MIT Open Access Articles

Quantum Mechanical Description of Electrostatics Provides a Unified Picture of Catalytic Action Across Methyltransferases

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

As Published: 10.1021/acs.jpcllett.9b01555

Publisher: American Chemical Society (ACS)

Persistent URL: <https://hdl.handle.net/1721.1/134774>

Version: Original manuscript: author's manuscript prior to formal peer review

Terms of Use: Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.

A Quantum Mechanical Description of Electrostatics

Provides a Unified Picture of Catalytic Action

Across Methyltransferases

Zhongyue Yang¹, Fang Liu¹, Adam H. Steeves¹, and Heather J. Kulik^{1,*}

¹*Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA*

02139

AUTHOR INFORMATION

Corresponding Author

*email: hjkulik@mit.edu, phone: 617-253-4584

ABSTRACT: Methyl transferases (MTases) are a well-studied class of enzymes for which competing enzymatic enhancement mechanisms have been suggested, ranging from structural methyl group C-H $\cdots$ X hydrogen bonds (HBs) to electrostatic- and charge-transfer-driven stabilization of the transition state (TS). We identified all Class I MTases for which reasonable resolution (< 2.0 Å) crystal structures could be used to form catalytically competent ternary complexes for multi-scale (i.e., quantum-mechanical/molecular-mechanical or QM/MM) simulation of the S_N2 methyl transfer reaction coordinate. The four Class I MTases studied have both distinct functions (e.g., protein repair or biosynthesis) and substrate nucleophiles (i.e., C, N, or O). While CH $\cdots$ X HBs stabilize all reactant complexes, no universal TS stabilization role is found for these interactions in MTases. A consistent picture is instead obtained through analysis of charge transfer and electrostatics, wherein the majority of cofactor-substrate charge separation is maintained in the TS region, and electrostatic potential is correlated with substrate nucleophilicity (i.e., intrinsic reactivity).

TOC GRAPHICS

Methyltransferases (MTases)¹ are among the most ubiquitous regulatory enzymes in the cell, catalyzing gene signaling², protein repair³⁻⁴, neurotransmitter regulation⁵, and natural product biosynthesis. Most MTases use the co-factor S-adenosyl methionine (SAM) to donate a methyl group to a range of substrate nucleophiles through a general S_N2 mechanism⁶⁻¹⁹. Despite this deceptively simple reaction, no universal role has been conclusively demonstrated for the greater enzyme environment in catalyzing methyl transfer. Commonly suggested models include those geometric in nature, e.g., compaction of the active site^{7, 20-21}, particularly through formation of methyl group C-H···O hydrogen bonds (HBs)²²⁻³⁰ or in dynamically formed near-attack conformations^{15, 23}. Alternatively, predominantly electrostatic contributions have been suggested³¹ including SAM–substrate charge transfer^{6-8, 32-35} or modification of the electrostatic potential³⁶⁻³⁸ experienced by the methyl group^{9, 14, 28, 39-41}, in accordance with recent observations on related enzymes.⁴²⁻⁴⁴

No broad conclusion can be reached on the relative weight of these roles in MTase activity because each study is typically carried out on a single MTase and associated substrate. One of the few exceptions to this is a recent systematic survey by Trievel and coworkers²³ of 46 SAM-bound MTase X-ray crystal structures with < 2.0 Å resolution, in which they showed the importance of methyl group C-H···O HBs in SAM binding, including through minimal model quantum-mechanical (QM) calculations. Unfortunately, minimal models are poor proxies for the enzyme environments in MTases: gas phase methyl transfer from SAM is a barrierless process, whereas enzyme-catalyzed methyl transfer barriers are significant (ca. 20 kcal/mol^{10, 12, 21, 45}). Prior studies on minimal QM models led to conclusions of charge neutralization of the oppositely charged substrates in the transition state⁴⁶, but subsequent larger scale studies have shown significant charge separation until the products are reached.^{6-7, 33-34} As a result, multi-scale

quantum-mechanical/molecular-mechanical (QM/MM) modeling is essential, with at least substrates and adjacent residues^{6, 32, 41, 47} treated with QM, owing to the poor treatment of non-covalent interactions (e.g., the C-H $\cdots$ O HBs of interest here) by force fields⁴⁸ or across the QM/MM boundary⁴⁹.

To address the question of the relative roles of competing models for MTase activity, we carried out a systematic survey of structurally characterized MTases to identify candidates for large-scale QM/MM simulation. These candidates were curated from the protein data bank (PDB) in January 2017 with a focus on the most abundant, Class-I MTases¹. We required 2.0 Å or better resolution of at least one crystal structure corresponding to a published manuscript, mandated that a catalytically competent ternary complex (i.e., with bound SAM and substrate) could be constructed with the SAM methyl group within 4.0 Å of the substrate methyl-accepting atom after superposition of available crystal structures without any adjustment, and finally selected for unique enzyme function (Supporting Information Tables S1-S2). From 72 unique crystal structures (resolution: 1.35-2.0 Å), 40 were determined to be published Class I MTases, only 6 of which formed valid ternary complexes (Supporting Information Tables S1-S2). This number is much smaller than the 46 crystal structures that had been identified in a prior survey²³ due to our more stringent criteria (Supporting Information Tables S1-S2).

After excluding catechol O-methyltransferase (COMT) due to exhaustive prior computational studies^{6, 8, 14, 17-18, 29, 31, 33, 41, 50-51} as well as one MTase for which a feasible reaction coordinate could not be obtained⁵², our curation yielded four Class-I MTases (Supporting Information Tables S1-S2). The MTases CMTr1⁵³, PIMT⁵⁴, PfPMT⁵⁵, and HcgC⁵⁶ play diverse regulatory⁵³⁻⁵⁴ and biosynthetic^{55,56} roles (Figure 1). CMTr1 is a human 2'-O-ribose MTase, catalyzing the 5'-capping of the first transcribed nucleotide of human messenger RNA⁵³

for translation and stability.^{53, 57-59} PIMT is a protein repair MTase from *Pyrococcus furiosus*, which methylates the carboxylate group of the isomerized and racemized aspartyl residues to tag age-damaged proteins.⁵⁴ Phosphoethanolamine MTase, PfPMT, methylates phosphoethanolamine (PEA) to form phosphocholine for membrane biogenesis in the malaria parasite, *Plasmodium falciparum*⁵⁵. HcgC is a C-MTase that methylates the enolate carbanion of 6-carboxymethyl-5-methyl-4-hydroxy-2-pyridinol (CMHP)⁵⁶. Thus, the methyl acceptor in each substrate is distinct: hydroxyl oxygen in CMTr1, carboxylate oxygen in PIMT, amine nitrogen in PfPMT, and enolate carbon in HcgC (Figure 1).

Figure 1. Active site structures for methyltransferases CMTr1, PIMT, PfPMT, and HcgC. SAM, G, IAS, PEA, and CMHP refer to S-adenosyl methionine, guanine, isoaspartate, phosphoethanolamine, and 6-carboxy-methyl-4-hydroxy-2-pyridinol, respectively. CMTr1, PIMT, PfPMT, and HcgC are colored in green, blue, orange and grey, respectively. The substrate and co-factor SAM are shown in blue sticks. The methyl donor sulfur and the methyl acceptor nucleophile are shown as spheres in all cases.

Although all four enzymes are Class I MTases, the key residues expected to facilitate methyl transfer are diverse (Figure 1). Charged residues (e.g., Lys239, Glu440, and Asp364) have been implicated in binding, deprotonating, and facilitating methyl transfer of the highly charged mRNA substrate in CMTr1, similar to COMT^{11, 60} (Figure 1). Polar residues, especially with hydroxyl functional groups, have been implicated in substrate binding and catalysis in the

remaining three enzymes (e.g., Thr73 and Thr165 in PIMT, Tyr19 and His132 in PfPMT, and Thr6', a water, or Ile115 backbone in HcgC) by crystallographic studies^{54, 56} and in some cases confirmed by mutagenesis and calorimetry⁵⁵ (Figure 1).

We next performed large-scale electronic structure^{6, 61-63} QM/MM^{7, 41, 64-69} steered-molecular dynamics (SMD)⁷⁰⁻⁷³ simulations (Supporting Information Figure S1). This approach, combined with a fast SMD pulling speed, was designed to maximize correspondence with prior work²³ that had focused on the role of crystal structure atomic positions in MTase activity, rather than on method-⁷⁴, model-^{6, 32, 41}, and sampling⁷⁵⁻⁷⁷-sensitive barrier height estimation. The 250-600 atom QM regions were rationally selected using charge-shift analysis^{6, 32-35}, wherein any residues that exhibit charge transfer to the central reacting species (here, SAM and the substrate) are included in the QM region (see Computational Details and Supporting Information Tables S3-S6). As in prior work,^{6, 8-9, 28, 33, 39} we used the difference between the C-S and C-X bond distances, $d_{C-S} - d_{C-X}$, as the SMD reaction coordinate (RC), which corresponds to the cleavage of the C-S bond in SAM and the formation of the C-X (here, X = C, N, or O) bond (Supporting Information Figures S2-S5). In lieu of a rigorously characterized transition state (TS), we define a TS region by the midpoint relative distance of the methyl group between donor and acceptor atoms as judged by the van der Waals (vdW) radii⁷⁸⁻⁷⁹ (see Computational Details).

To enable comparison to recent static analyses of MTase binary structures²³, we first geometrically quantified CH $\cdots$ X (X=N or O) HBs along the RC by setting a generous cutoff of < 2.5 Å for the non-bonded CH $\cdots$ X distance along with a simultaneous CH $\cdots$ X angle > 120°. Based on these criteria, CH $\cdots$ X HBs are observed for a portion of the RC for all enzymes studied (Figure 2 and Supporting Information Figure S6). However, relatively few persist at the short distances (C-H $\cdots$ X ca. 1.5-2.0 Å) observed in prior analysis²³, despite no bias in our QM

simulation against short interatomic distances.

Figure 2. Variation of distances (black) and angles (red) of representative $\text{CH}\cdots\text{X}$ ($\text{X} = \text{O}$ or N) bonds in CMTr1 (top, left), PIMT (top, right), PfPMT (bottom, left), and HcgC (bottom, right). Inset active site geometries display all $\text{CH}\cdots\text{X}$ geometric HBs with surrounding electronegative atoms, with the displayed HB shown in green. The cutoff for distance (2.5 Å, left axis) and angle (120°, right axis) is indicated as a single horizontal dashed line. Individual points are shown as translucent circles with a 40-point running average shown as a solid line. An approximate TS region is indicated by a ± 0.1 Å-width shaded region centered on the estimated TS (black dashed vertical line).

Although most potential HBs do not satisfy the geometric criteria for more than a small portion of the RC, the most enduring HB in CMTr1, a $\text{CH}\cdots\text{N}$ HB from the SAM methyl to the guanine nitrogen, persists until the RC reaches a product-like distance of ~ 0.9 Å ($d_{\text{C-S}} = 2.5$ Å, $d_{\text{C-N}} = 1.6$ Å, Figure 2). In PIMT and PfPMT, $\text{CH}\cdots\text{O}$ hydrogen bonds between the SAM methyl and adjacent hydroxyl groups of Thr165 and Tyr19, respectively, are present in the reactant complex but persist only until approximately the TS at ~ 0.2 Å ($d_{\text{C-S}} = 2.3$ Å, $d_{\text{C-O}} = 2.1$ Å). Unlike the other three cases, no protein-substrate HBs are present in HcgC, but the SAM methyl forms a weak $\text{CH}\cdots\text{O}$ HB to a conserved, crystallographic water molecule intermittently both

before and after the TS region. This observation in HcgC suggests a potential dual role of water molecules in stabilizing the substrate enolate and in promoting the methyl transfer process.

In order to quantify the energetic contribution of $\text{CH}\cdots\text{X}$ HBs to enzyme catalysis, we estimated HB energies⁸⁰⁻⁸¹ from bond critical point (BCP) potential energy densities (Figure 3 and Supporting Information Figures S7-S10). As motivated by geometric analysis, we group and sum HB energies from BCP analysis with the SAM methyl H atoms into two categories: i) those with methyl-acceptor substrate atoms (Sub-HBs) and ii) those with protein residues (Res-HBs). Consistent with geometric observations, multiple HBs occur for at least some portion of the RC for all four enzymes studied, with individual HBs reaching values as high as 8 kcal/mol, albeit for small portions of the RC (Supporting Information Figures S7-S10). Summed Sub-HB and Res-HB interactions reach maximum values of around 10 and 8 kcal/mol respectively but values averaged over the RC are much lower (< 5 kcal/mol), significantly weaker than typical (e.g., $\text{O}\cdots\text{H}$) HBs in proteins⁸²⁻⁸⁵ (Figure 3).

Figure 3. Variation of $\text{CH}\cdots\text{X}$ ($\text{X}=\text{O}/\text{N}$) hydrogen bonding energies (in kcal/mol) to the SAM methyl group in MTases: CMTr1 (top, left), PIMT (top, right), PfPMT (bottom, left), and HcgC (bottom, right). The $\text{CH}\cdots\text{X}$ HBs are grouped into Sub-HBs (cyan points) or Res-HBs (green

points), and the reported energies reflect the sum of all HBs of each type. 40-point running averages are shown as solid lines. Active site structures are shown in insets with BCPs of Sub-HBs (cyan) and Res-HBs (green) spheres, respectively. An approximate TS region is indicated by a ± 0.1 Å-width shaded region centered on the estimated TS (red dashed vertical line).

Sub-HB interactions are only observed in the O-type CMTr1 and PIMT MTases. The CMTr1 Sub-HBs to guanine and adenine of the mRNA substrate (ca. 7 kcal/mol) vary significantly across the RC (fluctuations ca. 3-4 kcal/mol) and are enhanced after the TS (to 9 kcal/mol) until they disappear in the product (Figure 3). This contribution is much larger than the Res-HB (2 kcal/mol) to nearby Asp364, which is more stable across the RC (Figure 3). Similarly strong Sub-HB interactions in the PIMT reactant complex (7 kcal/mol) drop rapidly (to 1 kcal/mol) in the TS (Figure 3). The weaker Res-HB interactions from nearby Thr165 and Thr73 HB acceptor hydroxyl oxygen atoms increase slightly from reactant (3 kcal/mol) to TS (4 kcal/mol, Figure 3).

In N-type PfPMT and C-type HcgC, Res-HB interactions are generally more significant (up to 6-7 kcal/mol) than in O-type MTases (Figure 3). For PfPMT, the nearby Tyr19 hydroxyl and Asp128 backbone carbonyl oxygen atoms act as HB acceptors, with increasing HB strength from the reactant (2 kcal/mol) to the TS (7 kcal/mol, Figure 3). In addition to the conserved crystallographic water we noted in geometric analysis of HcgC, other water molecules and an Ile115 backbone carbonyl act as HB acceptors with increasing strength from reactants (3 kcal/mol) to the TS (5 kcal/mol, Figure 3 and Supporting Information Figure S10). Over all MTases studied, the $\text{CH}\cdots\text{X}$ Res-HBs stabilize reactant complex SAM binding by 1-3 kcal/mol, in qualitative agreement with prior hydrogen bonding analysis on MTases.^{22-25, 27, 30} The expected effects on MTase activity however are variable: no or limited enhancement (i.e., by TS stabilization) is observed for the O-type CMTr1 and PIMT, whereas some (2-5 kcal/mol) enhancement is expected for N-type PfPMT and C-type HcgC (Figure 3).

Next, we computed the charge on SAM and the substrate along the RC to identify if charge transfer with the greater enzyme environment leads to favorable electrostatic contributions (Figure 4 and see Computational Details). In COMT^{6, 33} and the MTases studied here, both deviation of substrate charges from formal expected values and preservation of charge separation in the TS region are observed, suggesting a universal, significant role of the enzyme with respect to the uncatalyzed reaction (Figure 4). A consistent picture of charge separation in the TS region is observed across O-type CMTr1 (difference of reactant charges preserved in the TS: 68%) and PIMT (60%), albeit with persistent separation in the product only for CMTr1 (SAM: 0.3 and G: -0.7 a.u.) due to the -2 formal charge of the guanine reactant in comparison to the neutralized PIMT substrate (SAM and IAS both: -0.1 a.u.).

Figure 4. Charge transfer between co-factor SAM (red) and the substrates (blue) along the RC for CMTr1 (G, top, left), PIMT (IAS, top, right), PfPMT (OPE, bottom, left), and HcgC (MHP, bottom, right). Individual points are shown as circles as well as a running average 40 points in length shown as a solid line. The TS point along the RC was estimated by the difference of van der Waals radii of the methyl donor and acceptor (black vertical dashed line), and an

approximate transition state region is indicated by a ± 0.1 Å-width shaded region centered around this point. Skeleton structures of the RC are shown in inset for each enzyme.

Despite distinct nucleophiles, N-type PfPMT and C-type HcgC consistently exhibit charge separation in the TS (PfPMT: 58% and HcgC: 80%, see Figure 4). The lack of complete charge neutralization of HcgC products (SAM: 0.0 a.u., MHP: -1.0 a.u.) is due to the -3 formal charge on the reactants, whereas the PfPMT charge profile is more similar to the O-type PIMT (SAM: -0.3 a.u. OPE: -0.1 a.u., see Figure 4). Thus charge transfer between the substrate and MTase environment appears to be more strongly dictated by the formal charge of the reacting substrate than methyl acceptor identity. In PfPMT, significant fluctuations around the reactant complex (ca. -1.2 to -0.7 Å) are due to an $\text{OH}\cdots\text{N}$ HB between nearby Tyr19 and the OPE substrate that forms only for that portion of the RC, highlighting the role of directional hydrogen bonding in mediating substrate charges (Figure 4). On the whole, we conclude that in QM/MM simulations in which charge transfer is permitted between the substrate and greater enzyme environment, much (ca. 60-80%) of the reactant complex charge separation between SAM and the substrate is preserved well past the TS region, with variations dictated by methyl-acceptor substrate formal charge and temporary HB interactions arising along the RC (Figure 4).

Given the lack of charge neutralization at the TS, we finally investigated the electrostatic potential that the four MTases exert on the transferring methyl carbon in the vicinity of the TS region (-0.5 to 0.5 Å) of the RC (Figure 5 and see Computational Details). Indeed, in all four MTases, we observe negative electrostatic potentials that stabilize the positively charged methyl group, albeit with significant variations between enzymes that are larger than the fluctuations ($\sim 40 \text{ kJ}\cdot\text{mol}^{-1}\cdot\text{e}^{-1}$) between snapshots (Figure 5).

Figure 5. Distributions of electrostatic potential of the transferring methyl carbon in the four MTases. The dashed line refers to the average of the electrostatic potential, and the average value, standard deviation, and standard error (in parentheses) labeled. Potentials were computed over all snapshots with RC values ranging from -0.5 to 0.5 Å.

To rationalize why stronger electrostatic potentials are observed in some enzymes than others, we considered substrate differences. The pK_a values of the methyl-accepting functional groups or their conjugate acids provide a good metric for intrinsic (i.e., uncatalyzed) reactivity, where the conjugate base of a lower- pK_a functional group is usually a weaker nucleophile and should be less reactive for S_N2 methyl transfer. The functional group pK_a s span a wide range: 4 for the carboxylic acid⁸⁶ in PIMT, 11 for the protonated amine⁸⁶ in PfPMT, 14 for the hydroxyl⁸⁷ in CMTr1, and 19 for the enol group⁸⁸ in HcgC. The strongest averaged electrostatic potentials in PIMT ($-307 \text{ kJ}\cdot\text{mol}^{-1}\cdot\text{e}^{-1}$) or PfPMT ($-269 \text{ kJ}\cdot\text{mol}^{-1}\cdot\text{e}^{-1}$) correspond to the expected lower nucleophilicity of the methyl acceptors in these reactions (Figure 5).^{86, 89} Conversely, smaller electrostatic potentials for CMTr1 ($-197 \text{ kJ}\cdot\text{mol}^{-1}\cdot\text{e}^{-1}$) and HcgC ($-209 \text{ kJ}\cdot\text{mol}^{-1}\cdot\text{e}^{-1}$) are associated with the more nucleophilic substrates (Figure 5).⁸⁷⁻⁸⁸ Quantitatively, a good correlation is observed over all enzymes ($R^2 = 0.8$) suggesting that an increase in the pK_a of the conjugate acid by 5 units should correlate to enzyme electrostatic potential reduction by ca. 40 $\text{kJ}\cdot\text{mol}^{-1}\cdot\text{e}^{-1}$ (Supporting Information Figure S11). Qualitatively, these results suggest that the

enzymatic electrostatic environment compensates for otherwise chemically challenging methyl transfer reactions.

In conclusion, we carried out a systematic study of all Class I MTases with which reasonable resolution (<2.0 Å) crystal structures could be used to form catalytically competent ternary complexes for simulation of the MTase reaction coordinate. We employed large-scale QM/MM analysis to investigate universal features of the MTase enzyme environment across four curated enzymes with distinct function and substrates. We observed that $\text{CH}\cdots\text{X}$ HBs formed between the SAM methyl group and enzyme stabilized the reactant complex SAM binding, but $\text{CH}\cdots\text{X}$ HBs contributed to TS stabilization only for N-type PfPMT and C-type HcgC. A consistent picture was instead obtained through analysis of charge transfer and electrostatics, wherein the majority of SAM-substrate charge separation from the reactant complex was maintained in the TS region, and strongly favorable electrostatic potentials correlated with substrate nucleophilicity (i.e., intrinsic reactivity). Despite the lack of apparent conserved role of methyl-group HBs, active site HBs to the substrates and other polar interactions play a significant role in charge transfer and electrostatics that contribute universally to MTase activity. These observations provide a potential path for re-engineering MTases to modify substrate scope or toward using electrostatic- and charge-transfer-derived properties of binary MTase structures to decrypt MTases with unknown function/substrates.

Computational Details.

Apoprotein protonation states were assigned using the H++ webserver⁹⁰⁻⁹² assuming a pH of 7.0 and a dielectric constant of 10.0 with all other defaults applied, and remaining holoprotein active site charges were manually assigned (Supporting Information Tables S7-S10). MD simulations with AMBER⁹³ used the AMBER ff14SB force field⁹⁴ along with the generalized

AMBER force field (GAFF)⁹⁵ for non-standard residues. GAFF RESP⁹⁶ charges were obtained with the R.E.D. web server⁹⁷⁻⁹⁹ (i.e., with GAMESS-US¹⁰⁰ using Hartree-Fock theory/6-31G*¹⁰¹). Protein ternary complexes were solvated in a periodic rectangular box with $> 10 \text{ \AA}$ buffer of TIP3P¹⁰² water and neutralized with Na⁺ or Cl⁻ counterions. The enzyme was equilibrated using the PMEMD GPU-accelerated form of AMBER¹⁰³⁻¹⁰⁴ and a $200 \text{ kcal/mol}\cdot\text{\AA}^2$ force constant on the protein to maintain the crystal structure atomic positions (details provided in Supporting Information Text S1 and Table S11). QM/MM was carried out using TeraChem^{63, 105} for the QM portion and AMBER⁹³ for the MM component in the Terachem-AMBER interface⁶¹. These aperiodic QM/MM simulations used spherical boundary conditions with a restraint of $1.5 \text{ kcal/mol}\cdot\text{\AA}^2$ on extracted spheres that were optimized for 2000 cycles (Supporting Information Table S11). The QM region was modeled with density functional theory (DFT) using the ω PBEh¹⁰⁶ ($\omega=0.2 \text{ bohr}^{-1}$) functional and the 6-31G*¹⁰¹ basis. QM regions were obtained from CSA^{32, 34} with a charge shift threshold of 0.04 a.u. determined from by-residue-summed Voronoi deformation density (VDD) charges¹⁰⁷ as in previous work^{6, 32, 34} (Supporting Information Tables S3-S6).

SMD simulations employed a $1000 \text{ kcal/mol}\cdot\text{\AA}^2$ harmonic pulling force selected by trial and error to drive the reaction from reactant state to product state over 1.5 ps with a step size of 0.5 fs (Supporting Information Figure S1). The TS region is selected based on where the vDW radii of the S atom (1.81 $\text{\AA}$) and the methyl-accepting O (1.58 $\text{\AA}$), N (1.64 $\text{\AA}$), or C (1.77 $\text{\AA}$) atoms begin to overlap. This corresponds to a TS between ca. 0.04 $\text{\AA}$ (HcgC) and ca. 0.2 $\text{\AA}$ (CMTr1, PIMT, or PfPMT) on the RC, consistent with the TS at around 0.3 $\text{\AA}$ obtained from the maximum in COMT free energy simulations³³ (Supporting Information Figures S2-S5). Analyzed snapshots were extracted from SMD every 2-4 steps to ensure an even spacing of RC

across enzymes (Supporting Information Table S12). BCP⁸¹ locations and energies⁸⁰ were obtained with Multiwfn.¹⁰⁸ Charge transfer was estimated from sums of VDD charges over the relevant reacting molecules. The electrostatic potential of the transferring methyl carbon was computed by Coulomb's law with RESP charges⁹⁶ computed¹⁰⁹ in TeraChem for the QM atoms and MM charges from the force field parameters.

ASSOCIATED CONTENT

Supporting Information. MTase PDB curation results; SMD force tests; QM region definitions for all enzymes; reaction coordinate donor and acceptor distances; C-H...X geometric and energetic hydrogen bond analysis; pK_a and electrostatic potential comparison; protonation state of residues for MM simulation; extended computational details including details of solvent cap and snapshot collection. (PDF)

PDB structures and topology/restart files for each reactant and transition state structure of each enzyme. (ZIP)

This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*email: hjkulik@mit.edu phone: 617-253-4584

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

The authors acknowledge an NEC Corporation Grant from the MIT Research Support Committee. H.J.K. holds a Career Award at the Scientific Interface from the Burroughs Wellcome Fund and an AAAS Marion Milligan Mason Award, which also supported this work. This work was carried out in part using computational resources from the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1548562. This work used the XStream computational resource, supported by the National Science Foundation Major Research Instrumentation program (ACI-1429830).

References

- (1) Schubert, H. L.; Blumenthal, R. M.; Cheng, X. Many Paths to Methyltransfer: A Chronicle of Convergence. *Trends in biochemical sciences* **2003**, *28*, 329-335.
- (2) Razin, A.; Riggs, A. D. DNA Methylation and Gene Function. *Science* **1980**, *210*, 604-610.
- (3) Skinner, M. M.; Puvathingal, J. M.; Walter, R. L.; Friedman, A. M. Crystal Structure of Protein Isoaspartyl Methyltransferase: A Catalyst for Protein Repair. *Structure* **2000**, *8*, 1189-1201.
- (4) Visick, J. E.; Cai, H.; Clarke, S. The L-Isoaspartyl Protein Repair Methyltransferase Enhances Survival of Aging Escherichia Coli Subjected to Secondary Environmental Stresses. *Journal of bacteriology* **1998**, *180*, 2623-2629.
- (5) Axelrod, J.; Tomchick, R. Enzymatic O-Methylation of Epinephrine and Other Catechols. *Journal of Biological Chemistry* **1958**, *233*, 702-705.
- (6) 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*, 11381-11394.
- (7) 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*, 7954-7959.
- (8) 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*, e0161868.
- (9) Świderek, K.; Tuñón, I.; Williams, I. H.; Moliner, V. Insights on the Origin of Catalysis on Glycine N-Methyltransferase from Computational Modeling. *J. Am. Chem. Soc.* **2018**, *140*, 4327-4334.
- (10) Lautala, P.; Ulmanen, I.; Taskinen, J. Molecular Mechanisms Controlling the Rate and Specificity of Catechol O-Methylation by Human Soluble Catechol O-Methyltransferase. *Molecular Pharmacology* **2001**, *59*, 393-402.

- (11) Coward, J. K.; Slisz, E. P.; Wu, F. Y. H. Kinetic Studies on Catechol O-Methyltransferase, Product Inhibition and the Nature of the Catechol Binding Site. *Biochemistry* **1973**, *12*, 2291-2297.
- (12) Lotta, T.; Vidgren, J.; Tilgmann, C.; Ulmanen, I.; Melen, K.; Julkunen, I.; Taskinen, J. Kinetics of Human Soluble and Membrane-Bound Catechol O-Methyltransferase: A Revised Mechanism and Description of the Thermolabile Variant of the Enzyme. *Biochemistry* **1995**, *34*, 4202-4210.
- (13) Woodard, R. W.; Tsai, M. D.; Floss, H. G.; Crooks, P. A.; Coward, J. K. Stereochemical Course of the Transmethylation Catalyzed by Catechol O-Methyltransferase. *Journal of Biological Chemistry* **1980**, *255*, 9124-9127.
- (14) 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: Struct., Funct., Bioinf.* **2015**, *83*, 318-330.
- (15) Lau, E. Y.; Bruice, T. C. Importance of Correlated Motions in Forming Highly Reactive near Attack Conformations in Catechol O-Methyltransferase. *Journal of the American Chemical Society* **1998**, *120*, 12387-12394.
- (16) Lau, E. Y.; Bruice, T. C. Comparison of the Dynamics for Ground-State and Transition-State Structures in the Active Site of Catechol O-Methyltransferase. *Journal of the American Chemical Society* **2000**, *122*, 7165-7171.
- (17) Roca, M.; Moliner, V.; Tuñón, I.; Hynes, J. T. Coupling between Protein and Reaction Dynamics in Enzymatic Processes: Application of Grote-Hynes Theory to Catechol O-Methyltransferase. *Journal of the American Chemical Society* **2006**, *128*, 6186-6193.
- (18) 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*, 1240-1251.
- (19) Jurkowska, R. Z.; Jurkowski, T. P.; Jeltsch, A. Structure and Function of Mammalian DNA Methyltransferases. *Chembiochem* **2011**, *12*, 206-222.
- (20) Czarnota, S.; Johannissen, L. O.; Baxter, N. J.; Rummel, F.; Wilson, A. L.; Cliff, M. J.; Levy, C. W.; Scrutton, N. S.; Waltho, J. P.; Hay, S. Equatorial Active Site Compaction and Electrostatic Reorganization in Catechol-O-Methyltransferase. *ACS Catalysis* **2019**, *9*, 4394-4401.
- (21) Zhang, J.; Klinman, J. P. Enzymatic Methyl Transfer: Role of an Active Site Residue in Generating Active Site Compaction That Correlates with Catalytic Efficiency. *Journal of the American Chemical Society* **2011**, *133*, 17134-17137.
- (22) 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*, 19280-19287.
- (23) Horowitz, S.; Dirk, L. M. A.; 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. *Journal of the American Chemical Society* **2013**, *135*, 15536-15548.
- (24) Horowitz, S.; Yesselman, J. D.; Al-Hashimi, H. M.; Trievel, R. C. Direct Evidence for Methyl Group Coordination by Carbon-Oxygen Hydrogen Bonds in the Lysine Methyltransferase Set7/9. *Journal of Biological Chemistry* **2011**, *286*, 18658-18663.
- (25) 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*, 1692-1697.
- (26) Fick, R. J.; Kroner, G. M.; Nepal, B.; Magnani, R.; Horowitz, S.; Houtz, R. L.; Scheiner, S.; Trievel, R. C. Sulfur-Oxygen Chalcogen Bonding Mediates AdoMet Recognition in the Lysine Methyltransferase Set7/9. *ACS chemical biology* **2016**, *11*, 748-754.
- (27) Fick, R. J.; Clay, M. C.; Vander Lee, L.; Scheiner, S.; Al-Hashimi, H.; Trievel, R. C. Water-Mediated Carbon-Oxygen Hydrogen Bonding Facilitates S-Adenosylmethionine

Recognition in the Reactivation Domain of Cobalamin-Dependent Methionine Synthase. *Biochemistry* **2018**, *57*, 3733-3740.

(28) Roca, M.; Martí, S.; Andrés, J.; Moliner, V.; Tuñón, I.; Bertrán, J.; Williams, I. H. Theoretical Modeling of Enzyme Catalytic Power: Analysis of “Cratic” and Electrostatic Factors in Catechol O-Methyltransferase. *Journal of the American Chemical Society* **2003**, *125*, 7726-7737.

(29) 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*, 8634-8635.

(30) Wilson, P. B.; Williams, I. H. Influence of Equatorial CH $\cdots$ O Interactions on Secondary Kinetic Isotope Effects for Methyl Transfer. *Angewandte Chemie* **2016**, *128*, 3244-3247.

(31) Saez, D. A.; Zinovjev, K.; Tuñón, I.; Vöhringer-Martinez, E. Catalytic Reaction Mechanism in Native and Mutant Catechol-O-Methyltransferase from the Adaptive String Method and Mean Reaction Force Analysis. *The Journal of Physical Chemistry B* **2018**, *122*, 8861-8871.

(32) Karelina, M.; Kulik, H. J. Systematic Quantum Mechanical Region Determination in QM/MM Simulation. *Journal of Chemical Theory and Computation* **2017**, *13*, 563-576.

(33) Kulik, H. J. Large-Scale QM/MM Free Energy Simulations of Enzyme Catalysis Reveal the Influence of Charge Transfer. *Physical Chemistry Chemical Physics* **2018**, *20*, 20650-20660.

(34) Qi, H. W.; Karelina, M.; Kulik, H. J. Quantifying Electronic Effects in QM and QM/MM Biomolecular Modeling with the Fukui Function. *Acta Physico-Chimica Sinica* **2018**, *34*, 81-91.

(35) Yang, Z.; Mehmood, R.; Wang, M.; Qi, H. W.; Steeves, A. H.; Kulik, H. J. Revealing Quantum Mechanical Effects in Enzyme Catalysis with Large-Scale Electronic Structure Simulation. *Reaction Chemistry & Engineering* **2019**.

(36) Welborn, V. V.; Pestana, L. R.; Head-Gordon, T. Computational Optimization of Electric Fields for Better Catalysis Design. *Nature Catalysis* **2018**, *1*.

(37) Vaissier, V.; Sharma, S. C.; Schaettle, K.; Zhang, T.; Head-Gordon, T. Computational Optimization of Electric Fields for Improving Catalysis of a Designed Kemp Eliminase. *ACS Catalysis* **2017**, *8*, 219-227.

(38) Bhowmick, A.; Sharma, S. C.; Head-Gordon, T. The Importance of the Scaffold for De Novo Enzymes: A Case Study with Kemp Eliminase. *Journal of the American Chemical Society* **2017**, *139*, 5793-5800.

(39) 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*, 14917-14925.

(40) García-Meseguer, R.; Zinovjev, K.; Roca, M.; Ruiz-Pernía, J. J.; Tuñón, I. a. Linking Electrostatic Effects and Protein Motions in Enzymatic Catalysis. A Theoretical Analysis of Catechol O-Methyltransferase. *The Journal of Physical Chemistry B* **2014**, *119*, 873-882.

(41) 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*, 9913-9921.

(42) Suydam, I. T.; Snow, C. D.; Pande, V. S.; Boxer, S. G. Electric Fields at the Active Site of an Enzyme: Direct Comparison of Experiment with Theory. *Science* **2006**, *313*, 200-204.

(43) Lockhart, D. J.; Boxer, S. G. Electric Field Modulation of the Fluorescence from Rhodobacter Sphaeroides Reaction Centers. *Chemical Physics Letters* **1988**, *144*, 243-250.

(44) Fafarman, A. T.; Sigala, P. A.; Schwans, J. P.; Fenn, T. D.; Herschlag, D.; Boxer, S. G. Quantitative, Directional Measurement of Electric Field Heterogeneity in the Active Site of Ketosteroid Isomerase. *Proceedings of the National Academy of Sciences* **2012**, *109*, E299-E308.

- (45) Schultz, E.; Nissinen, E. Inhibition of Rat Liver and Duodenum Soluble Catechol-O-Methyltransferase by a Tight-Binding Inhibitor OR-462. *Biochemical Pharmacology* **1989**, *38*, 3953-3956.
- (46) Kuhn, B.; Kollman, P. A. QM- Fe and Molecular Dynamics Calculations on Catechol O-Methyltransferase: Free Energy of Activation in the Enzyme and in Aqueous Solution and Regioselectivity of the Enzyme-Catalyzed Reaction. *Journal of the American Chemical Society* **2000**, *122*, 2586-2596.
- (47) Kulik, H. J.; Luehr, N.; Ufimtsev, I. S.; Martinez, T. J. Ab Initio Quantum Chemistry for Protein Structures. *The Journal of Physical Chemistry B* **2012**, *116*, 12501-12509.
- (48) Paton, R. S.; Goodman, J. M. Hydrogen Bonding and π -Stacking: How Reliable Are Force Fields? A Critical Evaluation of Force Field Descriptions of Nonbonded Interactions. *Journal of Chemical Information and Modeling* **2009**, *49*, 944-955.
- (49) Mao, Y.; Shao, Y.; Dziedzic, J.; Skylaris, C.-K.; Head-Gordon, T.; Head-Gordon, M. Performance of the AMOEBA Water Model in the Vicinity of QM Solutes: A Diagnosis Using Energy Decomposition Analysis. *Journal of chemical theory and computation* **2017**, *13*, 1963-1979.
- (50) Rod, T. H.; Ryde, U. Quantum Mechanical Free Energy Barrier for an Enzymatic Reaction. *Physical Review Letters* **2005**, *94*, 138302.
- (51) Kuhn, B.; Kollman, P. A. QM-Fe and Molecular Dynamics Calculations on Catechol O-Methyltransferase: Free Energy of Activation in the Enzyme and in Aqueous Solution and Regioselectivity of the Enzyme-Catalyzed Reaction. *Journal of the American Chemical Society* **2000**, *122*, 2586-2596.
- (52) Bruender, N. A.; Thoden, J. B.; Kaur, M.; Avey, M. K.; Holden, H. M. Molecular Architecture of a C-3' -Methyltransferase Involved in the Biosynthesis of d-Tetronitrose. *Biochemistry* **2010**, *49*, 5891-5898.
- (53) Smietanski, M.; Werner, M.; Purta, E.; Kaminska, K. H.; Stepinski, J.; Darzynkiewicz, E.; Nowotny, M.; Bujnicki, J. M. Structural Analysis of Human 2' -O-Ribose Methyltransferases Involved in mRNA Cap Structure Formation. *Nature communications* **2014**, *5*, 3004.
- (54) Griffith, S. C.; Sawaya, M. R.; Boutz, D. R.; Thapar, N.; Katz, J. E.; Clarke, S.; Yeates, T. O. Crystal Structure of a Protein Repair Methyltransferase from *Pyrococcus Furiosus* with Its L-Isoaspartyl Peptide Substrate1. *Journal of molecular biology* **2001**, *313*, 1103-1116.
- (55) Lee, S. G.; Kim, Y.; Alpert, T. D.; Nagata, A.; Jez, J. M. Structure and Reaction Mechanism of Phosphoethanolamine Methyltransferase from the Malaria Parasite *Plasmodium Falciparum* an Antiparasitic Drug Target. *Journal of Biological Chemistry* **2012**, *287*, 1426-1434.
- (56) Bai, L.; Wagner, T.; Xu, T.; Hu, X.; Ermler, U.; Shima, S. A Water-Bridged H-Bonding Network Contributes to the Catalysis of the SAM-Dependent C-Methyltransferase HcgC. *Angewandte Chemie* **2017**, *129*, 10946-10949.
- (57) Wang, S. P.; Deng, L.; Ho, C. K.; Shuman, S. Phylogeny of mRNA Capping Enzymes. *Proceedings of the National Academy of Sciences* **1997**, *94*, 9573-9578.
- (58) Shuman, S. Structure, Mechanism, and Evolution of the mRNA Capping Apparatus **2000**.
- (59) Furuichi, Y.; Shatkin, A. J. Viral and Cellular mRNA Capping: Past and Prospects **2000**.
- (60) Vidgren, J.; Svensson, L. A.; Liljas, A. Crystal Structure of Catechol O-Methyltransferase. *Nature* **1994**, *368*, 354-358.
- (61) Isborn, C. M.; Götz, A. W.; Clark, M. A.; Walker, R. C.; Martínez, T. J. Electronic Absorption Spectra from MM and Ab Initio QM/MM Molecular Dynamics: Environmental Effects on the Absorption Spectrum of Photoactive Yellow Protein. *Journal of Chemical Theory and Computation* **2012**, *8*, 5092-5106.
- (62) Liu, F.; Luehr, N.; Kulik, H. J.; Martínez, T. J. Quantum Chemistry for Solvated Molecules on Graphical Processing Units Using Polarizable Continuum Models. *Journal of chemical theory and computation* **2015**, *11*, 3131-3144.

- (63) Ufimtsev, I. S.; Martinez, T. J. Quantum Chemistry on Graphical Processing Units. 3. Analytical Energy Gradients, Geometry Optimization, and First Principles Molecular Dynamics. *Journal of Chemical Theory and Computation* **2009**, *5*, 2619-2628.
- (64) Jeziorski, B.; Moszynski, R.; Szalewicz, K. Perturbation Theory Approach to Intermolecular Potential Energy Surfaces of Van Der Waals Complexes. *Chemical Reviews* **1994**, *94*, 1887-1930.
- (65) Lin, H.; Truhlar, D. QM/MM: What Have We Learned, Where Are We, and Where Do We Go from Here? *Theor Chem Acc* **2007**, *117*, 185-199.
- (66) Monari, A.; Rivail, J.-L.; Assfeld, X. Advances in the Local Self-Consistent Field Method for Mixed Quantum Mechanics/Molecular Mechanics Calculations. *Accounts of Chemical Research* **2012**, *46*, 596-603.
- (67) Murphy, R. B.; Philipp, D. M.; Friesner, R. A. A Mixed Quantum Mechanics/Molecular Mechanics (QM/MM) Method for Large Scale Modeling of Chemistry in Protein Environments. *Journal of Computational Chemistry* **2000**, *21*, 1442-1457.
- (68) Rosta, E.; Klahn, M.; Warshel, A. Towards Accurate Ab Initio QM/MM Calculations of Free-Energy Profiles of Enzymatic Reactions. *The Journal of Physical Chemistry B* **2006**, *110*, 2934-2941.
- (69) Warshel, A.; Levitt, M. Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme. *Journal of molecular biology* **1976**, *103*, 227-249.
- (70) Park, S.; Schulten, K. Calculating Potentials of Mean Force from Steered Molecular Dynamics Simulations. *The Journal of chemical physics* **2004**, *120*, 5946-5961.
- (71) Park, S.; Khalili-Araghi, F.; Tajkhorshid, E.; Schulten, K. Free Energy Calculation from Steered Molecular Dynamics Simulations Using Jarzynski's Equality. *The Journal of chemical physics* **2003**, *119*, 3559-3566.
- (72) Izrailev, S.; Stepaniants, S.; Isralewitz, B.; Kosztin, D.; Lu, H.; Molnar, F.; Wriggers, W.; Schulten, K. Steered Molecular Dynamics. In *Computational Molecular Dynamics: Challenges, Methods, Ideas*; Springer: 1999; pp 39-65.
- (73) Isralewitz, B.; Gao, M.; Schulten, K. Steered Molecular Dynamics and Mechanical Functions of Proteins. *Current opinion in structural biology* **2001**, *11*, 224-230.
- (74) Kromann, J. C.; Christensen, A. S.; Cui, Q.; Jensen, J. H. Towards a Barrier Height Benchmark Set for Biologically Relevant Systems. *PeerJ* **2016**, *4*, e1994.
- (75) Ryde, U. How Many Conformations Need to Be Sampled to Obtain Converged QM/MM Energies? The Curse of Exponential Averaging. *Journal of chemical theory and computation* **2017**, *13*, 5745-5752.
- (76) Li, Y.; Zhang, R.; Du, L.; Zhang, Q.; Wang, W. How Many Conformations of Enzymes Should Be Sampled for DFT/MM Calculations? A Case Study of Fluoroacetate Dehalogenase. *International Journal of Molecular Sciences* **2016**, *17*, 1372.
- (77) Yang, Z.; Yang, S.; Yu, P.; Li, Y.; Doubleday, C.; Park, J.; Patel, A.; Jeon, B.-s.; Russell, W. K.; Liu, H.-w. Influence of Water and Enzyme Spnf on the Dynamics and Energetics of the Ambimodal [6+ 4]/[4+ 2] Cycloaddition. *Proceedings of the National Academy of Sciences* **2018**, *115*, E848-E855.
- (78) Rowland, R. S.; Taylor, R. Intermolecular Nonbonded Contact Distances in Organic Crystal Structures: Comparison with Distances Expected from Van Der Waals Radii. *The Journal of Physical Chemistry* **1996**, *100*, 7384-7391.
- (79) Batsanov, S. S. Van Der Waals Radii of Elements. *Inorganic materials* **2001**, *37*, 871-885.
- (80) Espinosa, E.; Molins, E.; Lecomte, C. Hydrogen Bond Strengths Revealed by Topological Analyses of Experimentally Observed Electron Densities. *Chemical Physics Letters* **1998**, *285*, 170-173.
- (81) Bader, R. F. A Quantum Theory of Molecular Structure and Its Applications. *Chemical Reviews* **1991**, *91*, 893-928.

- (82) Qi, H. W.; Kulik, H. J. Evaluating Unexpectedly Short Non-Covalent Distances in X-Ray Crystal Structures of Proteins with Electronic Structure Analysis. *Journal of Chemical Information and Modeling* **2019**.
- (83) Mehmood, R.; Qi, H. W.; Steeves, A. H.; Kulik, H. J. The Protein's Role in Substrate Positioning and Reactivity for Biosynthetic Enzyme Complexes: The Case of SyrB2/SyrB1. *ACS Catalysis* **2019**.
- (84) Kresheck, G. C.; Klotz, I. M. Thermodynamics of Transfer of Amides from an Apolar to an Aqueous Solution. *Biochemistry* **1969**, *8*, 8-12.
- (85) Chothia, C.; Janin, J. Principles of Protein-Protein Recognition. *Nature* **1975**, *256*, 705.
- (86) Ellenbogen, E. Dissociation Constants of Peptides. I. A Survey of the Effect of Optical Configuration1. *Journal of the American Chemical Society* **1952**, *74*, 5198-5201.
- (87) Velikyan, I.; Acharya, S.; Trifonova, A.; Földesi, A.; Chattopadhyaya, J. The pKa's of 2'-Hydroxyl Group in Nucleosides and Nucleotides. *Journal of the American Chemical Society* **2001**, *123*, 2893-2894.
- (88) Chiang, Y.; Kresge, A. J.; Tang, Y. S.; Wirz, J. The pKa and Keto-Enol Equilibrium Constant of Acetone in Aqueous Solution. *Journal of the American Chemical Society* **1984**, *106*, 460-462.
- (89) Dippy, J. F. J.; Hughes, S. R. C.; Rozanski, A. 498. The Dissociation Constants of Some Symmetrically Disubstituted Succinic Acids. *Journal of the Chemical Society (Resumed)* **1959**, 2492-2498.
- (90) Anandakrishnan, R.; Aguilar, B.; Onufriev, A. V. H++ 3.0: Automating pK Prediction and the Preparation of Biomolecular Structures for Atomistic Molecular Modeling and Simulations. *Nucleic Acids Research* **2012**, *40*, W537-W541.
- (91) Gordon, J. C.; Myers, J. B.; Foltz, T.; Shoja, V.; Heath, L. S.; Onufriev, A. H++: A Server for Estimating pKas and Adding Missing Hydrogens to Macromolecules. *Nucleic Acids Research* **2005**, *33*, W368-W371.
- (92) Myers, J.; Grothaus, G.; Narayanan, S.; Onufriev, A. A Simple Clustering Algorithm Can Be Accurate Enough for Use in Calculations of pKs in Macromolecules. *Proteins: Structure, Function, and Bioinformatics* **2006**, *63*, 928-938.
- (93) D.A. Case, J. T. B., R.M. Betz, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, T. Luchko, R. Luo, B. Madej, K.M. Merz, G. Monard, P. Needham, H. Nguyen, H.T. Nguyen, I. Omelyan, A. Onufriev, D.R. Roe, A. Roitberg, R. Salomon-Ferrer, C.L. Simmerling, W. Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, D.M. York and P.A. Kollman Amber 2015, University of California, San Francisco. 2015.
- (94) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *Journal of Chemical Theory and Computation* **2015**, *11*, 3696-3713.
- (95) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *Journal of Computational Chemistry* **2004**, *25*, 1157-1174.
- (96) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *The Journal of Physical Chemistry* **1993**, *97*, 10269-10280.
- (97) F. Wang, J.-P. B., P. Cieplak, and F.-Y. Dupradeau. R.E.D. Python: Object Oriented Programming for Amber Force Fields, Université De Picardie - Jules Verne, Sanford|Burnham Medical Research Institute, Nov. 2013. <http://q4md-forcefieldtools.org/REDServer-Development/> (accessed 5/27/19).
- (98) Vanqualef, E.; Simon, S.; Marquant, G.; Garcia, E.; Klimerak, G.; Delepine, J. C.; Cieplak, P.; Dupradeau, F.-Y. R.E.D. Server: A Web Service for Deriving RESP and ESP Charges and Building Force Field Libraries for New Molecules and Molecular Fragments. *Nucleic Acids Research* **2011**, *39*, W511-W517.

- (99) Dupradeau, F.-Y.; Pigache, A.; Zaffran, T.; Savineau, C.; Lelong, R.; Grivel, N.; Lelong, D.; Rosanski, W.; Cieplak, P. The R.E.D. Tools: Advances in RESP and ESP Charge Derivation and Force Field Library Building. *Physical Chemistry Chemical Physics* **2010**, *12*, 7821-7839.
- (100) Gordon, M. S.; Schmidt, M. W. Advances in Electronic Structure Theory: GAMESS a Decade Later. *Theory and Applications of Computational Chemistry: the first forty years* **2005**, 1167-1189.
- (101) Harihara, P. C.; Pople, J. A. Influence of Polarization Functions on Molecular-Orbital Hydrogenation Energies. *Theor Chim Acta* **1973**, *28*, 213-222.
- (102) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *The Journal of Chemical Physics* **1983**, *79*, 926-935.
- (103) Götz, A. W.; Williamson, M. J.; Xu, D.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with Amber on GPUs. 1. Generalized Born. *Journal of Chemical Theory and Computation* **2012**, *8*, 1542-1555.
- (104) Salomon-Ferrer, R.; Götz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with Amber on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *Journal of Chemical Theory and Computation* **2013**, *9*, 3878-3888.
- (105) Petachem, Llc. <http://www.petachem.com> (accessed 5/27/19).
- (106) Rohrdanz, M. A.; Martins, K. M.; Herbert, J. M. A Long-Range-Corrected Density Functional That Performs Well for Both Ground-State Properties and Time-Dependent Density Functional Theory Excitation Energies, Including Charge-Transfer Excited States. *The Journal of chemical physics* **2009**, *130*, 054112.
- (107) Guerra, C. F.; Handgraaf, J. W.; Baerends, E. J.; Bickelhaupt, F. M. Voronoi Deformation Density (VDD) Charges: Assessment of the Mulliken, Bader, Hirshfeld, Weinhold, and VDD Methods for Charge Analysis. *Journal of Computational Chemistry* **2004**, *25*, 189-210.
- (108) Lu, T.; Chen, F. Multiwfn: A Multifunctional Wavefunction Analyzer. *Journal of Computational Chemistry* **2012**, *33*, 580-592.
- (109) Liu, F.; Sanchez, D. M.; Kulik, H. J.; Martinez, T. J. Exploiting Graphical Processing Units to Enable Quantum Chemistry Calculation of Large Solvated Molecules with Conductor-Like Polarizable Continuum Models. *International Journal of Quantum Chemistry* **2018**, *119*, e25760.

MTCTDraft_v8.pdf (2.70 MiB)

[view on ChemRxiv](#) • [download file](#)

What governs catalysis?

Supporting Information for

A Quantum Mechanical Description of Electrostatics Provides a Unified Picture of Catalytic Action Across Methyltransferases

Zhongyue Yang¹, Fang Liu¹, Adam H. Steeves¹, and Heather J. Kulik¹

¹Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139

Contents

Table S1 Results of MTase PDB curation: Type I MTases	Page S2
Table S2 Results of MTase PDB curation: Other MTases and unpublished structures	Page S5
Figure S1 Restraining force test in CMTr1	Page S6
Table S3 Selection of residues for the QM region in CMTr1	Page S7
Table S4 Selection of residues for the QM region in PIMT	Page S8
Table S5 Selection of residues for the QM region in PfPMT	Page S11
Table S6 Selection of residues for the QM region in HcgC	Page S12
Figure S2 Individual donor/acceptor distances for reaction coordinate in CMTr1	Page S13
Figure S3 Individual donor/acceptor distances for reaction coordinate in PIMT	Page S13
Figure S4 Individual donor/acceptor distances for reaction coordinate in PfPMT	Page S14
Figure S5 Individual donor/acceptor distances for reaction coordinate in HcgC	Page S14
Figure S6 Other CH $\cdots$ X hydrogen bonds in MTases	Page S15
Figure S7 CH $\cdots$ X hydrogen bond energies in CMTr1	Page S16
Figure S8 CH $\cdots$ X hydrogen bond energies in PIMT	Page S16
Figure S9 CH $\cdots$ X hydrogen bond energies in PfPMT	Page S17
Figure S10 CH $\cdots$ X hydrogen bond energies in HcgC	Page S17
Figure S11 Correlation of pK _a and electrostatic potential across MTases	Page S18
Table S7 Protonation state of residues for MM simulation in CMTr1	Page S19
Table S8 Protonation state of residues for MM simulation in PIMT	Page S19
Table S9 Protonation state of residues for MM simulation in PfPMT	Page S20
Table S10 Protonation state of residues for MM simulation in HcgC	Page S20
Text S1 Extended computational details	Page S21
Table S11 Details of solvent cap in MTases	Page S22
Table S12 Details of snapshots in MTases	Page S22
References	Page S23

Table S1. Summary of crystal structure analysis for 40 unique Type I MTase crystal structures with associated peer reviewed publications curated from the PDB in January 2017. The parent PDB code is listed in the first column and “other structures” refers to other PDB codes affiliated with that same paper. The closest C···O non-covalent distances (in Å) of the binary complex are reported under 3.5 Å, “Prev.” refers to whether the crystal structures were analyzed in Ref. 1 (Y = Yes, N = No), function is the putative function of the relevant enzyme, substrate refers to the class of substrates the enzyme methylates, and “Ternary” refers to whether a catalytically competent ternary complex could be formed. Yes** means that the protein was excluded due to being COMT, Yes* means the protein was excluded due to unphysical rearrangement of the protein (i.e., proton rearrangement) during the reaction, No* means that a ternary was formed but not with the products instead with substrates or a reaction coordinate was unclear, and No** means the substrate was just outside of the 4.0 Å distance cutoff or otherwise ambiguous. Of the remaining 6 candidate proteins, 4 were selected for subsequent study, as indicated in bold. The Ref. column lists the original reference.

PDB	Res.	Year	dC···O	Prev.	Function	Substrate	Other structures	Ternary	Ref.
5FHQ	1.63	2016	3.30	N	COMT (rat) catechol-O-methyltransferase	Catechol	5FHR	Yes**	http://dx.doi.org/10.1002/anie.201508287
2CL5	1.60	2006	3.24, 3.37, 3.38, 3.67	Y	COMT w/ inhibitor	Catechol	3BWM	Yes**	https://doi.org/10.1124/mol.106.023119
3NDI	1.50	2010	3.56, 3.56, 3.66	Y	C-3'-methyltransferase for D-tetronitrose	Sugar	3NDJ	Yes*	http://dx.doi.org/10.1021/bi100782b
3UJ7	1.55	2012	2.85	Y	phosphoethanolamine methyltransferase	Phosphoethanolamine	3UJ6, 3UJ8, 3UJ9, 3UJA, 3UJB, 3UJC, 3UJD	Yes	http://dx.doi.org/10.1074/jbc.M111.315267
1JG4	1.50	2001	2.81	Y	L-Isoaspartyl O-Mtase	Protein	1JG1, 1JG2, 1JG3	Yes	http://dx.doi.org/10.1006/jmbi.2001.5095
5D4U	2	2016	3.9	N	HcgC: Pyridinol Methyltransferase, aromatic C-methyltransferase	Pyridinol	5D4T, 5D4V, 5D5O, 5D5T, 5O4M, 5O4N, 5O4H, 5O4J	Yes	http://dx.doi.org/10.1002/anie.201604352
4N49	1.9	2014	3.60	N	CMT1/CMT2: 2'-O-ribose methyltransferase (RNA cap)	RNA/DNA	4N48, 4N4A	Yes	http://dx.doi.org/10.1038/ncomms4004
*4BLV	2	2013	3.30	N	RlmJ: rRNA M6A Methyltransferase	RNA/DNA	4BLU, 4BLW	No**	http://dx.doi.org/10.1093/nar/gkt719
3ELW	1.90	2009	3.40	Y	Wesselbron virus Mtase	RNA/DNA	3ELD, 3ELU, 3ELW, 3ELY, 3EMB, 3EMD	No**	http://dx.doi.org/10.1016/j.jmb.2008.10.028
3PFG	1.35	2011	3.39, 3.49, 3.49	Y	TylM1 dimethylation of C3' amino of hexose	Sugar	3PFH, 3PX2, 3PX3	No*	http://dx.doi.org/10.1021/bi101733v

PDB	Res.	Year	dC ⁺⁺ O	Prev.	Function	Substrate	Other structures	Ternary	Ref.
5KVA	1.83	2016	3.40	N	Caffeoyl-CoA O-Methyltransferase	caffeoyl-coenzyme A		No	http://dx.doi.org/10.1104/pp.16.00845
4YMG	1.9	2015	3.2	N	PaMTH1: small molecule O-methyltransferase (myricetin)	myricetin	4QVK 4YMG 4YMH	No	http://dx.doi.org/10.1074/jbc.M115.660829
5FA8	1.3	2016	3.00	N	aKMT: Protein Lysine N-methyltransferase	Protein	5FAD	No	http://dx.doi.org/10.1074/mcp.M115.057778
1RJD	1.80	2004	3.49, 3.54	Y	Leucine carboxymethyltransferase	Protein	1RJE, 1RJF, 1RJG	No	http://dx.doi.org/10.1074/jbc.M311484200
4PCL	1.85	2015	3.1	N	O-Methyltransferase	Protein	4OA5 4OA8 4PCA	No	http://dx.doi.org/10.1371/journal.ppat.1005248
2NXE	1.75	2007	2.82	Y	Trimethyltransferase PrmA	Protein	2NXC, 2NXC, 2NXJ, 2NXN	No	http://dx.doi.org/10.1038/sj.emboj.7601508
4QDJ	1.6	2014	3.10	N	ChlM: protoporphyrin IX methyltransferase	Protoporphyrin sidechain	4QDK	No	http://dx.doi.org/10.1074/jbc.M114.584920
3MTE	1.80	2010	3.02	Y	A1408 16S rRNA Mtase	RNA/DNA	3MQ2	No	http://dx.doi.org/10.1093/nar/gkq627
5E72	1.74	2016	3.8	N	aTrm11: tRNA m2G/m22G10 methyltransferase	RNA/DNA	5E71	No	http://dx.doi.org/10.1093/nar/gkw561
4CTJ	1.47	2014	3.20, 3.25, 3.32, 3.39, 3.40, 3.45	N	DENGUE 3 NS5 METHYLTRANSFERASE (RNA cap Mtase)	RNA/DNA	4CTK	No	http://dx.doi.org/10.1016/j.antiviral.2014.03.013
5EKX	1.76	2017	3.30	N	DENGUE 3 NS5 METHYLTRANSFERASE RNA cap Guanine N7 and RNA 2-O Mtase	RNA/DNA	5E9Q, 5EC8, 5EHG, 5EHI, 5EIF, 5EIW	No	http://dx.doi.org/10.1016/j.eimech.2016.10.007
1G60	1.74	2003	3.00, 3.00	Y	DNA Mtase MBoIIA (beta-class)	RNA/DNA		No	http://dx.doi.org/10.1093/nar/gkg713
2DPM	1.80	1998	3.25	Y	DpnM DNA adenine Mtase	RNA/DNA		No	http://dx.doi.org/10.1016/S0969-2126(98)00154-3
4DF3	1.73	2012	3.80	N	Fibrillarin: snoRNA methyltransferase	RNA/DNA		No	http://dx.doi.org/10.1107/S1744309112026528
3GCZ	1.70	2009	3.17	Y	Flavivirus Mtase	RNA/DNA		No	http://dx.doi.org/10.1016/j.bbrc.2009.03.008
3DMF	1.58		3.21	Y	Guanosine rRNA Mtase	RNA/DNA	3DMG, 3DMH	No	http://dx.doi.org/10.1074/jbc.M804005200
5K7U	1.7	2016	3.4	N	METTL3: adenine m6 N-methyltransferase	RNA/DNA	5K7M, 5K7W	No	http://dx.doi.org/10.1016/j.molcel.2016.05.041
5IL1	1.71	2016	3.3	N	METTL3: adenine m6 N-methyltransferase	RNA/DNA	5IL0, 5IL2	No	http://dx.doi.org/10.1038/nature18298
5L6E	1.9	2016	3.4	N	METTL3: adenine m6 N-methyltransferase	RNA/DNA	5L6D	No	http://dx.doi.org/10.7554/eLife.18434
4FZV	2	2012	3.50	N	NSUN4: rRNA methyltransferase	RNA/DNA		No	http://dx.doi.org/10.1016/j.str.2012.08.027
4DMG	1.7	2012	3.70	N	RlmO: 5-Methylcytidine (m5C) Methyltransferase (RNA)	RNA/DNA		No	http://dx.doi.org/10.1074/jbc.M112.376160

PDB	Res.	Year	dC ⁺⁺ O	Prev.	Function	Substrate	Other structures	Ternary	Ref.
1EIZ	1.50	2000	3.38, 3.59	Y	RNA Methyltransferase FtsJ from heat shock proteins	RNA/DNA	1EJ0 (3.49, 3.56)	No	http://dx.doi.org/10.1016/S1097-2765(00)00035-6
3M6W	1.30	2010	3.32, 3.41	Y	RsmF 16S rRNA Mtase (methylcytidine)	RNA/DNA	3M6U, 3M6V, 3M6W, 3M6X	No	http://dx.doi.org/10.1261/rna.2088310
3G89	1.50	2009	3.12, 3.22, 3.25, 3.38, 3.47	Y	RsmG Mtase 16S rRNA - N7 methylation G527	RNA/DNA	3G88, 3G8A, 3G8B	No	http://dx.doi.org/10.2210/pdb3g89/pdb
3DXY	1.50	2009	2.92	Y	tRNA (m(7)G46) Mtase	RNA/DNA	3DXX, 3DXZ	No	http://dx.doi.org/10.1002/prot.22413
3MB5	1.60	2010	3.28, 3.40, 3.53	Y	tRNA m1A58 Mtase	RNA/DNA	3LGA, 3LHD	No	http://dx.doi.org/10.1093/nar/gkq381
3TM4	1.95		3.47	Y	tRNA:m2G6 methyltransferase Trm14/TrmN	RNA/DNA	3TLJ, 3TM5, 3TMA	No	http://dx.doi.org/10.1093/nar/gks163
3P97	1.70	2011	2.30, 2.94, 3.34, 3.40	Y	Dengue flavivirus Mtase	RNA/DNA	3P8Z	No	http://dx.doi.org/10.1074/jbc.M110.179184
1N2X	1.90	2003	2.43, 3.16	Y	TM0872 Mtase	Unk	1M6Y	No	http://dx.doi.org/10.1110/ps.0302403
2WA2	1.80	2009	3.46	Y	Virus Mtase	Unk	2WA1	No	https://doi.org/10.1107/S0907444909017260

Table S2. Type II-VI MTases from crystal structure analysis (total 18) along with 14 for which no published manuscript is available but were included in prior analysis in Ref. 1. The parent PDB code is listed in the first column and “other structures” refers to other PDB codes affiliated with that same paper. The closest C⁺O non-covalent distances (in Å) of the binary complex are reported under 3.5 Å, “Prev.” refers to whether the crystal structures were analyzed in Ref. 1 (Y = Yes, N = No), function is the putative function of the relevant enzyme, substrate refers to the class of substrates the enzyme methylates. The Ref. column lists the original reference. The unpublished structures are shaded in gray.

PDB	Res.	Year	dC ⁺ O	Prev.	Type	Function	Other structures	Ref.
1MSK	1.80	1996	3.02, 3.50	Y	II	Methionine synthase		http://dx.doi.org/10.1016/S0969-2126(96)00135-9
1V2X	1.85	2004	2.59, 2.96	Y	IV	tRNA(GM18) Mtase: acts on Guanosine18 of tRNA		http://dx.doi.org/10.1016/j.str.2004.03.003
4YVG	1.55	2015	3.32	N	IV	TrmD: tRNA (guanine-N(1))-methyltransferase	4YVH 4YVI 4YVJ 4YVK	http://dx.doi.org/10.1073/pnas.1422981112
3AIA	1.40	2010	3.06, 3.21	Y	IV	RNA Mtase	3AIG	http://dx.doi.org/10.1093/mcb/miq034
3RQ4	1.85	2013	3.22, 3.34, 3.34	Y	V	Human histone lysine Mtase	3S8P (3.10, 3.18, 3.40)	http://dx.doi.org/10.1016/j.febslet.2013.10.020
5ARF	1.92	2016	2.63	N	V	Smyd2: Histone lysine N-methyltransferase	5AR G	http://dx.doi.org/10.1021/acs.jmedchem.5b01890
5CCL	1.5	2016	2.89, 3.18, 3.20, 3.46	N	V	SMYD3: Histone-lysine N-methyltransferase	5CC M	http://dx.doi.org/10.1021/acsmedchemlett.5b00272
4J83	1.7	2013	3.01, 3.02, 3.44	N	V	SET7/9: histone lysine N-methyltransferase	4J7I, 4J8O	http://dx.doi.org/10.1021/ja407140k
4JLG	1.9	2014	3.04, 3.10, 3.34	N	V	SETD7: protein lysine N methyltransferase		http://dx.doi.org/10.1073/pnas.1407358111
5JIY	1.48	2016	3.1	N	V	G9a: Histone lysine N-methyltransferase	5JHN, 5JIN, 5JJ0	http://dx.doi.org/10.1073/pnas.1605523113
3OOI	1.75	2010	3.11, 3.49	Y	V	Human Histone-Lysine N-methyltransferase SET domain		http://dx.doi.org/10.1074/jbc.M110.204115
1N6A		2003	3.00, 3.48	Y	V	SET7/9-Adomet		http://dx.doi.org/10.1093/emboj/cdq025
5TOM	1.9	2016	3.2	N	V	Clr4: Histone lysine N-methyltransferase	5TOK	http://dx.doi.org/10.7554/eLife.17903
5FEP	1.45	2016	--	N	VI	HydE: radical SAM Mtase	5FES 5FE W 5FEX 5FEZ 5FF0 5FF2 5FF3 5FF4	http://dx.doi.org/10.1038/nchem.2490
4NJH	1.9	2014	3.32	N	VI	QueE: Radical SAM	4NJG 4NJI 4NJJ 4NJK	http://dx.doi.org/10.1038/nchembio.1426
4M7T	1.56	2013	3.48	N	VI	BtrN: radical SAM	4M7S	http://dx.doi.org/10.1073/pnas.1312228110

PDB	Res.	Year	dC...O	Prev.	Type	Function	Other structures	Ref.
4WCX	1.59	2015	--	N	VI	HydG: radical SAM		http://dx.doi.org/10.1073/pnas.1417252112
3RFA	2.05	2011	3.26	Y	VI	Radical SAM Mtase	3RF9	http://dx.doi.org/10.1126/science.1205358
1ZQ9			3.09	Y				Not published
2AVD			3.27, 3.51	Y				Not published
2B9E			3.44	Y				Not published
2NPN			3.08, 3.08	Y				Not published
2NYU			3.42	Y				Not published
2PLW			3.22	Y				Not published
2ZUL			3.15	Y				Not published
3CKK			2.90	Y				Not published
3DLC			3.29	Y				Not published
3DOU			3.61, 3.68	Y				Not published
3IHT			3.53	Y				Not published
3KKZ			2.83, 3.28	Y				Not published
3QWP			3.08, 3.25, 3.34	Y				Not published
4HTF			3.21	Y				Not published

Figure S1. The evolution of the reaction coordinate ($d_{C-S}-d_{C-X}$) under different restraining force constants in CMTr1. Values ranging from 250-1000 enforce the MT coordinate, whereas lower values do not.

Table S3. List of QM-treated residues considered for possible QM region inclusion in CSA analysis and those that were preserved after CSA are indicated with “Y” in “CSA?” column for CMTr1. A QM region of 484 atoms + 24 link atoms (508 total) is selected with a charge of -2 from an initial region of 934 atoms + 24 link atoms (958 total) and a charge of -2.

Res. #	Res.	# At	Charge	CSA?	Res. #	Res.	# At	Charge	CSA?
63	LYS	22	1	Y	197	THR	14	0	--
64	SER	11	0	--	224	ASP	12	-1	Y
66	PHE	20	0	Y	225	GLY	7	0	--
67	ASP	12	-1	--	226	GLY	7	0	--
74	MET	17	0	--	227	PHE	20	0	--
75	ARG	24	1	--	228	SER	11	0	--
78	ARG	24	1	Y	233	GLU	15	-1	Y
93	LEU	19	0	Y	234	ASN	14	0	--
94	ASN	14	0	Y	236	GLN	17	0	--
95	ARG	24	1	Y	243	LEU	19	0	Y
96	ALA	10	0	--	264	LYS	22	1	--
97	ALA	10	0	--	266	PHE	20	0	--
99	LYS	22	1	Y	296	ARG	24	1	--
137	CYS	11	0	--	298	ALA	10	0	--
138	ALA	10	0	--	299	ASN	14	0	Y
139	GLY	7	0	--	300	SER	11	0	Y
140	PRO	14	0	--	301	GLU	15	-1	--
141	GLY	7	0	Y	367	GLN	17	0	--
142	GLY	7	0	--	371	LEU	19	0	--
143	PHE	20	0	--	408	SAM	50	1	Y
160	MET	17	0	--	409	MGT	44	-1	Y
161	THR	14	0	--	410	G	34	-1	Y
162	LEU	19	0	Y	411	A	33	-1	Y
166	ASN	14	0	--	412	U	30	-1	Y
194	GLY	7	0	--	413	C3	32	-1	Y
195	ASP	12	-1	--				All	CSA
196	ILE	19	0	Y	Totals		# atoms	934	484
							with link	958	508
							charge	-2	-2

Table S4. List of QM-treated residues considered for possible QM region inclusion in CSA analysis and those that were preserved after CSA are indicated with “Y” in “CSA?” column for PIMT. Thr60 and Thr152 were manually incorporated due to their possible involvement in the formation of CH $\cdots$ O hydrogen bond with the SAM methyl group. This leads to a QM region with 452 atoms, 482 atoms including link atoms, and a net charge of -3 after CSA from an initial QM region of 1539 atoms, 1575 including link atoms, and a net charge of -3.

Res. #	Res.	# At	Charge	CSA?	Res. #	Res.	# At	Charge	CSA?
20	ILE	19	0	--	283	WAT	3	0	--
38	SER	11	0	--	287	WAT	3	0	--
43	TYR	21	0	Y	288	WAT	3	0	--
50	ASP	12	-1	Y	289	WAT	3	0	--
51	GLU	15	-1	Y	296	WAT	3	0	--
52	PRO	14	0	Y	297	WAT	3	0	--
53	LEU	19	0	--	300	WAT	3	0	--
54	PRO	14	0	Y	302	WAT	3	0	--
55	ILE	19	0	--	303	WAT	3	0	--
57	ALA	10	0	--	306	WAT	3	0	--
58	GLY	7	0	--	307	WAT	3	0	--
59	GLN	17	0	Y	309	WAT	3	0	--
60	THR	14	0	Y	315	WAT	3	0	--
61	VAL	16	0	--	321	WAT	3	0	--
62	SER	11	0	--	323	WAT	3	0	--
63	ALA	10	0	--	326	WAT	3	0	--
65	HIP	18	1	Y	330	WAT	3	0	--
66	MET	17	0	--	331	WAT	3	0	--
67	VAL	16	0	--	339	WAT	3	0	--
69	ILE	19	0	--	344	WAT	3	0	--
70	MET	17	0	--	349	WAT	3	0	--
84	GLU	15	-1	Y	351	WAT	3	0	--
85	VAL	16	0	--	363	WAT	3	0	--
86	GLY	7	0	--	365	WAT	3	0	--
87	THR	14	0	Y	374	WAT	3	0	--
88	GLY	7	0	--	376	WAT	3	0	--
89	SER	11	0	--	382	WAT	3	0	--
90	GLY	7	0	--	385	WAT	3	0	--
91	TRP	24	0	--	391	WAT	3	0	--
92	ASN	14	0	Y	404	WAT	3	0	--
93	ALA	10	0	--	407	WAT	3	0	--
94	ALA	10	0	--	412	WAT	3	0	--
95	LEU	19	0	--	417	WAT	3	0	--
96	ILE	19	0	--	427	WAT	3	0	--
106	THR	14	0	--	431	WAT	3	0	--
107	ILE	19	0	--	435	WAT	3	0	--
108	GLU	15	-1	Y	437	WAT	3	0	--
109	ARG	24	1	Y	439	WAT	3	0	--
110	ILE	19	0	--	440	WAT	3	0	--
113	LEU	19	0	--	446	WAT	3	0	--
132	ILE	19	0	--	447	WAT	3	0	--
134	GLY	7	0	--	458	WAT	3	0	--
135	ASP	12	-1	Y	466	WAT	3	0	--
136	GLY	7	0	--	467	WAT	3	0	--

Res. #	Res.	# At	Charge	CSA?	Res. #	Res.	# At	Charge	CSA?
137	SER	11	0	--	473	WAT	3	0	--
138	LYS	22	1	--	479	WAT	3	0	--
150	ILE	19	0	--	485	WAT	3	0	--
151	VAL	16	0	--	495	WAT	3	0	--
152	THR	14	0	Y	516	WAT	3	0	--
153	ALA	10	0	--	521	WAT	3	0	--
154	GLY	7	0	--	529	WAT	3	0	--
175	PRO	14	0	--	535	WAT	3	0	--
180	HID	17	0	Y	542	WAT	3	0	--
181	LEU	19	0	--	543	WAT	3	0	--
183	GLN	17	0	--	564	WAT	3	0	--
185	LEU	19	0	--	581	WAT	3	0	--
200	HIE	17	0	Y	585	WAT	3	0	--
201	GLY	7	0	--	601	WAT	3	0	--
202	GLY	7	0	Y	618	WAT	3	0	--
203	VAL	16	0	Y	630	WAT	3	0	--
204	ALA	10	0	--	635	WAT	3	0	--
205	PHE	20	0	--	639	WAT	3	0	--
206	VAL	16	0	--	645	WAT	3	0	--
207	PRO	14	0	--	667	WAT	3	0	--
208	LEU	19	0	--	688	WAT	3	0	--
209	ILE	19	0	Y	689	WAT	3	0	--
210	GLY	7	0	--	690	WAT	3	0	--
211	GLU	15	-1	--	692	WAT	3	0	--
212	TYR	21	0	--	696	WAT	3	0	--
213	GLY	7	0	--	706	WAT	3	0	--
214	TRP	24	0	--	718	WAT	3	0	--
216	SAM	50	1	Y	735	WAT	3	0	--
217	VAL	18	0	Y	755	WAT	3	0	--
218	TYR	21	0	Y	758	WAT	3	0	--
219	PRO	14	0	Y	784	WAT	3	0	--
220	IAS	12	-1	Y	791	WAT	3	0	--
221	HIE	17	0	Y	813	WAT	3	0	--
222	ALA	11	0	Y	836	WAT	3	0	--
231	WAT	3	0	--	841	WAT	3	0	--
232	WAT	3	0	--	864	WAT	3	0	--
234	WAT	3	0	--	937	WAT	3	0	--
235	WAT	3	0	--	959	WAT	3	0	--
240	WAT	3	0	--	977	WAT	3	0	--
241	WAT	3	0	--	985	WAT	3	0	--
244	WAT	3	0	--	1061	WAT	3	0	--
249	WAT	3	0	--	1095	WAT	3	0	--
250	WAT	3	0	--	1145	WAT	3	0	--
252	WAT	3	0	--	1171	WAT	3	0	--
256	WAT	3	0	--	1196	WAT	3	0	--
259	WAT	3	0	--	1202	WAT	3	0	--
261	WAT	3	0	--	1335	WAT	3	0	--
263	WAT	3	0	--	1402	WAT	3	0	--
264	WAT	3	0	--	1454	WAT	3	0	--
273	WAT	3	0	--	1486	WAT	3	0	--

Res. #	Res.	# At	Charge	CSA?	Res. #	Res.	# At	Charge	CSA?
277	WAT	3	0	--	1576	WAT	3	0	--
279	WAT	3	0	--				All	CSA
280	WAT	3	0	--	Totals		# atoms	1539	452
281	WAT	3	0	--			with link	1575	482
282	WAT	3	0	--			charge	-3	-3

Table S5. List of QM-treated residues considered for possible QM region inclusion in CSA analysis and those that were preserved after CSA are indicated with “Y” in “CSA?” column for PfPMT. HID123 was manually added to the CSA-selected region due to its catalytic role suggested by experiment.² This leads to a QM region of 264 atoms, 284 with link atoms and a net charge of -1 from a starting QM region of 1019 atoms, 1067 atoms with link atoms, and a charge of -1.

Res. #	Res.	# At	Charge	CSA?	Res. #	Res.	# At	Charge	CSA?
4	LEU	19	0	--	122	HID	17	0	Y
8	GLN	17	0	Y	123	LEU	19	0	--
9	TYR	21	0	Y	131	LEU	19	0	--
14	VAL	16	0	--	150	TYR	21	0	--
17	TYR	21	0	Y	165	TYR	21	0	Y
21	PHE	20	0	--	169	ARG	24	1	Y
24	ASN	14	0	--	171	TYR	21	0	--
25	TYR	21	0	--	200	TRP	24	0	--
26	ILE	19	0	--	204	LEU	19	0	--
27	SER	11	0	--	230	LEU	19	0	--
28	SER	11	0	--	233	GLY	7	0	--
30	GLY	7	0	--	234	TRP	24	0	--
34	THR	14	0	--	236	ARG	24	1	--
51	ASP	12	-1	--	237	LYS	22	1	Y
52	ILE	19	0	--	246	GLN	17	0	--
53	GLY	7	0	--	257	OPE	14	-2	Y
54	SER	11	0	Y	258	SAM	50	1	Y
55	GLY	7	0	--	263	WAT	3	0	--
56	LEU	19	0	--	266	WAT	3	0	--
57	GLY	7	0	--	267	WAT	3	0	Y
58	GLY	7	0	--	268	WAT	3	0	--
59	GLY	7	0	--	271	WAT	3	0	--
74	ILE	19	0	--	272	WAT	3	0	--
75	ASP	12	-1	Y	275	WAT	3	0	--
76	ILE	19	0	--	276	WAT	3	0	--
77	CYS	11	0	--	280	WAT	3	0	--
80	ILE	19	0	--	284	WAT	3	0	--
81	VAL	16	0	--	286	WAT	3	0	--
99	ASN	14	0	--	291	WAT	3	0	--
100	ASP	12	-1	--	292	WAT	3	0	--
101	ILE	19	0	Y	294	WAT	3	0	--
102	LEU	19	0	--	295	WAT	3	0	--
103	THR	14	0	--	300	WAT	3	0	--
115	TYR	21	0	--	321	WAT	3	0	--
116	SER	11	0	--	335	WAT	3	0	--
117	ARG	24	1	--	346	WAT	3	0	--
118	ASP	12	-1	Y	370	WAT	3	0	--
119	ALA	10	0	--		All	CSA		
120	ILE	19	0	--	Totals	# atoms	1019	264	
121	LEU	19	0	--		with link	1067	284	
						charge	-1	-1	

Table S6. List of QM-treated residues considered for possible QM region inclusion in CSA analysis and those that were preserved after CSA are indicated with “Y” in “CSA?” column for HcgC. SER10 was manually added due to its formation of double hydrogen bond with Thr8 that was shown experimentally to influence significantly the enzymatic catalytic performance. This leads to a QM region of 517 atoms, 545 with link atoms, and a net charge of -4 from an initial QM region of 1000 atoms, 1043 with link atoms, and a net charge of -3.

Res. #	Res.	# At	Charge	CSA?	Res. #	Res.	# At	Charge	CSA?
1	SAM	50	1	Y	238	GLU	15	-1	Y
2	9KH	19	-3	Y	242	PHE	20	0	Y
6	GLY	7	0	--	300	WAT	3	0	Y
7	ILE	19	0	Y	301	WAT	3	0	Y
8	THR	14	0	Y	302	WAT	3	0	--
10	SER	11	0	Y	304	WAT	3	0	Y
11	VAL	16	0	--	306	WAT	3	0	--
20	SER	11	0	--	308	WAT	3	0	Y
23	LEU (N)	21	1	--	309	WAT	3	0	Y
24	TYR	21	0	--	310	WAT	3	0	Y
54	VAL	16	0	--	312	WAT	3	0	Y
55	VAL	16	0	--	313	WAT	3	0	Y
58	LYS	22	1	Y	314	WAT	3	0	Y
77	PHE	20	0	--	317	WAT	3	0	Y
78	GLY	7	0	--	318	WAT	3	0	Y
79	ALA	10	0	--	324	WAT	3	0	--
80	TYR	21	0	Y	327	WAT	3	0	Y
81	LEU	19	0	--	329	WAT	3	0	Y
82	SER	11	0	Y	330	WAT	3	0	Y
83	GLY	7	0	--	331	WAT	3	0	Y
86	ILE	19	0	--	332	WAT	3	0	Y
100	VAL	16	0	--	334	WAT	3	0	Y
101	ASP	12	-1	Y	335	WAT	3	0	Y
102	ILE	19	0	Y	338	WAT	3	0	Y
103	GLN	17	0	Y	347	WAT	3	0	--
105	HIE	17	0	Y	348	WAT	3	0	--
106	LEU	19	0	--	356	WAT	3	0	--
109	ILE	19	0	--	369	WAT	3	0	--
117	PHE	20	0	--	372	WAT	3	0	--
120	LEU	19	0	Y	377	WAT	3	0	--
141	LEU	19	0	--	381	WAT	3	0	--
142	THR	14	0	Y	447	WAT	3	0	--
143	GLY	7	0	Y	457	WAT	3	0	--
144	ILE	19	0	Y	472	WAT	3	0	--
145	GLY	7	0	Y	489	WAT	3	0	--
146	GLY	7	0	Y	515	WAT	3	0	--
147	VAL	16	0	Y	531	WAT	3	0	--
148	SER	11	0	Y	573	WAT	3	0	--
151	LEU	19	0	--	591	WAT	3	0	--
163	GLU	15	-1	Y	599	WAT	3	0	--
165	PRO	14	0	--	651	WAT	3	0	--
192	LEU	19	0	--	663	WAT	3	0	--
194	THR	14	0	Y	664	WAT	3	0	--
203	THR	14	0	--	706	WAT	3	0	--

Res. #	Res.	# At	Charge	CSA?	Res. #	Res.	# At	Charge	CSA?
204	SER	11	0	--	708	WAT	3	0	--
205	GLY	7	0	--	780	WAT	3	0	--
206	THR	14	0	Y	850	WAT	3	0	--
207	MET	17	0	Y	920	WAT	3	0	--
208	THR	14	0	Y	961	WAT	3	0	--
209	LEU	19	0	Y				All	CSA
210	VAL	16	0	--	Totals		# atoms	1000	517
211	VAL	16	0	--			with link	1043	545
							charge	-3	-4

Figure S2. Evolution of C-S and C-O distances (in Å) along the steered-MD trajectory in CMTr1. The differences of these two quantities is the SMD reaction coordinate.

Figure S3. Evolution of C-S and C-O distances (in Å) along the steered-MD trajectory in PIMT. The differences of these two quantities is the SMD reaction coordinate.

Figure S4. Evolution of C-S and C-N distances (in Å) along the steered-MD trajectory in PfPMT. The differences of these two quantities is the SMD reaction coordinate.

Figure S5. Evolution of C-S and C-C distances (in Å) along the steered-MD trajectory in HcgC. The differences of these two quantities is the SMD reaction coordinate.

(a) CMTr1

(b) PIMT

(c) PfPMT

(d) HcgC

Figure S6. Typical CH $\cdots$ X hydrogen bond distances (black lines) and angles (red lines) that are not listed in the main text in (a) CMTr1, (b) PIMT, (c) PfPMT (Tyr19 or Asp128 shown), and (d) HcgC.

Figure S7. Decomposed hydrogen bond energies in the CMTr1 obtained from QTAIM analysis and estimation of the hydrogen bond energy from half of the potential energy density at the bond critical point.

Figure S8. Decomposed hydrogen bond energies in the PIMT obtained from QTAIM analysis and estimation of the hydrogen bond energy from half of the potential energy density at the bond critical point.

Figure S9. Decomposed hydrogen bond energies in the PfPMT obtained from QTAIM analysis and estimation of the hydrogen bond energy from half of the potential energy density at the bond critical point.

Figure S10. Decomposed hydrogen bond energies in the HcgC obtained from QTAIM analysis and estimation of the hydrogen bond energy from half of the potential energy density at the bond critical point.

Figure S11. Correlation of conjugate acid functional group pK_a with computed electrostatic potential in the TS region. The best fit line ($R^2 = 0.8$) is $ESP = 7.36 \cdot pK_a - 333.77$.

Table S7. Residue protonation states in CMTr1. The residue with a charge 0 is left blank. The residue ASP364 is manually protonated, and the G(methyl acceptor) is manually deprotonated.

Res	q	Res	q	Res	q	Res	q	Res	q	Res	q	Res	q	Res	q	Res	q	Res	q	Res	q	Res	q
VAL141		VAL177		GLU213	-1	ARG249	1	GLU285	-1	PHE321		GLY357		SER393	q	LEU429		ILE465		GLU501	-1	ARG537	1
ASP142	-1	GLY178		MET214		MET250		TYR286		GLU322	-1	VAL358		ILE394		PHE430		LYS466	1	SER502		LEU538	
TRP143		LYS179	1	ARG215	1	PHE251		VAL287		PRO323		HIE359		VAL395		LYS431	1	LEU467		HID503		TRP539	
ARG144	1	ARG180	1	ARG216	1	THR252		LEU288		TYR324		PHE360		ARG396	1	PRO432		ASN468		CYS504		GLY540	
ASP145	-1	LYS181	1	ALA217		ASN253		TRP289		TYR325		LEU361		THR397		ILE433		GLN469		SER505		ILE541	
GLU146	-1	MET182		ARG218	1	PRO254		ARG290	1	GLY326		MET362		GLY398		THR434		LEU470		LEU506		PRO542	
PRO147		ILE183		THR219		ARG255	1	LYS291	1	GLU327	-1	ALA363		GLY399		SER435		ARG471	1	GLN507		ASP543	-1
GLU148	-1	ILE184		ARG220	1	ASP256	-1	LYS292	1	GLY328		ASP364	0	HIE400		ARG436	1	ASN472		ILE508		GLN544	
PRO149		GLU185	-1	ALA221		SER257		TRP293		GLY329		GLY365		PHE401		PRO437		THR473		LYS509	1	ALA545	
SER150		ASP186	-1	ASN222		TYR258		HIE294		ILE330		GLY366		ILE402		ALA438		ASP474	-1	ALA510		ARG546	1
ALA151		GLU187		PRO223		GLY259		ALA295		ASP331	-1	PHE367		CYS403		ASN439		SER475		LEU511		VAL547	
CYS152		THR188		TYR224		LYS260	1	LYS296	1	GLY332		SER368		LYN404		SER440		ASP476	-1	ALA512		SAM	1
GLU153	-1	GLU189	-1	GLU225	-1	PRO261		GLY297		ASP333	-1	VAL369		THR405		GLH441		VAL477		LYS513	1	MGT	-1
GLN154		PHE190		MET226		LEU262		PHE298		GLY334		GLU370	-1	PHE406		ARG442	1	ASN478		ILE514		G	-2
VAL155		CYS191		ILE227		VAL263		GLY299		ASP335	-1	GLY371		ASP407	-1	TYR443		LEU479		ILE515		A	-1
SER156		GLY192		ARG228	1	LYS264	1	MET300		ILE336		GLN372		LEU408		VAL444		VAL480		ALA516		U	-1
TRP157		GLU193	-1	GLY229		ASP265	-1	THR301		THR337		GLU373	-1	PHE409		VAL445		VAL481		PHE517		C	-1
PHE158		LEU194	-1	VAL230		ARG266	1	LEU302		ARG338	1	ASN374		THR410		CYS446		PRO482		VAL518			
PRO159		LEU195		PHE231		GLU267	-1	LYS303	1	PRO339		LEU375		PRO411		LYS447	1	LEU483		GLN519			
GLU160	-1	LEU196		PHE232		ALA268		GLY304		GLU340	-1	GLN376		PHE412		GLY448		GLU484	-1	ASP520	-1		
CYS161		HIP197	1	LEU233		GLU269	-1	PRO305		ASN341		GLU377	-1	SER413		LEU449		VAL485		THR521			
THR162		SER198		ASN234		LEU270		ASN306		ILE342		ILE378		VAL414		LYS450	1	ILE486		THR522			
THR163		VAL199		ARG235	1	LEU271		ASP307	-1	SER343		LEU379		GLY415		VAL451		LYS487	1	LEU523			
GLU164	-1	LEU200		ALA236		TYR272		PHE308		ALA344		SER380		LEU416		GLY452		GLY488		SER524			
ILE165		GLN201		ALA237		PHE273		LYS309	1	PHE345		LYS381	1	VAL417		ILE453		ASP489	-1	GLU525	-1		
PRO166		CYS202		MET238		ALA274		LEU310		ARG346	1	GLN382		TYR418		ASP454	-1	HIP490	1	PRO526			
ASP167	-1	LYS203	1	LYS239	1	ASP275	-1	GLU311	-1	ASN347		LEU383		LEU419		ASP455	-1	GLU491	-1	ARG527	1		
THR168		SER204		MET240		VAL276		ASP312	-1	PHE348		LEU384		LEU420		VAL456		PHE492		GLN528			
GLN169		VAL205		ALA241		CYS277		PHE313		VAL349		LEU385		TYR421		ARG457	1	THR493		ALA529			
GLU170	-1	PHE206		ASN242		ALA278		TYR314		LEU350		CYS386		CYS422		ASP458	-1	ASP494	-1	GLU530	-1		
MET171		ASP207	-1	MET243		GLY279		SER315		ASP351	-1	GLN387		CYS423		TYR459		TYR495		ILE531			
SER172		VAL208		ASP244	-1	PRO280		ALA316		ASN352		PHE388		PHE424		LEU460		MET496		ARG532	1		
ASP173	-1	LEU209		PHE245		GLY281		SER317		THR353		LEU389		GLU425	-1	PHE461		ILE497		LYS533	1		
TRP174		ASP210	-1	VAL246		GLY282		SER318		ASP354	-1	MET390		ARG426	1	ALA462		ARG498	1	GLU534	-1		
MET175		GLY211		PHE247		PHE283		GLU319	-1	ARG355	1	ALA391		VAL427		VAL463		SER499		CYS535			
VAL176		GLU212	-1	ASP248	-1	SER284		LEU320		LYS356	1	LEU392		CYS428		ASN464		ASN500		LEU536			

Table S8. Residue protonation states in PIMT. The residue with a charge 0 is left blank.

Res	q	Res	q	Res	q	Res	q	Res	q	Res	q	Res	q	Res	q	Res	q	Res	q	Res	q	Res	q
GLU14	-1	ILE33		VAL52		GLY71		LYS90	1	ILE109		GLU128	-1	GLY147		ALA166		LEU185		THR204		GLY223	
LYS15	1	ILE34		GLU53	-1	GLN72		PRO91		SER110		PHE129		ASP148	-1	GLY167		ILE186		LYS205	1	GLU224	-1
GLU16	-1	ARG35	1	ASP54	-1	THR73		GLY92		GLU111	-1	ALA130		GLY149		ALA168		ILE187		ASP206	-1	TYR225	
LEU17		SER36		LYS55	1	VAL74		MET93		ILE112		LYS131	1	SER150		PRO169		PRO188		GLY207		GLY226	
TYR18		LYS37	1	TYR56		SER75		ASN94		VAL113		ARG132	1	LYS151	1	LYS170	1	VAL189		ILE208		TRP227	
GLU19	-1	GLU38	-1	LYS57	1	ALA76		ILE95		LYS114	1	ASN133		GLY152		ILE171		GLY190		LYS209	1	LYS228	1
LYS20	1	VAL39		LYS58	1	PRO77		LEU96		THR115		LEU134		PHE153		PRO172		SER191		ILE210		SAM229	1
TRP21		GLU40	-1	TYR59		HIP78	1	GLU97	-1	ASP116	-1	GLU135	-1	PRO154		GLU173	-1	TYR192		LYS211	1	VAL230	
MET22		ARG41	1	ALA60		MET79		VAL98		VAL117		ARG136	1	PRO155		PRO174		HID193		ASN212		TYR231	
ARG23	1	ALA42		HIE61		VAL80		GLY99		TYR118		ALA137		LYS156	1	LEU175		LEU194		HIE213		PRO232	
THR24		PHE43		ILE62		ALA81		THR100		THR119		GLY138		ALA157		ILE176		TRP195		GLY214		IAS233	-1
VAL25		LEU44		ASP63	-1	ILE82		GLY101		ILE120		VAL139		PRO158		GLU177	-1	GLN196		GLY215		HIE234	
GLU26	-1	LYS45	1	GLU64	-1	MET83		SER102		GLU121	-1	LYS140	1	TYR159		GLN178		GLU197	-1	VAL216		ALA235	
MET27		TYR46		PRO65		LEU84		GLY103		ARG122	1	ASN141		ASP160	-1	LEU179		LEU198		ALA217			
LEU28		PRO47		LEU66		GLU85	-1	TRP104		ILE123		VAL142		VAL161		LYS180	1	LEU199		PHE218			
LYS29	1	ARG48	1	PRO67		ILE86		ASN105		PRO124		HIP143	1	ILE162		ILE181		GLU200	-1	VAL219			
ALA30		TYR49		ILE68		ALA87		ALA106		GLU125	-1	VAL144		ILE163		GLY182		VAL201		PRO220			
GLU31	-1	LEU50		PRO69		ASN88		ALA107		LEU126		ILE145		VAL164		GLY183		ARG202	1	LEU221			
GLY32		SER51		ALA70		LEU89		LEU108		VAL127		LEU146		THR165		LYS184	1	LYS203	1	ILE222			

Table S9. Residue protonation states in PfPMT. The residue with a charge 0 is left blank.

Res	q	Res	q	Res	q	Res	q	Res	q	Res	q	Res	q	Res	q	Res	q	Res	q	Res	q		
LYS11	1	GLU33	-1	GLU55	-1	TYR77		SER99		PHE121		GLN143		LYS165	1	GLU187	-1	TYR209		LEU231		ARG253	1
THR12		ASN34		ASN56		GLY78		GLY100		ASP122	-1	LYS144	1	GLU166	-1	GLU188	-1	TRP210		PHE232		LYS254	1
PHE13		TYR35		SER57		ALA79		ASN101		LEU123		CYS145		ASN167		TYR189		ASN211		SER233		MET255	
LEU14		ILE36		LYS58	1	HID80		ASN102		ILE124		TYR146		TRP168		ALA190		GLN212		GLU234	-1	GLN256	
GLU15	-1	SER37		VAL59		THR81		LYS103	1	TYR125		LYS147	1	ASP169	-1	ASP191	-1	LEU213		LYS235	1	ARG257	1
ASN16		SER38		LEU60		HIE82		ILE104		SER126		TRP148		ASP170	-1	ILE192		LEU214		LYS236	1	TRP258	
ASN17		GLY39		ASP61	-1	GLY83		ILE105		ARG127	1	LEU149		GLU171	-1	LEU193		GLU215	-1	PHE237		GLY259	
GLN18		GLY40		ILE62		ILE84		PHE106		ASP128	-1	LYS150	1	PHE172		THR194		VAL216		ILE238		TYR260	
TYR19		LEU41		GLY63		ASP85	-1	GLU107	-1	ALA129		PRO151		LYS173	1	ALA195		GLU217	-1	SER239		PHE261	
THR20		GLU42	-1	SER64		ILE86		ALA108		ILE130		THR152		GLU174	-1	CYS196		HIP218	1	LEU240		LYS262	1
ASP21	-1	ALA43		GLY65		CYS87		ASN109		LEU131		GLY153		TYR175		ASN197		LYS219	1	ASP241	-1	ALA263	
GLU22	-1	THR44		LEU66		SER88		ASP110	-1	HID132		THR154		VAL176		PHE198		TYR220		ASP242	-1	THR264	
GLY23		LYS45	1	GLY67		ASN89		ILE111		LEU133		LEU155		LYS177	1	LYS199	1	LEU221		GLY243		LYS265	1
VAL24		LYS46	1	GLY68		ILE90		LEU112		SER134		LEU156		GLN178		ASN200		HIP222	1	TRP244		ASN266	
LYS25	1	ILE47		GLY69		VAL91		THR113		LEU135		ILE157		ARG179	1	VAL201		GLU223	-1	SER245		OPE267	-2
VAL26		LEU48		CYS70		ASN92		LYS114	1	GLU136	-1	THR158		LYS180	1	VAL202		ASN224		ARG246	1	SAM268	1
TYR27		SER49		MET71		MET93		GLU115	-1	ASN137		ASP159	-1	TYR181		SER203		LYS225	1	LYS247	1		
GLU28	-1	ASP50	-1	TYR72		ALA94		PHE116		LYS138	1	TYR160		THR182		LYS204	1	GLU226	-1	ILE248			
PHE29		ILE51		ILE73		ASN95		PRO117		ASN139		CYS161		LEU183		ASP205	-1	GLU227	-1	LYS249	1		
ILE30		GLU52	-1	ASN74		GLU96	-1	GLU118	-1	LYS140	1	ALA162		ILE184		LEU206		PHE228		ASP250	-1		
PHE31		LEU53		GLU75	-1	ARG97	1	ASN119		LEU141		THR163		THR185		SER207		LEU229		SER251			
GLY32		ASN54		LYS76	1	VAL98		ASN120		PHE142		GLU164	-1	VAL186		ASP208	-1	LYS230	1	LYS252	1		

Table S10. Residue protonation states in HcgC. The residue with a charge 0 is left blank.

Res	q	Res	q	Res	q	Res	q	Res	q	Res	q	Res	q	Res	q	Res	q	Res	q	Res	q	Res	q
SAM	1	ILE24'		VAL51		VAL76		ASP101	-1	GLU126	-1	LEU151		ILE176		SER201		SER226		LEU251		SER276	
9KH	-3	PRO25'		SER52		PHE77		ILE102		LEU127		ILE152		ASP177	-1	LYS202	1	VAL227		THR252		LYS277	1
MET1'		ASN26'		ASP53	-1	GLY78		GLN103		ARG128	1	SER153		ASN178		THR203		LEU228		GLU253	-1	ASN278	
ASN2'		LEU27'		VAL54		ALA79		PRO104		ASN129		LYS154	1	THR179		SER204		TYR229		LEU254		ILE279	
TYR3'		MET30		VAL55		TYR80		HIE105		GLY130		PHE155		ASP180	-1	GLY205		THR230		ASN255		SER280	
GLY4'		ASN31		GLU56	-1	LEU81		LEU106		THR131		ASN156		LYS181	1	THR206		ILE231		MET256		ARG281	1
ILE5'		TYR32		LYS57	1	SER82		LYS107	1	SER132		PRO157		ARG182	1	MET207		PRO232		SER257		VAL282	
THR6'		GLY33		LYS58	1	GLY83		ASP108	-1	ILE133		LYS158	1	LEU183		THR208		ASN233		ALA258		ASP283	-1
GLU7'	-1	ILE34		ALA59		SER84		ILE109		ASN134		VAL159		CYS184		LEU209		LEU234		ILE259		SER284	
SER8'		THR35		ASN60		TYR85		LEU110		PRO135		LEU160		VAL185		VAL210		LYS235	1	THR260		PHE285	
VAL9'		GLU36	-1	ALA61		ILE86		PHE111		ASP136	-1	ILE161		GLY186		VAL211		TYR236		VAL261		VAL286	
LYS10'1	1	SER37		ILE62		ALA87		ASN112		LEU137		VAL162		ALA187		ASP212	-1	PHE237		SER262		LYS287	1
THR11'1		VAL38		LYS63	1	TYR88		ASP113	-1	VAL138		GLU163	-1	LYS188	1	ILE213		GLU238	-1	SER263		GLU288	-1
THR12'1		LYS39	1	TYR64		SER89		GLY114		ILE139		ASP164	-1	LYS189	1	ILE214		GLY239		ILE264		PHE289	
ARG13'1	1	THR40		PHE65		LEU90		ILE115		ASP140	-1	PRO165		GLY190		MET215		THR240		ASP265	-1	ASP290	-1
SER14'1		THR41		LEU66		LEU91		LYS116	1	LEU141		LYS166	1	VAL191		ASP216	-1	VAL241		HIE266			
ILE15'1		ARG42	1	GLU67	-1	LYS92	1	PHE117		THR142		GLY167		LEU192		SER217		PHE242		VAL267			
THR16'1		SER43		GLY68		ASP93	-1	MET118		GLY143		ASN168		LYS193	1	CYS218		HIE243		GLU268	-1		
VAL17'1		LYS44	1	GLU69	-1	CYS94		ASP119	-1	ILE144		HIE169		THR194		ARG219	1	GLU244	-1	TYR269			
SER18'1		ILE45		GLU70	-1	GLU95	-1	LEU120		GLY145		ASP170	-1	TYR195		GLU220	-1	LYS245	1	GLU270	-1		
SER19'1		LYS46	1	PHE71		GLU96	-1	ASN121		GLY146		LYS171	1	ARG196	1	ILE221		ASN246		LEU271			
ILE20'1		ILE47		LYS72	1	VAL97		LYS122	1	VAL147		GLY172		SER197		ASN222		VAL247		GLU272	-1		
LEU21'1		LYS48	1	GLN73		ILE98		LEU123		SER148		ILE173		SER198		GLU223	-1	LYS248	1	GLU273	-1		
TYR22'1		ASP49	-1	ALA74		ILE99		GLN124		PRO149		SER174		LYS199	1	LEU224		LYS249	1	ILE274			
THR23'1		ILE50		ILE75		VAL100		LEU125		ASP150	-1	LYS175	1	PHE200		ASP225	-1	PHE250		LEU275			

Text S1. Extended methods discussion.

Structures for MTases were obtained from the protein data bank, combining a SAM-bound binary complex with an SAH- and substrate/product-bound ternary structure, respectively: CMTr1³ (PDB IDs: 4N49, 4N48); PIMT⁴ (PDB IDs: 1JG1, 1JG4); PfPMT² (PDB IDs: 3UJ7, 3UJB); and HcgC⁵ (PDB IDs: 5O4H, 5O4J). SAM was preserved from the binary structure and used to replace SAH in the ternary structure, and all crystallizing agents were removed. The resulting holoenzymes are of varying size and carry a net charge: 6705 atoms and -17 charge in CMTr1, 3616 atoms and +5 charge in PIMT, 4245 atoms and -4 charge in PfPMT, and 8545 atoms and -8 charge in HcgC (homodimer). Final system sizes of 104,264, 64,923, 76,627, and 76,667 atoms in CMTr1, PIMT, PfPMT and HcgC, respectively. Equilibration steps were as follows: i) 3000 minimization steps, ii) 10-ps NVT heating to 300 K with a Langevin thermostat with collision frequency of 1.0 ps⁻¹ and a random seed, iii) 250-ps NPT equilibration using the Berendsen barostat with a pressure relaxation time of 2 ps, and iv) 10-ns NPT production run. The SHAKE algorithm⁶ was applied in combination with a 2-fs timestep for all MD simulations. For the long-range electrostatics, the particle mesh Ewald method was used with a 10-Å electrostatic cutoff. The equilibrated snapshots were truncated to a spherical shape that included 8,000-10,000 waters but no counterions using the closestwater command in the AMBER cpptraj utility, where the range corresponds to differences in MTase size and shape. To ensure that the system was neutralized in the sphere, the counterions were then added and a spherical boundary condition was applied with the solvatecap command in tleap centered at the droplet center of mass and circumscribing the extracted water molecules. In HcgC, a truncated model was used to include one monomer and the water-shielding loops of the second monomer. The final QM/MM systems consisted of 31,940, 33,072, 31,265, and 31,687 atoms in CMTr1, PIMT, PfPMT and HcgC. Structures of all enzymes as well as prmtop/inpcrd files for further simulation are provided in the supporting information ZIP file. HBs were obtained from QTAIM by computing them as half the potential energy density at the BCP⁷.

Table S11. Details of solvent caps in the construction of AMBER parameter files in the model MTases. For HcgC, the active site resides in the interface of a homodimer. The model is thus truncated to include one monomer and some loop residues shielding the active site.

	Protein and Ligand atom #	Counter Ions #	Sphere radius (Å)	Sphere center	water #	Total atom #
CMTTrl	6705	17 (Na ⁺)	43.5	(46.48, 53.95, 52.82)	8406	31940
PIMT	3616	5 (Cl ⁻)	44.4	(40.92, 46.60, 43.11)	9817	33072
PfPMT	4245	4 (Na ⁺)	43.5	(45.93, 46.10, 45.58)	9025	31324
HcgC	4697	5 (Na ⁺)	41	(50.34, 45.93, 38.31)	7998	28696

Table S12. Snapshot extraction interval for single point electronic structure calculations in the MTases.

	Range of RC	SMD Steps	Snapshot interval
CMTTrl	[-1.25, 2.0]	2280	2
PIMT	[-1.50, 2.0]	2500	2
PfPMT	[-1.25, 2.0]	2280	2
HcgC	[-1.5, 1.5]	3500	3

References

1. Horowitz, S.; Dirk, L. M. A.; 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. *Journal of the American Chemical Society* **2013**, *135* (41), 15536-15548.
2. Lee, S. G.; Kim, Y.; Alpert, T. D.; Nagata, A.; Jez, J. M., Structure and Reaction Mechanism of Phosphoethanolamine Methyltransferase from the Malaria Parasite *Plasmodium falciparum* AN ANTIPARASITIC DRUG TARGET. *Journal of Biological Chemistry* **2012**, *287* (2), 1426-1434.
3. Smietanski, M.; Werner, M.; Purta, E.; Kaminska, K. H.; Stepinski, J.; Darzynkiewicz, E.; Nowotny, M.; Bujnicki, J. M., Structural analysis of human 2'-O-ribose methyltransferases involved in mRNA cap structure formation. *Nature communications* **2014**, *5*, 3004.
4. Griffith, S. C.; Sawaya, M. R.; Boutz, D. R.; Thapar, N.; Katz, J. E.; Clarke, S.; Yeates, T. O., Crystal structure of a protein repair methyltransferase from *Pyrococcus furiosus* with its l-isoaspartyl peptide substrate. *Journal of molecular biology* **2001**, *313* (5), 1103-1116.
5. Bai, L.; Wagner, T.; Xu, T.; Hu, X.; Ermler, U.; Shima, S., A Water-Bridged H-Bonding Network Contributes to the Catalysis of the SAM-Dependent C-Methyltransferase HcgC. *Angewandte Chemie* **2017**, *129* (36), 10946-10949.
6. Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C., Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics* **1977**, *23* (3), 327-341.
7. Espinosa, E.; Molins, E.; Lecomte, C., Hydrogen bond strengths revealed by topological analyses of experimentally observed electron densities. *Chemical Physics Letters* **1998**, *285* (3), 170-173.

MTCTSI_v5.pdf (2.99 MiB)

[view on ChemRxiv](#) • [download file](#)

Other files

structures.zip (16.03 MiB)

[view on ChemRxiv](#) • [download file](#)