MTases is -306.2 MV/cm versus -254.9 MV/cm for metal MTases. The consistently tighter EF strength spread and higher range of metal-containing MTases indicates that MTases have consistently evolved to directly offset the negative EF strength contribution of the metal ions contained within.

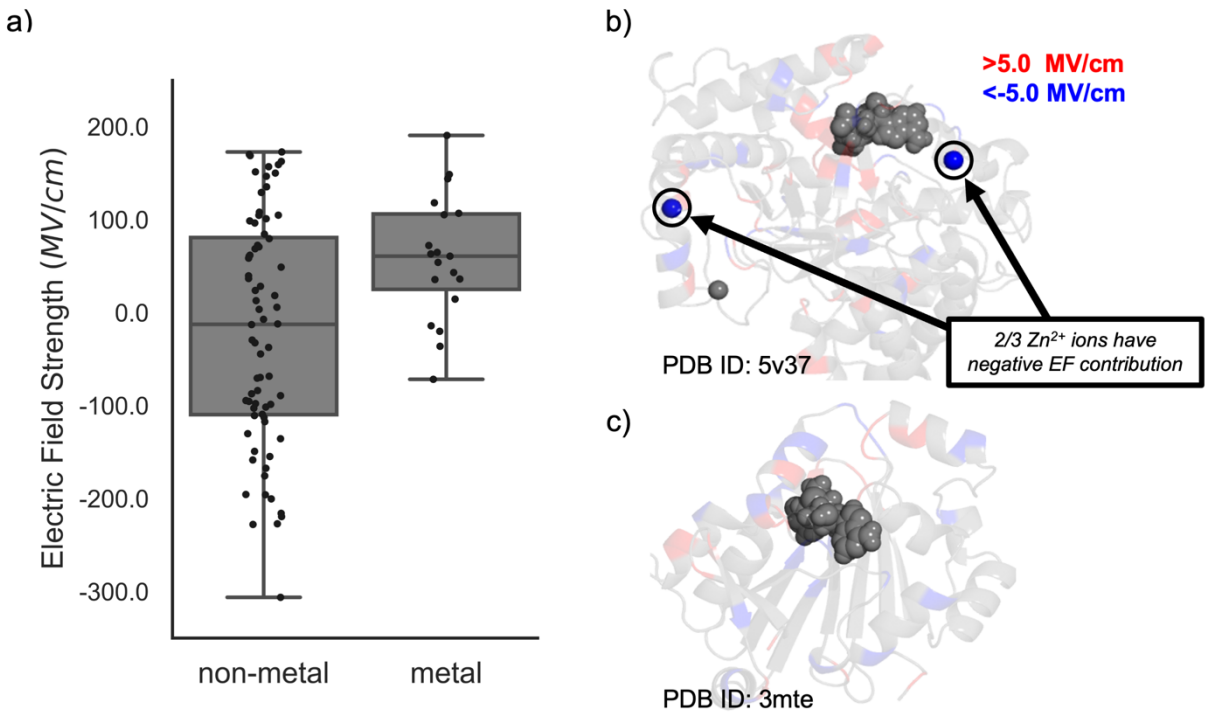


Figure 5. Distributions of electric fields for MTases as a function of containing metal ions. The plot on the left visualizes raw values as black dots and box plots in grey. Inlay (b) shows an example of a structure with three metal ions where two provide negative electric field contribution. The transparent amino acids are colored red and blue when they contribute > 5.0 MV/cm and < -5.0 MV/cm to the electric field, respectively. Inlay (c) shows a structure that has no metal ions present. SAM is rendered as grey spheres in both (b) and (c).

Conclusion

We carried out a high-throughput analysis of 91 SAM MTase structures focusing on how these enzymes have evolved to achieve enzymatic efficiency across a wide range of substrates. We first generated cavity volume and electric field strength values for each structure and then determined the catalytic function of each protein. When looking at *O*- and *N*- targeting MTases, we found there was not significant difference in cavity volumes but that electric field strength and electronegativity of target atom are inversely correlated at a statistically significant level. This trend was also conserved when looking only at MTases that target RNA or DNA substrates, including a small number of *C*-targeting MTases. Comparing values for MTases targeting RNA/DNA and histones, we observed variation in both electric field strength and cavity volume between these categories. More variation was seen in the electric field strength and cavity volume values for RNA/DNA-targeting MTases than histone-targeting MTases which mirrors the associated diversity with each class of substrates. We observed that MTases targeting sp^3 hybridized nitrogen atoms have more positive electric field strengths than those targeting sp^2 atoms, likely to help prepare nitrogen centers to accept a methyl group. Lastly, we investigated the role of metal ions and found that they largely have a negative contribution to electric field strength and that the sequences of metal containing enzymes have evolved to offset this effect.

ASSOCIATED CONTENT

Supporting Information. Values for activation barriers; the list of PDB structures used; the PDB query used to obtains structures; enzymatic function annotations of structures; AMBER minimization settings; explanation of sequence clustering protocol.

Data and Software Availability. The code, structures, and parameter files can all be found at https://github.com/chrisjurich/sam-pockets. The code of EnzyHTP can be found at https://github.com/ChemBioHTP/EnzyHTP. AMBER 19 is available from http://ambermd.org/. Mole2.0 can be found at https://webchem.ncbr.muni.cz/Platform/App/Mole.

AUTHOR INFORMATION

Corresponding Author

*Email: zhongyue.yang@vanderbilt.edu phone: 615-343-9849

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

The authors thank Prof. Jens Meiler for his suggestions and comments. This research was supported by the startup grant from Vanderbilt University. C. Jurich is supposed by Vanderbilt Chemistry-Biology Interface Training Grant (T32GM065086). Z. J. Yang is supported by the National Institute of General Medical Sciences of the National Institutes of Health under award number R35GM146982 and Rosetta Common Seed grant.

References

1. Bügl, H.; Fauman, E. B.; Staker, B. L.; Zheng, F.; Kushner, S. R.; Saper, M. A.; Bardwell, J. C.; Jakob, U., RNA methylation under heat shock control. *Mol Cell* **2000,** *6* (2), 349-60.

2. Nai, Y.-S.; Huang, Y.-C.; Yen, M.-R.; Chen, P.-Y., Diversity of Fungal DNA Methyltransferases and Their Association With DNA Methylation Patterns. *Frontiers in Microbiology* **2021,** *11*.
3. Dhe-Paganon, S.; Syeda, F.; Park, L., DNA methyl transferase 1: regulatory mechanisms and implications in health and disease. *Int J Biochem Mol Biol* **2011,** *2* (1), 58-66.
4. Zhang, H.; Lang, Z.; Zhu, J.-K., Dynamics and function of DNA methylation in plants. *Nature Reviews Molecular Cell Biology* **2018,** *19* (8), 489-506.
5. Katada, S.; Sassone-Corsi, P., The histone methyltransferase MLL1 permits the oscillation of circadian gene expression. *Nature Structural & Molecular Biology* **2010,** *17* (12), 1414-1421.
6. Min, J.; Feng, Q.; Li, Z.; Zhang, Y.; Xu, R.-M., Structure of the Catalytic Domain of Human DOT1L, a Non-SET Domain Nucleosomal Histone Methyltransferase. *Cell* **2003,** *112* (5), 711-723.
7. Moore, L. D.; Le, T.; Fan, G., DNA methylation and its basic function. *Neuropsychopharmacology* **2013,** *38* (1), 23-38.
8. Falnes, Pål Ø.; Jakobsson, Magnus E.; Davydova, E.; Ho, A.; Małecki, J., Protein lysine methylation by seven-β-strand methyltransferases. *Biochemical Journal* **2016,** *473* (14), 1995-2009.
9. Winter, D. L.; Hart-Smith, G.; Wilkins, M. R., Characterization of Protein Methyltransferases Rkm1, Rkm4, Efm4, Efm7, Set5 and Hmt1 Reveals Extensive Post-Translational Modification. *J Mol Biol* **2018,** *430* (1), 102-118.
10. Paik, W. K.; Kim, S.; Lim, I. K., Protein methylation and interaction with the antiproliferative gene, BTG2/TIS21/Pc3. *Yonsei Med J* **2014,** *55* (2), 292-303.
11. Männistö, P. T.; Kaakkola, S., Catechol-O-methyltransferase (COMT): Biochemistry, Molecular Biology, Pharmacology, and Clinical Efficacy of the New Selective COMT Inhibitors. *Pharmacological Reviews* **1999,** *51* (4), 593-628.
12. Witte, A. V.; Flöel, A., Effects of COMT polymorphisms on brain function and behavior in health and disease. *Brain Research Bulletin* **2012,** *88* (5), 418-428.
13. Akil, M.; Kolachana, B. S.; Rothmond, D. A.; Hyde, T. M.; Weinberger, D. R.; Kleinman, J. E., Catechol-O-methyltransferase genotype and dopamine regulation in the human brain. *J Neurosci* **2003,** *23* (6), 2008-13.
14. Liscombe, D. K.; Louie, G. V.; Noel, J. P., Architectures, mechanisms and molecular evolution of natural product methyltransferases. *Natural Product Reports* **2012,** *29* (10), 1238-1250.
15. Ohashi, M.; Liu, F.; Hai, Y.; Chen, M.; Tang, M. C.; Yang, Z.; Sato, M.; Watanabe, K.; Houk, K. N.; Tang, Y., SAM-dependent enzyme-catalysed pericyclic reactions in natural product biosynthesis. *Nature* **2017,** *549* (7673), 502-506.
16. Ohashi, M.; Tan, D.; Lu, J.; Jamieson, C. S.; Kanayama, D.; Zhou, J.; Houk, K. N.; Tang, Y., Enzymatic cis-Decalin Formation in Natural Product Biosynthesis. *Journal of the American Chemical Society* **2023**.
17. Patra, N.; Ioannidis, E. I.; Kulik, H. J., Computational Investigation of the Interplay of Substrate Positioning and Reactivity in Catechol O-Methyltransferase. *PLOS ONE* **2016,** *11* (8), e0161868.
18. Zhang, J.; Kulik, H. J.; Martinez, T. J.; Klinman, J. P., Mediation of donor–acceptor distance in an enzymatic methyl transfer reaction. *Proceedings of the National Academy of Sciences* **2015,** *112* (26), 7954-7959.

19. Ruggiero, G. D.; Williams, I. H.; Roca, M.; Moliner, V.; Tuñón, I., QM/MM Determination of Kinetic Isotope Effects for COMT-Catalyzed Methyl Transfer Does Not Support Compression Hypothesis. *Journal of the American Chemical Society* **2004,** *126* (28), 8634-8635.
20. Soriano, A.; Castillo, R.; Christov, C.; Andrés, J.; Moliner, V.; Tuñón, I., Catalysis in glycine N-methyltransferase: testing the electrostatic stabilization and compression hypothesis. *Biochemistry* **2006,** *45* (50), 14917-25.
21. Horowitz, S.; Adhikari, U.; Dirk, L. M. A.; Del Rizzo, P. A.; Mehl, R. A.; Houtz, R. L.; Al-Hashimi, H. M.; Scheiner, S.; Trievel, R. C., Manipulating Unconventional CH-Based Hydrogen Bonding in a Methyltransferase via Noncanonical Amino Acid Mutagenesis. *ACS Chemical Biology* **2014,** *9* (8), 1692-1697.
22. Roca, M.; Williams, I. H., Transition-State Vibrational Analysis and Isotope Effects for COMT-Catalyzed Methyl Transfer. *Journal of the American Chemical Society* **2020,** *142* (36), 15548-15559.
23. Świderek, K.; Tuñón, I.; Williams, I. H.; Moliner, V., Insights on the Origin of Catalysis on Glycine N-Methyltransferase from Computational Modeling. *Journal of the American Chemical Society* **2018,** *140* (12), 4327-4334.
24. Zhang, J.; Balsbaugh, J. L.; Gao, S.; Ahn, N. G.; Klinman, J. P., Hydrogen deuterium exchange defines catalytically linked regions of protein flexibility in the catechol O-methyltransferase reaction. *Proceedings of the National Academy of Sciences* **2020,** *117* (20), 10797-10805.
25. Zhang, J.; Klinman, J. P., Convergent Mechanistic Features between the Structurally Diverse N- and O-Methyltransferases: Glycine N-Methyltransferase and Catechol O-Methyltransferase. *Journal of the American Chemical Society* **2016,** *138* (29), 9158-9165.
26. 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.; Trievel, R. C., Conservation and functional importance of carbon-oxygen hydrogen bonding in AdoMet-dependent methyltransferases. *J Am Chem Soc* **2013,** *135* (41), 15536-48.
27. Couture, J.-F.; Hauk, G.; Thompson, M. J.; Blackburn, G. M.; Trievel, R. C., Catalytic Roles for Carbon-Oxygen Hydrogen Bonding in SET Domain Lysine Methyltransferases *. *Journal of Biological Chemistry* **2006,** *281* (28), 19280-19287.
28. Trievel, R. C.; Flynn, E. M.; Houtz, R. L.; Hurley, J. H., Mechanism of multiple lysine methylation by the SET domain enzyme Rubisco LSMT. *Nature Structural & Molecular Biology* **2003,** *10* (7), 545-552.
29. Yang, Z.; Liu, F.; Steeves, A. H.; Kulik, H. J., Quantum Mechanical Description of Electrostatics Provides a Unified Picture of Catalytic Action Across Methyltransferases. *The Journal of Physical Chemistry Letters* **2019,** *10* (13), 3779-3787.
30. Yan, B.; Ran, X.; Gollu, A.; Cheng, Z.; Zhou, X.; Chen, Y.; Yang, Z. J., IntEnzyDB: an Integrated Structure–Kinetics Enzymology Database. *Journal of Chemical Information and Modeling* **2022,** *62* (22), 5841-5848.
31. Yan, B.; Ran, X.; Jiang, Y.; Torrence, S. K.; Yuan, L.; Shao, Q.; Yang, Z. J., Rate-Perturbing Single Amino Acid Mutation for Hydrolases: A Statistical Profiling. *The Journal of Physical Chemistry B* **2021,** *125* (38), 10682-10691.

32. Sousa, S. F.; Calixto, A. R.; Ferreira, P.; Ramos, M. J.; Lim, C.; Fernandes, P. A., Activation Free Energy, Substrate Binding Free Energy, and Enzyme Efficiency Fall in a Very Narrow Range of Values for Most Enzymes. *ACS Catalysis* **2020,** *10* (15), 8444-8453.
33. Rod, T. H.; Ryde, U., Accurate QM/MM Free Energy Calculations of Enzyme Reactions: Methylation by Catechol O-Methyltransferase. *Journal of Chemical Theory and Computation* **2005,** *1* (6), 1240-1251.
34. Shao, Q.; Jiang, Y.; Yang, Z. J., EnzyHTP: A High-Throughput Computational Platform for Enzyme Modeling. *Journal of Chemical Information and Modeling* **2022,** *62* (3), 647-655.
35. Berman, H. M.; Gierasch, L. M., How the Protein Data Bank changed biology: An introduction to the JBC Reviews thematic series, part 1. *Journal of Biological Chemistry* **2021,** *296*, 100608.
36. Vasina, M.; Velecký, J.; Planas-Iglesias, J.; Marques, S. M.; Skarupova, J.; Damborsky, J.; Bednar, D.; Mazurenko, S.; Prokop, Z., Tools for computational design and high-throughput screening of therapeutic enzymes. *Advanced Drug Delivery Reviews* **2022,** *183*, 114143.
37. Lameira, J.; Bora, R. P.; Chu, Z. T.; Warshel, A., Methyltransferases do not work by compression, cratic, or desolvation effects, but by electrostatic preorganization. *Proteins* **2015,** *83* (2), 318-30.
38. Oanca, G.; Asadi, M.; Saha, A.; Ramachandran, B.; Warshel, A., Exploring the Catalytic Reaction of Cysteine Proteases. *The Journal of Physical Chemistry B* **2020,** *124* (50), 11349-11356.
39. Ji, Z.; Kozuch, J.; Mathews, I. I.; Diercks, C. S.; Shamsudin, Y.; Schulz, M. A.; Boxer, S. G., Protein Electric Fields Enable Faster and Longer-Lasting Covalent Inhibition of β-Lactamases. *Journal of the American Chemical Society* **2022,** *144* (45), 20947-20954.
40. Zheng, C.; Mao, Y.; Kozuch, J.; Atsango, A. O.; Ji, Z.; Markland, T. E.; Boxer, S. G., A two-directional vibrational probe reveals different electric field orientations in solution and an enzyme active site. *Nature Chemistry* **2022,** *14* (8), 891-897.
41. Webb, B.; Sali, A., Protein Structure Modeling with MODELLER. *Methods in molecular biology (Clifton, N.J.)* **2014,** *1137*, 1-15.
42. ULC, C. C. G., Molecular Operating Environment (MOE). **2022.02**.
43. Coudert, E.; Gehant, S.; de Castro, E.; Pozzato, M.; Baratin, D.; Neto, T.; Sigrist, C. J. A.; Redaschi, N.; Bridge, A.; Consortium, T. U., Annotation of biologically relevant ligands in UniProtKB using ChEBI. *Bioinformatics* **2022,** *39* (1).
44. Chang, A.; Jeske, L.; Ulbrich, S.; Hofmann, J.; Koblitz, J.; Schomburg, I.; Neumann-Schaal, M.; Jahn, D.; Schomburg, D., BRENDA, the ELIXIR core data resource in 2021: new developments and updates. *Nucleic Acids Research* **2020,** *49* (D1), D498-D508.
45. Kamenik, A. S.; Handle, P. H.; Hofer, F.; Kahler, U.; Kraml, J.; Liedl, K. R., Polarizable and non-polarizable force fields: Protein folding, unfolding, and misfolding. *The Journal of Chemical Physics* **2020,** *153* (18), 185102.
46. Sehnal, D.; Svobodová Vařeková, R.; Berka, K.; Pravda, L.; Navrátilová, V.; Banáš, P.; Ionescu, C.-M.; Otyepka, M.; Koča, J., MOLE 2.0: advanced approach for analysis of biomacromolecular channels. *Journal of Cheminformatics* **2013,** *5* (1), 39.
47. Jindal, G.; Warshel, A., Exploring the Dependence of QM/MM Calculations of Enzyme Catalysis on the Size of the QM Region. *The Journal of Physical Chemistry B* **2016,** *120* (37), 9913-9921.

48. Kulik, H. J.; Zhang, J.; Klinman, J. P.; Martínez, T. J., How Large Should the QM Region Be in QM/MM Calculations? The Case of Catechol O-Methyltransferase. *The Journal of Physical Chemistry B* **2016,** *120* (44), 11381-11394.
49. Bogani, D.; Morgan, M. A.; Nelson, A. C.; Costello, I.; McGouran, J. F.; Kessler, B. M.; Robertson, E. J.; Bikoff, E. K., The PR/SET domain zinc finger protein Prdm4 regulates gene expression in embryonic stem cells but plays a nonessential role in the developing mouse embryo. *Mol Cell Biol* **2013,** *33* (19), 3936-50.
50. Klug, A.; Schwabe, J. W. R., Zinc fingers. *The FASEB Journal* **1995,** *9* (8), 597-604.

Table of Contents Graphic