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

Bernd Kuhn and Peter A. Kollman*

Contribution from the Department of Pharmaceutical Chemistry, University of California at San Francisco, San Francisco, California 94143-0446

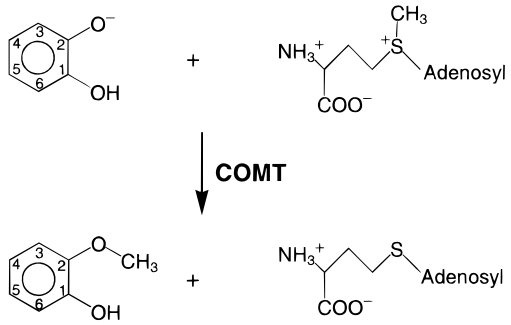
Received June 28, 1999. Revised Manuscript Received January 13, 2000

Abstract: We present combined quantum mechanical and free energy (QM−FE) as well as molecular dynamics (MD) calculations to investigate the methyl transfer reaction catalyzed by the enzyme catechol *O*-methyltransferase (COMT). The calculated transition state free energy, $\Delta G^{\ddagger}$, of 24.5 kcal/mol for the enzymatic reaction is in reasonable agreement with experiment (18 kcal/mol), and the level of agreement improves when the substrates are allowed to slightly deviate from their molecular mechanical energy minimized interpolated geometries. The calculated $\Delta G^{\ddagger}$ for the reaction in water is 5 kcal/mol higher than that found in the enzyme when cratic free energy terms are not considered, and this difference increases to 14−18 kcal/mol when they are included, in very good agreement with the estimated rate enhancement due to enzyme catalysis of 10^{11} (corresponding to $\Delta\Delta G^{\ddagger} \approx 15$ kcal/mol). In contrast to trypsin, studied earlier by QM−FE methods, the calculated gas phase $\Delta G^{\ddagger}$ is significantly lower than observed in the enzyme or in solution; thus, in COMT the enzyme must avoid increasing the transition state barrier as much as is observed in solution. In addition to these free energy calculations, we describe MD simulations on various monosubstituted catechols and we use such calculations to rationalize the significant regioselectivity for attack at the meta rather than para OH group relative to the substituent.

I. Introduction

Catechol *O*-methyltransferase (COMT, EC 2.1.1.6) is an essential enzyme that inactivates important neurotransmitters with catechol structure such as dopamine, noradrenaline, and adrenaline. It utilizes the enzymatic cofactor *S*-adenosyl-L-methionine (SAM) as a CH_3 donor to transform one of the two aromatic OH groups of catechols to OCH_3. A number of mechanistic studies has shown that the rate-limiting step in this reaction is the attack of a methyl group originally attached to the sulfur of SAM on an aromatic O^- of the catechol in a direct S_N2 process.[1−4] This transmethylation reaction for the unsubstituted catechol yielding catechol methyl ether and *S*-adenosyl-L-homocysteine (SAH) is depicted in Scheme 1.

Scheme 1



Recently, Bruice et al. carried out both quantum mechanical (QM) and molecular dynamics (MD) calculations relevant to the reaction catalyzed by COMT.[5,6] In the quantum mechanical study on $(CH_3)_3S^+$ reacting with catecholate in the gas phase and in continuum solvent, they pointed out the important role of solvation/desolvation for this reaction. The calculated kinetic isotope effects for the gas phase model reaction were found in good agreement with experimental values,[1] suggesting similar transition state (TS) geometries in COMT and in the nonenzymatic reaction. In molecular dynamics studies on the enzyme complexed with cofactor and substrate, they showed that during a 1 ns simulation, the attacking methyl of SAM and the O^- of the catechol were in a near attack configuration (NAC)[7] a large portion of the time, thus giving some insight into how the enzyme catalyzes this reaction.

Nonetheless, Bruice et al. did not directly study the energetics of the enzyme-catalyzed reaction and compare this to the same reaction in solution. It is important that this be done in order to fully describe the basis for how the enzyme catalyzes this

* To whom correspondence should be addressed. Phone: (415) 476-4637. Fax: (415) 476-0688. E-mail: pak@cgl.ucsf.edu.

(1) Hegazi, M. F.; Borchardt, R. T.; Schowen, R. L. *J. Am. Chem. Soc.* **1979**, *101*, 4359−4365.

(2) Woodard, R. W.; Tsai, M. D.; Floss, H. G.; Crooks, P. A.; Coward, J. K. *J. Biol. Chem.* **1980**, *255*, 9124−9127.

(3) Knipe, J. O.; Vasquez, P. J.; Coward, J. K. *J. Am. Chem. Soc.* **1982**, *104*, 3202−3209.

(4) Rodgers, J.; Femec, D. A.; Schowen, R. L. *J. Am. Chem. Soc.* **1982**, *104*, 3263−3268.

(5) Zheng, Y. J.; Bruice, T. C. *J. Am. Chem. Soc.* **1997**, *119*, 8137−8145.

(6) Lau, E. Y.; Bruice, T. C. *J. Am. Chem. Soc.* **1998**, *120*, 12387−12394.

(7) Lightstone, F. C.; Bruice, T. C. *J. Am. Chem. Soc.* **1996**, *118*, 2595−2605.

10.1021/ja992218v CCC: $19.00
Published on Web 03/02/2000

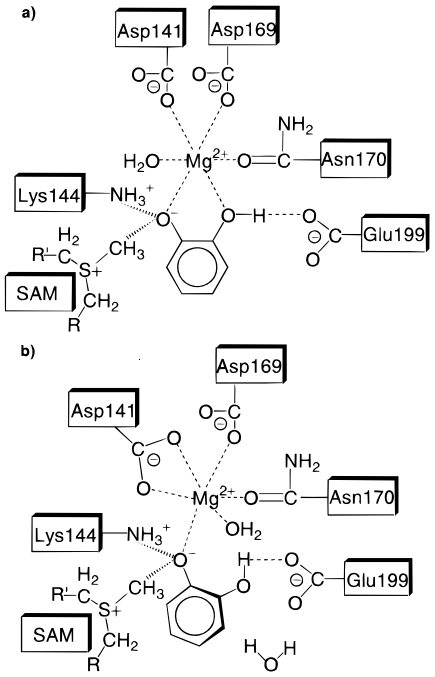


Figure 1. Schematic representation of the active site of COMT. (a) Starting geometry for our simulations: X-ray structure of the COMT−inhibitor complex,[10] with the inhibitor 3,5-dinitrocatechol replaced by catechol. (b) Structure of the Michaelis complex after MD equilibration and minimization.

reaction by increasing its rate by $\sim 10^{11}$ compared to that found in solution.[8] We have used the combination of high-level quantum mechanical and classical free energy calculations (QM−FE) to study the reaction pathway for *S*-methyl transfer in aqueous solution and in the active site of COMT. The free energies of activation for both the enzyme and solution reaction are calculated in reasonable agreement with experiment.

This result enables us to compare our computations with the only comparable previous QM−FE calculation in the literature, the formation of a tetrahedral intermediate in the amide hydrolysis in trypsin compared to the corresponding reaction in solution.[9] While in that previous study we only compared the free energies between the Michaelis complex and the tetrahedral intermediate as a model for the transition state of acylation, we are able to go beyond that in the work here and calculate free energies for the entire pathway of methyl transfer from SAM to catechol. It is particularly interesting to compare the results of those two studies, since, in the trypsin case, one creates charge in the transition state whereas, in COMT, one is annihilating charge in the transition state. How do the enzymes both stabilize more polar (trypsin) and less polar (COMT) transition states relative to water? We are able to give a detailed answer to this question below.

Both Bruice's molecular dynamics[6] and this study make use of the X-ray structure of COMT complexed with the inhibitor 3,5-dinitrocatechol (DNC), which is apparently forming a bidentate coordination of the two hydroxyl groups with the Mg^{2+} ion in the active site of COMT.[10] Bruice et al. have assumed a similar coordination for the unsubstituted catechol substrate, but unrestrained simulations suggest monodentate coordination of the deprotonated O^- of catechol with the neighboring OH protonated and hydrogen bonding to Glu 199. Bruice et al. have used the results of their simulation to rationalize the observed meta:para ratios for the methyl attack on the catechol for different substituents at the R_4 position of the ring;[11,12] given that our coordination geometry is different, we provide an alternate explanation.

II. Methods

In this section, we give details of the computational procedure of our QM−FE and molecular dynamics simulations. The theoretical background of our QM−FE approach for enzymatic reactions, which is an extension of the work of Jorgensen et al. who pioneered QM−FE calculations for organic reactions,[13,14] has been described previously.[9] Briefly, we combine ab initio calculations for the reactive part of the enzyme−substrate complex with a classical treatment of the free energy of interaction between the QM system and the classical environment as well as within the classical environment itself. In the first step, we cut out the residues of the complex that are crucial in the enzymatic reaction and saturate dangling bonds with hydrogen atoms. Geometry optimizations of this quantum mechanical model system are performed for different points along the reaction pathway, each with constraints on a few selected internal coordinates, which ensure that the relative orientation of the model fragments stays close to the preorganized geometry in the enzyme. It is important that these constraints are included in QM treatments of enzymatic reaction centers because pure gas phase optimizations in which the fragments are allowed to move freely often yield an energetically very different reaction pathway.[9] On the basis of the QM-optimized structures we derive atomic partial charges using the restrained electrostatic potential (RESP) fitting methodology[15] for each geometry along the reaction path and reinsert the model fragments into the protein coordinate frame. The use of RESP fitted atomic charges for the QM atoms is advantageous in two respects. First, since the same technique has been used for the derivation of the molecular mechanical (MM) charges, we treat electrostatic interactions between all atoms of the system on an equal level. Second, the possibility to define general constraints in the fitting process allows us to reduce the charges on the link atoms, thereby minimizing partial charge artifacts at the boundary between QM and MM regions. Using the standard force field representation for the QM atoms extended by a few additional potential terms to restrain the reinserted model system to its QM-optimized structure, we calculate the free energy of interaction. The overall free energy change, ΔG_{tot}, between two points along the reaction path is approximated as

$$\Delta G_{tot} = \Delta E_{QM} + \Delta G_{int} \quad (1)$$

where ΔE_{QM} denotes the difference in ab initio energy and ΔG_{int} is the difference of the free energy of interaction.

In order to compare the calculated reaction barrier heights with those derived from experimental reaction rates, one has to take into account the free energy contribution of bringing the reactants together in a reactive geometry. While, in the enzyme, this cratic free energy term is absorbed in the binding free energy of the Michaelis complex, it has to be calculated for the solution reaction. We use both quantum mechanical polarizable continuum model (PCM)[16] and potential of mean force (PMF) calculations as independent approaches to estimate this contribution.

1. Building of the COMT Model. We used the coordinates of the 2.0 Å X-ray crystal structure of rat liver COMT complexed with the inhibitor 3,5-dinitrocatechol as starting point for our simulations.[10] The active site of COMT, after replacing the two nitro groups of DNC with

(8) Mihel, I.; Knipe, J. O.; Coward, J. K.; Schowen, R. L. *J. Am. Chem. Soc.* **1979**, *101*, 4349−4351.
(9) Stanton, R. V.; Perakyla, M.; Bakowies, D.; Kollman, P. A. *J. Am. Chem. Soc.* **1998**, *120*, 3448−3457.
(10) Vidgren, J.; Svensson, L. A.; Liljas, A. *Nature* **1994**, *368*, 354−358.
(11) Creveling, C. R.; Dalgard, N.; Shimizu, H.; Daly, J. W. *Mol. Pharmacol.* **1970**, *6*, 691−696.
(12) Creveling, C. R.; Morris, N.; Shimizu, H.; Ong, H. H.; Daly, J. *Mol. Pharmacol.* **1972**, *8*, 398−409.
(13) Chandrasekhar, J.; Smith, S. F.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1985**, *107*, 154−162.
(14) Chandrasekhar, J.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1985**, *107*, 2974−2975.
(15) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269−10280.
(16) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117.

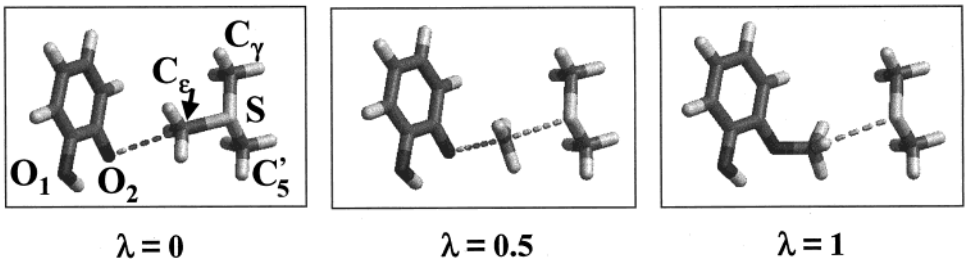
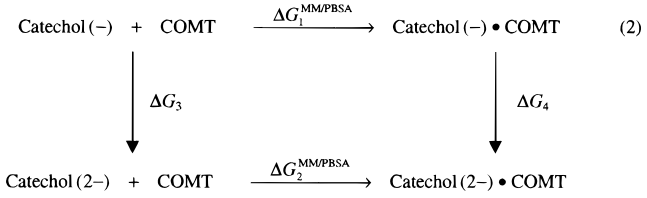


Figure 2. Quantum mechanical model system at three values of the reaction path variable λ. The atom labeling at $\lambda = 0$ is used throughout this work.

hydrogen atoms to model the unsubstituted catechol, is schematically illustrated in Figure 1a. Central in this region is an Mg^{2+} ion, which is octahedrally coordinated to the side chain oxygen atoms of residues Asp141, Asp169, and Asn170 as well as to the two O atoms of catechol and a water molecule. The enzyme nicely orients the transferring methyl group of SAM and one catecholic hydroxyl group in a reactive conformation.

The correct protonation state of catechol bound to COMT has been a subject of discussion.[5,17] Since the pK_a of this substrate in aqueous solution is 9.9,[18] catechol predominantly binds in the neutral form to the enzyme. Supported by experimental studies on methylase models,[3] it was suggested that COMT effects the methyl transfer reaction via general base catalysis. Presumably, the Lys144 residue has a lowered pK_a due to the presence of the Mg^{2+} ion and the positively charged SAM moiety, and acts as a base for proton abstraction from one catecholic hydroxyl group. Similar base catalysis function of lysine residues has been suggested for other enzymatic systems.[19] Semiempirical gas phase model calculations of the active site even predict doubly ionized catechol to be most stable when bound to COMT,[17] but should be interpreted with caution because of the neglect of solvation yielding unrealistically high energy differences between protonation states. There are two reasons that make it unlikely that catechol is bound to the enzyme in the dianionic form. First, the second pK_a of catechol is expected to be near 12.3,[20] which is 2.4 pK units higher than the first pK_a. More importantly, the hydroxyl group of catechol that is not involved in the reaction has no appropriate basic functional group for proton abstraction close to it, in contrast to the case for the other OH group of catechol, which has the neighboring Lys144 residue. As can be seen from Figure 1a, the acidic side chain of Glu199 (pK_a of side chain of free Glu = 4.3)[21] would have to be protonated to stabilize a doubly ionized catechol ligand. It is unrealistic to expect that the enzyme environment would overcome a difference of 8 pK_a units and, hence, that catechol is a dianion. Although metal ions can shift the pK_a by 6−7 units, this occurs only for water molecules directly bound to divalent cations, where the stabilization of an OH^- is much larger than with any other molecule, including catechol.

To further justify our choice of the protonation state of catechol in the enzyme, we performed free energy calculations using the MM/PBSA approach.[22] This method is based on a combination of MM complex interaction energies and a solvation free energy part, which consists of a continuum solvent Poisson−Boltzmann (PB) and a surface area (SA) dependent nonpolar contribution. Details of this computational technique, which proved powerful in calculating free energy differences between protein−ligand complexes,[23] have been described previously.[24] Given the following thermodynamic cycle,



one can estimate the free energy for abstracting the second hydroxyl proton from the catechol monoanion in the enzyme−ligand complex as

$$\Delta G_4 = \Delta G_3 + \Delta G_2^{\mathrm{MM/PBSA}} - \Delta G_1^{\mathrm{MM/PBSA}} \qquad (3)$$

Here, $\Delta G_1^{MM/PBSA}$ and $\Delta G_2^{MM/PBSA}$ are the binding free energies between COMT and monoanionic and dianionic catechol, respectively, which can be calculated using our MM/PBSA approach. ΔG_3 is the free energy for forming the catechol dianion from the monoanion in neutral aqueous solution and can be derived from the difference between the second pK_a of catechol and pH = 7, which is $\Delta p = 5.3$, to yield $\Delta G_3 = 2.303\ RT\ \Delta p = 7.3$ kcal/mol. Based on eq 3, our calculations[25] of $\Delta G_1^{MM/PBSA} = -20.8$ kcal/mol and $\Delta G_2^{MM/PBSA} = -18.4$ kcal/mol lead to a $\Delta G_4 = 9.7$ kcal/mol. Thus, our calculations predict that the monoanionic form of catechol is 10 kcal/mol thermodynamically more stable in the protein than the dianion.

Given our free calculations and the unlikely event that COMT shifts the second pK_a by 8 units, we have modeled catechol as a monoanion. The same protonation state for catechol has been assumed by Bruice et al.[5,6]

2. MM Parameters and Equilibration. We used the force field representation of AMBER 5[26] together with new parameters for the ligands which were derived as outlined previously.[27] The magnesium ion had an atomic charge of +2 in our simulations, and its van der Waals parameters were adjusted from Åqvist ($R^*_{Mg} = 0.7867$ Å, ϵ_{Mg}

(17) Ovaska, M.; Yliniemela, A. *J. Comput.- Aided Mol. Des.* **1998**, *12*, 301−307.
(18) Barratt, M. D. *Toxicol. Lett.* **1995**, *80*, 69−74.
(19) Paetzel, M.; Dalbey, R. E. *Trends Biochem. Sci.* **1997**, *22*, 28−31.
(20) Perrin, D. D.; Dempsey, B.; Serjeant, E. P. *pK_a Prediction for Organic Acids and Bases*; Chapman and Hall: London and New York, 1981.
(21) Stryer, L. *Biochemistry*, 3rd ed.; W. H. Freeman and Company: New York, 1988.
(22) Srinivasan, J.; Cheatham, T. E.; Cieplak, P.; Kollman, P. A.; Case, D. A. *J. Am. Chem. Soc.* **1998**, *120*, 9401−9409.
(23) Kuhn, B.; Kollman, P. A. *J. Am. Chem. Soc.*, in press.
(24) Massova, I.; Kollman, P. A. *J. Am. Chem. Soc.* **1999**, *121*, 8133−8143.

(25) The complex binding free energy was calculated from the difference in MM/PBSA free energies for the ligand−protein complex and uncomplexed reactants. Ensembles of structures (25 snapshots) for the MM/PBSA calculation were obtained from 150 ps MD simulations of the solvated complex between COMT and monoanionic and dianionic catechol, respectively, using the MD protocol described in the Methods section II.2. The MM/PBSA free energy of a molecule, G_{mol}, has contributions from $G_{mol} = E_{MM} + G_{solv} - TS$. E_{MM} denotes the MM energies of a molecule in the gas phase and was calculated using AMBER with no cutoff for the evaluation of nonbonded interactions. The solvation energy, G_{solv}, consists of a nonpolar term which was computed from the solvent accessible surface area and of an electrostatic term which was calculated using a finite-difference Poisson−Boltzmann (PB) model. In the PB treatment, we used Cornell et al. charges[26] and standard PARSE atomic radii (Sitkoff, D.; Sharp, K. A.; Honig, B. *J. Phys. Chem.* **1994**, *98*, 1978−1988), augmented by a Mg^{2+} ion radius of $r = 1.44$ Å, which was chosen to reproduce the experimental solvation free energy of −456 kcal/mol.[28] Solute entropy corrections $-TS$ were neglected because they are expected to be similar for the monoanionic and dianionic complex formations and would thus cancel in the thermodynamic cycle (2) and in eq 3.
(26) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(27) Fox, T.; Kollman, P. A. *J. Phys. Chem. B* **1998**, *102*, 8070−8079.

= 0.8751 kcal/mol).[28] Apart from the metal ion a new atom type was defined for the positively charged sulfur atom of SAM. From test calculations of the minimum energy distance between the S and O atoms of $(CH_3)_3S^+$ monohydrate, using Møller−Plesset second-order perturbation theory (MP2) calculations as reference, we derived its optimal nonbonded parameters (R^*_S = 2.0000 Å, ϵ_S = 0.2500 kcal/mol), which are similar to those of neutral S.[26]

The complex formed by COMT, SAM, catechol, and the crystallographic water was further solvated with a 20 Å sphere of TIP3P[29] water molecules centered at the transferring methyl group, and two Na^+ counterions were placed at the locations of largest negative electrostatic potential to neutralize the system. The counterions, which had distances larger than 16 Å from the active site, were fixed in space during all simulations to avoid artifactual long-range electrostatic effects on the calculated free energies. After energy minimization of the water positions for 1500 steps and MD equilibration of the water sphere with fixed solute for 20 ps we minimized the whole system with progressively smaller positional restraints on the solute [from 25 to 0 kcal/(mol Å^2) for a total of 4000 steps]. We subsequently equilibrated our solvated COMT model for 20 ps and, after filling up the 20 Å water sphere with additional TIP3P molecules to compensate for the loss due to diffusion, continued the MD equilibration for a further 20 ps. The equilibrated system was finally minimized for 2000 steps and yielded our model for the Michaelis complex. During all steps of the equilibration procedure we used a nonbonded cutoff of 16 Å and only allowed protein residues within 16 Å of the active site to move. The MD simulations were performed at constant T = 300 K using the Berendsen coupling scheme,[30] and with a time step of 1.5 fs and the SHAKE algorithm.[31] In order to keep the two reactants in a reactive geometry during the MD simulations, we constrained the distance between the transferring methyl carbon and the O^- of catechol, r_{OC}, to be 2.9 Å using a 20 kcal/(mol Å^2) harmonic restraint force constant. During equilibration, we found this distance to fluctuate around r_{OC} = (2.97 ± 0.09) Å. Even when we do not use this restraint, r_{OC} stays at (3.07 ± 0.15) Å, indicating no distorting effect of the bond length restraint. Finally, it should be noted that this restraint has no effect on the reaction energetics because the final minimization to obtain the Michaelis complex was done without constraints.

3. Quantum Mechanical Optimization. The choice of the quantum mechanical region has to be a compromise between including all the atoms of residues that are affecting the electronic structure of the reaction center at an accurate theoretical level and the computational cost of the QM calculation. Since semiempirical methods are hampered by the lack of quantitative accuracy in dealing with transition states,[32] we used large-scale ab initio and density functional theory (DFT) calculations. Our QM system consists of the catecholate anion and a trimethylsulfonium cation as a model for the reactive part of SAM (see Figure 2). In our view, this system maintains the essential aspects of the transmethylation reaction while allowing us to use high-level ab initio approaches. Further incorporating the Mg^{2+} ion and its coordinated residues, as in Figure 1a, into the QM region is currently tractable only with semiempirical methods[17] or ab initio methods using more limited basis sets. We believe that the effects of those residues on the reaction energetics can be described classically using free energy calculations. In our view, it is more important to use an ab initio approach that involves electron correlation and a basis set with sufficient flexibility to accurately describe the zwitterionic character of the complex.

In order to take into account the preorganization of the substrates by COMT we defined several constraints for the QM optimization. Starting from the MM minimized enzyme−substrate complex we cut out our QM model and put hydrogen atoms at the free valences of the QM system. We then used neighboring MM atoms as anchor points and fixed selected internal coordinates that cross the QM/MM boundary during the quantum mechanical optimization.[33] This set of *Z*-matrix constraints allows us to optimize the reaction center quantum mechanically while conserving the relative geometry of the reactants from the enzyme environment.

While the finding of a good reaction coordinate is difficult for most reactions and requires a potential energy scan in multiple dimensions, the relative simplicity of the direct S_N2 methyl transfer suggests a linear interpolation scheme to be a reasonable approximation. We generated the enzymatic reaction path in the following way. First, we obtained a model of the COMT−product complex by forming the bond between the transferring methyl carbon and the O^- of catechol in the optimized structure of the Michaelis complex. This geometry was subsequently minimized for 1000 steps using a conjugate gradient algorithm. We then defined the constraints along the reaction path by linearly interpolating the anchor points and the internal coordinate constraints for the QM optimization between the minimized enzyme−reactant and enzyme−product structure. To simulate the methyl transfer from $(CH_3)_3S^+$ to catechol, we further applied a constraint to the distance r_{SC} between the transferring methyl carbon and the sulfur atom which was interpolated between the QM-optimized distances of 1.820 and 3.279 Å for the reactant and product dimer, respectively. Introducing the reaction path variable λ, which is 0 for the reactant and 1 for the product state, r_{SC} is given as r_{SC} = 1.820 Å + 1.459λ Å. Using the interpolated constraints we calculated the reaction energy profile by pointwise QM optimizations of the dimer for different values of λ. Figure 2 shows snapshots of the QM model system along this reaction coordinate.

We did constrained geometry optimizations at the Hartree−Fock (HF) and DFT/B3LYP (Becke three-parameter Lee−Yang−Parr)[34−36] levels both using the 6-31+G* basis set.[37] The HF-optimized structures were subsequently used in MP2 single point calculations with 6-31+G* and aug-cc-pVDZ[38] basis sets. Diffuse functions were included in the computations to better describe the anionic character of catechol. All quantum mechanical calculations in this work have been performed using Gaussian 98.[39]

4. Free Energy Calculations. To perform the classical free energy of interaction calculation, we first defined MM parameters for the QM model system. Atomic partial charges, which change considerably

(28) Aqvist, J. *J. Phys. Chem.* **1990**, *94*, 8021−8024.
(29) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(30) Berendsen, H. J. C.; Potsma, J. P. M.; van Gunsteren, W. F.; DiNola, A. D.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(31) Ryckaert, J. P.; Cicotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(32) Mulholland, A. J.; Richards, W. G. *Int. J. Quantum Chem.* **1994**, *51*, 161−172.

(33) (a) Reactants: To define the orientation of the catecholate anion we used the Mg atom and the side chain N_ϵ and C_ϵ atoms of Lys144 as MM anchor points (dummy atoms during QM optimization). Referring to the atom labeling of Figure 2, the fixed *Z*-matrix coordinates are as follows: $r(O_2,Mg)$ = 1.93 Å, $\varphi(O_2,Mg,N_\epsilon)$ = 72.1°, $\varphi(C_2,O_2,Mg)$ = 156.8°, $\tau(O_2,Mg,N_\epsilon, C_\epsilon)$ = −29.0°, $\tau(C_2,O_2,Mg,N_\epsilon)$ = 31.4°, $\tau(C_1,C_2,O_2,Mg)$ = 102.3°. Analogously, the trimethylsulfonium cation is related to the positions of the C_4', O_4', and C_1' atoms of the adenosyl fragment of SAM by the following: $r(C_5',C_4')$ = 1.52 Å, $\varphi(C_5',C_4',O_4')$ = 107.6°, $\varphi(S,C_5',C_4')$ = 115.2°, $\tau(C_5',C_4',O_4',C_1')$ = 149.0°, $\tau(S,C_5',C_4',O_4')$ = 168.3°, $\tau(C_\gamma,S,C_5',C_4')$ = −85.9°. (b) Products: $r(O_2,Mg)$ = 2.95 Å, $\varphi(O_2,Mg,N_\epsilon)$ = 64.6°, $\varphi(C_2,O_2,Mg)$ = 150.5°, $\tau(O_2,Mg,N_\epsilon,C_\epsilon)$ = −38.0°, $\tau(C_2,O_2,Mg,N_\epsilon)$ = 42.9°, $\tau(C_1,C_2,O_2,Mg)$ = 90.6°, $r(C_5',C_4')$ = 1.53 Å, $\varphi(C_5',C_4',O_4')$ = 108.4°, $\varphi(S,C_5',C_4')$ = 116.0°, $\tau(C_5',C_4',O_4',C_1')$ = 148.3°, $\tau(S,C_5',C_4',O_4')$ = 172.1°, $\tau(C_\gamma,S,C_5',C_4')$ = −83.7°.

(34) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
(35) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(36) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623−11627.
(37) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; John Wiley: New York, 1986.
(38) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796−6806.
(39) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*; Gaussian, Inc.: Pittsburgh, PA, 1998.

during the charge annihilation reaction, were calculated individually for each HF/6-31+G*-optimized geometry along the reaction coordinate using the RESP procedure. All link hydrogen atoms were forced to have zero charge to avoid electrostatic artifacts, and the sum of the charges for the catecholate anion and the $(CH_3)_3S^+$ cation was constrained to be neutral. This latter constraint required refitting the charges of the MM atoms of the enzymatic cofactor with its sum being zero. We derived a single set of RESP charges for the MM atoms, from simultaneous fitting to both SAM and SAH, which was maintained throughout all free energy simulations. All RESP charges of the QM fragments and of SAM and SAH are tabulated in the Appendices 1 and 2 (Supporting Information). Our standard force field parametrization was used for the QM model system extended by a few harmonic restraints on key internal coordinates to keep the geometry close to the structures of the QM optimization.[40]

The free energy of interaction, ΔG_{int}, was calculated both in the protein environment and in aqueous solution using the thermodynamic integration (TI) method.[41] For the protein simulation, we reinserted the model fragments into the enzyme and equilibrated the system for 60 ps. Initial TI calculations with different nonbonded cutoffs showed strong fluctuations of ΔG_{int} with predominant electrostatic contributions. This strong dependence is due to the use of a residue-based cutoff in AMBER, which results in the two ionic ligands of very different size feeling a different electrostatic environment during the charge annihilation reaction. We used two distinct electrostatic treatments to overcome this problem. In the first series of simulations, no nonbonded cutoff was used for the QM model atoms while a cutoff of 16 Å was applied to all MM atoms (model I). This approach allows us to include all interactions between the protein residues and the atoms whose charges change during the free energy simulation. Our second approach is based on the assumption that charged residues far away from the active site cancel out their effects on the reaction energetics. Here, we neutralized all charged residues outside of ~10 Å of the active site by reassigning small charges to their respective CO_2^-, NH_3^+, or $NHC(NH_2)_2^+$ groups, while conserving overall neutrality of the system (model II). Equilibration and free energy calculations in model II were done with a nonbonded cutoff of 12 Å. It is interesting to note that COMT has an unusual charge distribution between 12 and 20 Å around the reaction center in that the overall charge in this sphere decreases steeply from $-2e^-$ at 12 Å to $-7e^-$ at 14 Å and increases again from 16 Å to $-1e^-$ at 20 Å. Our reduced-charge model allows us to investigate the importance of this charge distribution on the reaction energetics and to test the reproducibility of the free energies of model I with a computationally less intensive simulation.

To calculate the free energy of interaction for the solution reaction we placed the QM fragments in a periodic box of water molecules. Initial heating and equilibration of this system for 60 ps resulted in box dimensions of $\sim 38 \times 38 \times 39$ Å^3. We found the nonbonded cutoff dependence of the free energy to be converged to within 1 kcal/mol at 14 Å and performed TI calculations of this system for different simulation times using the *NpT* ensemble ($p = 1$ atm, $T = 300$ K). In both protein and water calculations, no interactions between QM atoms were included in the free energy of interaction because they are already contained in the ΔE_{QM} energy term (eq 1).

5. Cratic Free Energy Contribution. In the solution reaction, we need to quantitatively assess the free energy terms of bringing the reactants together in a Michaelis complex-like geometry in order to directly compare the reaction energetics with the enzyme environment. This cratic free energy term, ΔG_{cratic}, has contributions from the changes of solvation energy, solvent and solute entropy, and from the work to change the cavity in the water bulk when forming the reactant complex. We estimated ΔG_{cratic} by two different methods.

In the first approach, we used Tomasi's polarized continuum model (PCM),[16] which combines a Hartree−Fock self-consistent reaction field calculation[16] of the electric polarization of the solvent with separate estimates of the cavity[42] and dispersion−repulsion[43,44] contributions to the free energy. We used this model to calculate the free energy of complex formation, $\Delta G^*_{complex}$, between the $(CH_3)_3S^+$ cation and the catecholate anion in continuum water. However, this calculation, as indicated by the asterisk, does not yield the full free energy change in complex formation since it does not take into account the change in solute entropy. This is because the solvation effect is calculated for single monomer and complex structures, each. We approximated ΔS_{solute} from classical statistical formulas and normal-mode analysis of the two monomers and the dimer in the gas phase each and calculated the cratic free energy as

$$\Delta G_{cratic} = \Delta G^*_{complex} - T\Delta S_{solute} \quad (4)$$

$\Delta G^*_{complex}$ was obtained by minimizing the molecules in the gas phase and subsequent single point calculations in the reaction field of water, both at the HF/6-31+G* level. Spheres with united atom Hartree−Fock (UAHF) radii[45] were used to build the cavity, and the polarization charge was normalized by means of additional effective charge, distributed according to the solute electronic density (option ICOMP = 4 in the PCM version implemented in Gaussian 98).[39] Since basis set superposition errors are likely to be small for this kind of reaction, we did not include any correction terms into our treatment.

In our second approach, we estimated ΔG_{cratic} from a classical PMF calculation in explicit water. Using molecular dynamics, this yields the free energy change as a function of the separation of the two monomers. The particle mesh Ewald (PME) technique[46,47] was used in the PMF calculation to accurately describe the electrostatic interactions of the two molecular ions upon dimer formation. We performed these simulations in a periodic box of size $\sim 41 \times 40 \times 46$ Å^3 filled with 2380 water molecules and used a cutoff radius of 9 Å for the separation of short-range and long-range interactions. Cubic spline interpolation on a charge grid of $45 \times 45 \times 48$ Å^3 and a direct sum tolerance of 10^{-5} were applied. After equilibration of the two monomers for 90 ps, with the distance between O_2 and C_ϵ, r_{OC}, fixed at 8 Å, we calculated the potential of mean force for approaching the two ions from $r_{OC} =$ 8−3 Å. Additional harmonic restraints of 40 kcal/(mol radian^2) were applied to the internal coordinates $\varphi(O_2,C_\epsilon,S) = 159.1°$ and $\varphi(C_2,O_2,C_\epsilon) = 86.0°$ to ensure stable dynamical behavior during the PMF run.

6. MD Simulations of Meta- and Para-Substituted Catechols. To investigate the factors affecting the stereoselectivity of O-methylation in substituted catechols, several MD calculations with different substitutions R in meta or para position (4 and 5 in Scheme 1, respectively) were performed. The simulations included the L-enantiomer of the amino acid 3,4-dihydroxyphenylalanine or dopa ($R = CH_2CH(NH_3^+)$-CO_2^-) as well as a catechol acid ($R = CH_2CO_2^-$). Since those substitutions have little effect on the acidity of the catecholic hydroxyl groups, we assumed the same protonation states as in catechol (positions 1 and 2 in Scheme 1) for all compounds. Starting structures of these molecules in the enzyme were prepared by replacing catechol in the Michaelis complex with the substituted compound, minimization of the new ligand, and equilibration of the solvated enzyme−substrate

(40) Internal coordinate harmonic restraints with equilibrium values from the QM optimization and force constants of 100 kcal/(mol Å^2) and 40 kcal/(mol radian^2) for bonds and angles, respectively, were defined for the following intermolecular coordinates (atom labeling, see Figure 2). (a) Reactants: $r(O_2,C_\epsilon) = 3.03$ Å, $\varphi(O_2,C_\epsilon,S) = 159.1°$, $\varphi(C_2,O_2,C_\epsilon) = 86.0°$. (b) Products: $r(C_\epsilon,S) = 3.28$ Å, $\varphi(O_2,C_\epsilon,S) = 158.3°$, $\varphi(C_\epsilon,S,C_\gamma) = 94.7°$, $\varphi(C_\epsilon,S,C_5') = 107.8°$. Parameters for these coordinates at intermediate points of the reaction path were obtained from linear interpolation of the corresponding values between reactants and products. We additionally changed the standard force field equilibrium angles of $\varphi(H_\epsilon,C_\epsilon,S)$, $\varphi(H_\epsilon,C_\epsilon,O_2)$, and $\varphi(H_\epsilon,C_\epsilon,H_\epsilon)$ to their QM-optimized values for each point of the reaction path to reproduce the inversion of the transferring methyl group.

(41) Beveridge, D. L.; DiCapua, F. M. *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 431−492.

(42) Cossi, M.; Tomasi, J.; Cammi, R. *Int. J. Quantum Chem.* **1995**, 695−702.

(43) Floris, F. M.; Tomasi, J.; Ahuir, J. L. P. *J. Comput. Chem.* **1991**, *12*, 784−791.

(44) Cossi, M.; Mennucci, B.; Cammi, R. *J. Comput. Chem.* **1996**, *17*, 57−73.

(45) Barone, V.; Cossi, M.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3210−3221.

(46) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092.

(47) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577−8593.

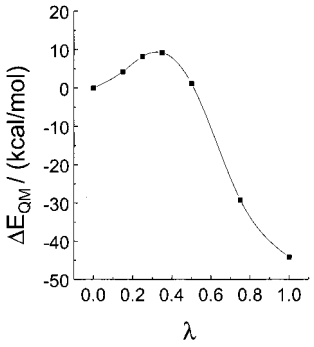


Figure 3. HF/6-31+G*//MP2/aug-cc-pVDZ-calculated energy profile for the reaction between catecholate anion and trimethylsulfonium cation. The points are interpolated by a cubic spline fit. The constraints used during the geometry optimization are explained in the text of the Methods section. Snapshots of the complex along the reaction path variable λ are depicted in Figure 2.

complex for 30 ps. Subsequently, we collected MD trajectories for 100−200 ps for analysis.

III. Results and Discussion

A. Enzymatic Catalysis. 1. MM Equilibration To Define Reactant/Product States for the QM−FE Calculation. Starting from the X-ray structure of the COMT−inhibitor complex, we find a significant conformational change in the Mg coordination when simulating the ligand catechol instead of the inhibitor 3,5-dinitrocatechol. This new coordination, which is formed during the initial heating phase to $T = 300$ K, and which remains stable at this temperature for at least 500 ps, is schematically depicted in Figure 1b. Interestingly, catechol prefers a monodentate rather than a bidentate coordination with the magnesium ion. The vacant metal coordination site is occupied with the second carboxylate oxygen of Asp141, initiating a positional shift of the Mg-coordinated water molecule. The new conformation of catechol in the enzyme, forming a triad of hydrogen bonds to (a) the intramolecular O^-, (b) the COO^- group of Glu199, and (c) a water molecule of the solvation shell, represents a physically reasonable arrangement. Important consequences of this finding will be discussed later in this section.

2. QM Optimization. The calculation of the ab initio energy differences, ΔE_{QM}, between structures along the reaction path variable λ yields an energy profile as is illustrated in Figure 3 for the HF/6-31+G*//MP2/aug-cc-pVDZ method. We find a transition state located at $\lambda = 0.35$ with a barrier of approximately 10 kcal/mol. The overall QM reaction energy is strongly negative in vacuo because the charges of the two reactants get annihilated in a gas phase environment. The screening of these charges, which are effective in the protein and in solution, will be taken into account in our study by the free energy of interaction ΔG_{int}. Numerical values of the QM reaction profile for different methods and basis sets are compared in Table 1. Differences exist in the height of the TS barrier, with the HF and MP2 methods both yielding similar and by 6−8 kcal/mol higher values than the DFT/B3LYP calculation. The overall reaction energy is less exothermic for the methods that include electron correlation. Previously, Radom et al. investigated S_N2 reactions at saturated carbon atoms and found that the B3LYP functional significantly underestimated the TS barrier compared to experiment.[48] In contrast, their MP2 calculations, using similar basis sets as our study, yielded satisfactory agreement with measurements. These results suggest that our MP2-calculated reaction barriers are more reasonable than the DFT/B3LYP values.

Table 1. Relative Energies ΔE_{QM} (kcal/mol) along the Reaction Path Variable λ for Different Quantum Mechanical Methods and Basis Sets[a]

	λ				
method	0.25	0.35	0.50	0.75	1.0
HF/6-31+G*	9.55	10.58	2.94	−32.58	−59.23
HF/aug-cc-pVDZ	9.34	10.85	3.98	−30.63	−55.96
MP2/6-31+G*	10.35	11.20	2.84	−28.97	−46.19
MP2/aug-cc-pVDZ	8.15	9.20	1.15	−29.20	−44.19
B3LYP/6-31+G*	3.43	2.37	−4.43	−30.41	−48.29

[a] All HF and MP2 energies are based on geometry optimizations at the HF/6-31+G* level. Absolute energies at $\lambda = 0$ are as follows: $E_{HF/6-31+G*} = -896.088278$ hartree, $E_{HF/aug-cc-pVDZ} = -896.168248$ hartree, $E_{MP2/6-31+G*} = -897.757340$ hartree, $E_{MP2/aug-cc-pVDZ} = -897.991095$ hartree, $E_{B3LYP/6-31+G*} = -899.943644$ hartree. The constraints used during the geometry optimization are explained in the text of the Methods section.

It is interesting to compare our QM energy profile, which takes into account the enzymatic preorganization of the substrates, with the unconstrained gas phase calculations of the same model system by Zheng and Bruice.[5] Using the similar 6-31+G** basis set, they computed TS barriers of 22.3, 21.1, and 15.3 kcal/mol for the HF, MP2, and DFT/B3LYP methods, respectively. These barrier heights are 10−12 kcal/mol larger than our values of Table 1, mainly because the minimal energy configuration of the unconstrained reactant complex has the deprotonated O atom of catechol symmetrically hydrogen bonded to all three methyl groups of the $(CH_3)_3S^+$ ion. This configuration is not productive in the enzyme and is naturally excluded from our QM optimization through the MM constraints. Test calculations showed that the energy difference between the unconstrained optimized dimer and a minimized dimer in which constraints on two intermolecular angular coordinates (as defined below) ensured that only a single methyl group was pointing to the O^- of catechol is 14.7 and 13.4 kcal/mol for the HF/6-31+G* and MP2/6-31+G* methods, respectively. Applying this correction to the values of Zheng yields TS barriers for the unconstrained gas phase calculation which are ∼3 kcal/mol lower than our enzyme-constrained optimization.

Comparing the transition state geometry, we find bond distances of 2.33 Å between C_ϵ and S and 2.42 Å between C_ϵ and O_2, which are longer by ∼0.2 Å for each bond compared to the unconstrained gas phase calculation of Zheng and Bruice. In their study, they were able to accurately reproduce experimental kinetic isotope effects for the COMT-catalyzed methylation of 3,4-dihydroxyacetophenone,[1,4] which suggests that our TS geometry may be somewhat too loose (calculated, $k_{H/D} = 0.80$, $k_{C12/C13} = 1.06$; experimental, $k_{H/D} = 0.83 \pm 0.05$, $k_{C12/C13} = 1.09 \pm 0.05$). It should be kept in mind that our approximation to the transition state, having the advantage of reduced computational expense, is based on a simple, linear interpolation scheme between the MM-optimized reactant− and product−enzyme complexes. Rather than performing a more elaborate, multidimensional TS search, we estimated the energetic effect of a tighter transition state on our energy profile. To this end, we calculated the energy difference between the unconstrained, tight TS geometry of Zheng and the relaxed reactant complex, in which only the soft intermolecular coordinates $\varphi(S,C_\epsilon,O_2)$ and $\tau(S,C_\epsilon,O_2,C_2)$ were constrained to their X-ray values of 173.1° and −7.4°, respectively. This constraint keeps the two moieties in an enzyme-like configuration with a single methyl

(48) Glukhovtsev, M. N.; Bach, R. D.; Pross, A.; Radom, L. *Chem. Phys. Lett.* **1996**, *260*, 558−564.

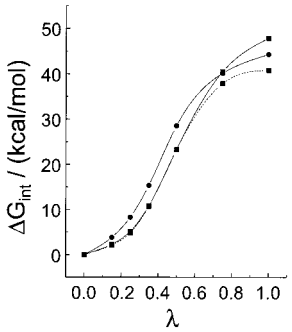


Figure 4. Free energy of interaction in the protein (squares) and in water (circles) as a function of the reaction path variable λ. The two protein simulations, model I (solid line) and model II (dashed line), differ in the treatment of electrostatic interactions with residues far away from the active site (see text in the Methods section for details). The simulation time is 180 ps.

group pointing toward the O^- of the catecholate. At the HF/6-31+G* and HF/6-31+G*//MP2/aug-cc-pVDZ levels, we obtained barrier heights of $\Delta E = 6.85$ and 6.36 kcal/mol, respectively. These barriers are 3−4 kcal/mol lower than the corresponding values in Table 1 and suggest that a tighter transition state provides a few kcal/mol more stabilization energy.

Due to the constraints imposed in our QM optimization, it is not physically meaningful to include vibrational corrections to the reaction profile. However, a normal-mode analysis of the unconstrained reactant and TS complex, as was used by Bruice, yields a small increase of the gas phase TS barrier, due to zero point energy, thermal, and entropic effects, by ≈1 kcal/mol.

3. Free Energy Calculations in the Enzyme and in Aqueous Solution. Figure 4 shows the free energy of interaction, ΔG_{int}, along the reaction path for the protein and water environment. In both environments the free energy increases as a function of λ because stabilizing electrostatic interactions between the charged reactants and the surroundings vanish during the charge annihilation reaction. As was found from an analysis of the free energy components, the change in the electrostatic interactions contributes more than 95% to the total ΔG_{int}. Comparing the two charge models for the protein simulation, we find a very similar behavior up to $\lambda = 0.5$ and a 7 kcal/mol smaller overall free energy change for the model in which charged residues far away from the active site are neutralized (model II). This suggests that the unusual charge distribution of the protein in the range of 12−20 Å from the active site, which is taken into account in model I, may help to release the product back into solution after the reaction took place. As one would expect, the water environment destabilizes the early phase of the reaction relative to the enzyme by up to 5 kcal/mol. This destabilization gets inverted around $\lambda = 0.75$ for the enzyme model I, and the product appears to be less well bound in the protein than in water. To test for convergence of our ΔG_{int} values, we performed several TI calculations of different lengths for both protein and water environments. As can be seen from Table 2, little hysteresis was found from forward and backward simulations and all numerical values are converged to better than 2 kcal/mol.

Adding the relative energies of ΔE_{QM} and ΔG_{int}, according to eq 1, we obtain the free energy profile of the transmethylation reaction in COMT and for the same reaction in water (Figure 5). The calculated transition state is located around $\lambda = 0.5$, and the barrier height in the enzyme is 24.5 kcal/mol. This is 6.5 kcal/mol higher than the experimental $\Delta G^{\ddagger} = 18$ kcal/mol, which was derived from kinetic measurements of the turnover

Table 2. Free Energy of Interaction, ΔG_{int}, of Different Simulation Length (100 Windows) in COMT and in Aqueous Solution[a]

	simulation time (ps)			
	90	150	180	300
protein (model I)	48.8/47.5	47.8/—	48.4/47.3	48.6/—
protein (model II)			41.6/40.0	
solution	41.6/42.6	42.7/—	43.2/45.4	42.7/42.6

[a] Numerical values of ΔG_{int} (kcal/mol) are from forward and reverse runs. The two charge models for the protein simulation are explained in the text of the Methods section.

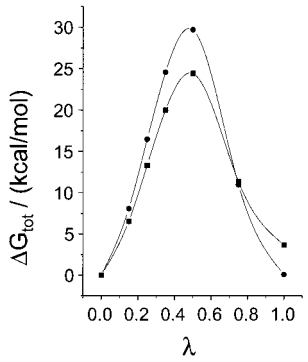


Figure 5. Free energy profile for the transmethylation reaction in COMT (squares), using charge model I, and for the analogous reaction in aqueous solution (circles).

number (24 min^{-1}) for the reaction between rat liver COMT, SAM, and the catechol derivative 3,4-dihydroxybenzoic acid.[49] If we allow for a slightly tighter transition state, as was seen in the QM optimization with relaxed constraints, our best estimate for $\Delta G^{\ddagger}$ (MP2/aug-cc-pVDZ) is 21.7 kcal/mol, which is in better agreement with experiment. Clearly, this comparison with experiment is only valid if the yet unmeasured rate constant of the unsubstituted catechol is similar to that of 3,4-dihydroxybenzoic acid. For the analogous reaction in solution, we calculated a TS barrier of 29.5 kcal/mol, hence a transition state stabilization of 5 kcal/mol in COMT compared to the reaction in water. The overall free energy of forming the product complex is slightly endergonic in the protein ($\Delta G_{tot} = +3.5$ kcal/mol) and almost thermoneutral in solution.

4. Cratic Free Energy Contribution in Water. We evaluated the cratic free energy contribution in aqueous solution, which comprises the important changes in solvation and entropy involved in complex formation, using quantum mechanical PCM, classical PMF, and normal-mode analysis calculations. The results of two approximations to the cratic free energy are summarized in Table 3. In the first method (see eq 4), our PCM calculations yielded a positive $\Delta G^*_{complex}$ of 3.9 kcal/mol, indicating that the loss in hydration energy when forming the complex overcomes the attractive interaction energy between the two approaching molecular ions. Our correction for the loss in solute entropy is unfavorable by an additional 9.4 kcal/mol. We are aware that this estimation of ΔS_{solute}, which is based on a classical treatment of translation and rotation, and on a normal-mode analysis of the vibrational degrees of freedom, at a single, minimized geometry is approximate. In our previous work on trypsin,[9] we did more elaborate calculations on this, allowing some flexibility in the "reactive geometry", and obtained 5−6 kcal/mol for orienting two molecules in a reactive geometry. Clearly, even if we used the previously calculated entropy correction, we compute a $\Delta G^*_{complex} - T\Delta S_{solute}$ of the order

(49) Schultz, E.; Nissinen, E. *Biochem. Pharmacol.* **1989**, *38*, 3953−3956.

Table 3. Cratic Free Energy Contributions (kcal/mol) for the Solution Reaction[a]

	method 1	method 2
$\Delta G^*_{complex}$	3.9	4−5[b]
$-T\Delta S_{solute}$	9.4[c]	
ΔG_{cratic}	13.3	

[a] $\Delta G^*_{complex}$ was calculated using quantum mechanical PCM (method 1) and classical PMF (method 2) calculations. Classical statistical formulas and a normal mode analysis were used to calculate the solute entropy correction. [b] PMF free energies from forward/reverse runs for different simulation times: 6.5/3.7 (75 ps), 4.2/4.6 (150 ps), 5.0 (300 ps). For the meaning of this value, see the paragraph about the cratic free energy in the Results and Discussion section. [c] Components of the change in solute entropy: translational (+11.1 kcal/mol), rotational (+6.3 kcal/mol), vibrational (−8.0 kcal/mol).

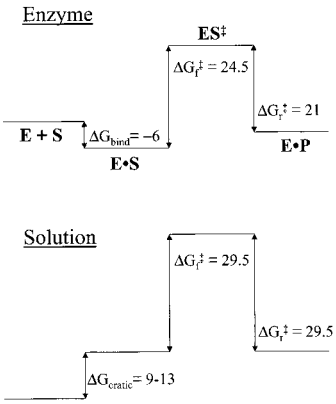


Figure 6. Comparison of schematic free energy diagrams for the reaction in the enzyme and in water. E·S and E·P are the enzyme−substrate and enzyme−product complexes, respectively, and ES‡ denotes the transition state. The enzymatic free energy of binding, ΔG_{bind}, has been inferred from the experimental Michaelis constant, K_M.[51,52] The subscripts f and r stand for forward and reverse, respectively. All energy units are kilocalories/mol.

of 10 kcal/mol, which is a considerable contribution to the $\Delta G^‡$ for the solution reaction.

In our second method, PMF simulations of different simulation lengths yielded a free energy of complex formation of 4−5 kcal/mol. Because of the dynamical nature of PMF simulations, part of the change in solute entropy is included in this free energy value. However, it is difficult to assess this contribution quantitatively due to the several restraints that we had to impose in our simulation (see Methods section), consequently reducing the number of solute degrees of freedom that were sampled during MD. Hence, we did not attempt to estimate an entropy correction factor for method 2.

We should mention that Zheng and Bruice calculated the solvation effect for forming the same ion pair using the continuum PM3-SM3 method[50] and obtained 7.7 kcal/mol.[5] They did not correct for the change in solute entropy, presumably because their method is parametrized to reproduce experimental hydration free energies. Moreover, they used the unconstrained minimum structure of the dimer which differs considerably from our enzyme-constrained structure in that all three methyl groups of the cation point toward the O^- of the anion. Consequently their dimer should have a lower hydration free energy relative to the monomers than ours.

5. Comparison of Reaction Energetics in COMT and in Aqueous Solution. We summarize our QM−FE calculations in the schematic free energy diagram of Figure 6. The major

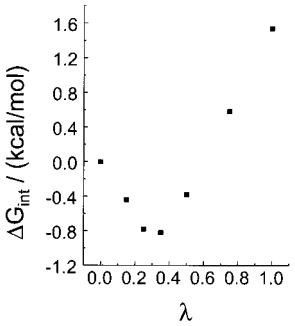


Figure 7. Free energy of interaction between the residue Met40 and the QM model system as a function of the reaction path variable λ.

difference between the reaction in COMT and in aqueous solution consists of the formation of the Michaelis complex (E·S). In solution, a considerable free energy price of 9−13 kcal/mol has to be paid to desolvate the reactants and to bring them together in a reactive geometry. Unfavorable free energy contributions due to solvation changes and due to entropy loss upon complex formation occur also in the enzyme, but are overcome by attractive binding interactions leading to an experimental free energy of binding of $\Delta G_{bind} = -(5-6)$ kcal/mol.[51,52] The second reason why the reaction in COMT is faster than in solution lies in a lowering of the transition state barrier in the enzyme by 5 kcal/mol compared to water.

It is interesting to compare the enzymatic catalysis in COMT with our recent QM−FE calculations on the amide hydrolysis in trypsin and aqueous solution.[9] In the trypsin study, we found a stabilization free energy of 6 kcal/mol for our transition state model in the enzyme compared to solution. Although both enzymes reveal a quantitatively similar lowering of the TS barrier compared to solution, they differ drastically in the way they achieve this. In trypsin, an additional charge on the amide is created in the transition state, which is stabilized by the preoriented dipoles of Asp102 and of the residues of the oxyanion hole. In contrast, parts of the charges of the reactants get annihilated in the transition state of COMT. This is clearly a more difficult reaction type for the enzyme to catalyze because no obvious extremum in substrate−enzyme electrostatic interactions along the reaction path exists, as in the trypsin case. To understand how COMT achieves a transition state stabilization of 5 kcal/mol compared to the same reaction in water, we analyzed the individual contributions to the free energy of the 20 enzymatic residues closest to the reaction center. We found the residue Met40, whose free energy of interaction with the reactants is displayed in Figure 7, to have its most favorable interactions along the reaction path at the transition state. In the enzyme−reactant complex, the sulfur atom of Met40 has van der Waals contact with the transferring methyl unit. This interaction is maintained throughout the reaction although the CH_3^+ moiety is moved by 1.6 Å toward the catecholate anion. Interestingly, the RESP-derived charges for the methyl group along the reaction path are $+0.32e^-$, $+0.49e^-$, and $+0.27e^-$ for $\lambda = 0$, 0.5, and 1, respectively. It is the positive charge of the transferring methyl group, which is at a maximum at the TS, which is stabilized by the negatively charged sulfur atom of Met40. Although the stabilization of the TS by Met40 is rather small, due to the large polarizability of sulfur, this interaction could likely be more favorable if explicit polarization were included in our calculation. However, it is unlikely to be increased much beyond 1.5 kcal/mol in total. Rate constant

(50) Cramer, C. J.; Truhlar, D. G. *J. Comput.-Aided Mol. Des.* **1992**, *6*, 629−666.

(51) Lotta, T.; Vidgren, J.; Tilgmann, C.; Ulmanen, I.; Melen, K.; Julkunen, I.; Taskinen, J. *Biochemistry* **1995**, *34*, 4202−4210.

(52) Borges, N.; VieiraCoelho, M. A.; Parada, A.; SoaresdaSilva, P. *J. Enzyme Inhib.* **1998**, *13*, 473−483.

measurements of a COMT mutant in which Met40 is replaced by a residue without the sulfur functionality would help quantify the importance of Met40 on COMT catalysis. In the subclass of SAM dependent methyltransferases for which X-ray structures are available, COMT seems to be unique in the presence of a Met residue in the active site. Although many members of this class, like the DNA methyltransferases, share a common protein structure,[53] they differ considerably from COMT in their active site residues. Only glycine *N*-methyltransferase has a methionine residue, Met90, close to the reaction center, with $r_{S-C} = 6.95$ Å (COMT: $r_{S-C} = 3.57$ Å).[54]

Apart from the TS stabilization effect of Met40, what are other factors important for the catalytic action of COMT? To have a lower $\Delta G^{\ddagger}$ than in water, it is advantageous for the enzyme not to stabilize the positive and negative charges of the two substrates SAM and catechol too much. As is apparent in the X-ray structure, several hydrophobic residues, such as Trp38 and Pro174 for catechol, and Trp143, Tyr68, and Met40 for SAM, are used to bind the reactants. Furthermore, the interaction between the reactive O of catechol and the Mg^{2+} ion appears to be of great importance for the course of the enzymatic reaction. In the reactant stage, the strong interaction between these two oppositely charged atoms helps to overcome the cratic free energy penalties and to bind catechol to the enzyme. This interaction is greatly reduced after methylation of the oxygen atom, promoting the release of the product back into solution. Interestingly, the RESP-derived partial charge on the reactive O of catechol, which is $-0.90e^-$ for the reactant and $-0.41e^-$ for the product, changes only by $\sim 20\%$ up to the transition state. This indicates that the Mg···O interaction destabilizes the transition state compared to the reactant complex much less than expected.

It can be seen in Figure 7 that the apparent free energy of activation is 24.5 kcal/mol in the enzyme and 38.5−42.5 kcal/mol in solution. This free energy difference of $\Delta\Delta G^{\ddagger} = 14-18$ kcal/mol corresponds to a rate enhancement of $\sim 10^{10}-10^{13}$ in the enzyme compared to the solution reaction. Our calculation is in very good agreement with predictions of Mihel et al.,[8] who estimated an enzymatic acceleration factor of 10^{11} based on experimental rate constants for the reaction between $C_6H_5O^-$ and $(CH_3)_3S^+$ in water.[55] Zheng and Bruice argued that the reaction rate between the catecholate ion and the trimethylsulfonium ion in water is expected to be slower than that of the phenolate ion because of its lower basic strength and the positive Brønstedt constant for nucleophilic attack, which would slightly increase the rate enhancement.

6. MD Calculations on 3,5-Dinitrocatechol. The metal coordination change found in the MM equilibration of the enzyme−catechol system motivated us to perform MD simulations for the COMT−inhibitor complex for which the X-ray crystal structure is known. The two nitro groups are located at positions 3 and 5 of the catechol ring, as indicated in a schematic view of a part of the active site in Figure 8. Since the pK_a of the related 2,4-dinitrophenol in water is 4.1,[56] it is very likely that the oxygen O_2 which is ortho and para to the two nitro groups, respectively, is deprotonated in the active site. The electron-withdrawing effect of the NO_2 groups is less pronounced at the position of O_1 [pK_a (3,5-dinitrophenol) = 6.7 vs pK_a (phenol) = 10.0],[56] and structure b in Figure 8 appears to be the most reasonable protonation state of 3,5-dinitrocatechol on the basis of the free pK_a values in water. However, it is known that pK_a's in an enzymatic environment can differ considerably from those in aqueous solution, and especially the strong electric field of the Mg^{2+} ion might increase the acid strength of the catecholic hydroxyls. As a consequence, we performed MD calculations of 100 ps for the two protonation states of DNC, which are displayed together with the RESP-derived charges in Figure 8b,c. Our simulations showed that the singly deprotonated DNC in Figure 8b, with Glu199 anionic, changes its Mg coordination from bidentate to monodentate, being replaced by the second carboxylate O of Asp141, in analogy to catechol. In contrast, dianionic 3,5-dinitrocatechol, with Glu199 protonated, keeps its X-ray orientation. This suggests that in the COMT−inhibitor complex, the proton is preferentially located at Glu199 rather than at O_1 of DNC.

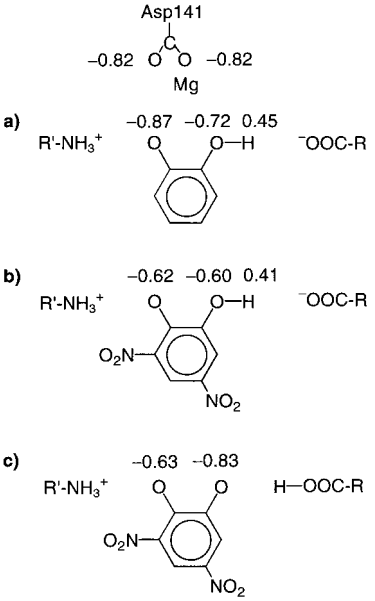


Figure 8. Schematic view of a part of the active site of COMT with RESP-derived atomic partial charges for the ligands (a) catechol, (b) 3,5-dinitrocatechol (monoanionic), and (c) 3,5-dinitrocatechol (dianionic). R = Glu199 and R′ = Lys144.

B. Regioselectivity in Monosubstituted Catechols. Experimental product ratio studies on various monosubstituted catechols showed that polar substituents R favor the O-methylation in COMT to occur at the meta rather than para catecholic hydroxyl group relative to R.[11,12] For example, dopamine ($R = CH_2CH_2NH_3^+$) or L- and D-dopa ($R = CH_2CH(NH_3^+)CO_2^-$) possess meta:para ratios of 6.9, 19.8, and 3.4, respectively. In contrast, nonpolar substitutions, such as $R = CH_2CH_3$, show almost no preference for either of the two O-methylation sites. No X-ray structure of COMT was available at that time, and it was suggested that a hydrophobic region is present in the active site which repels polar substrates in such a way that meta O-methylation is preferred.[11] Recently, Lau and Bruice supported this explanation by showing in MD simulations of the COMT−catechol complex that Trp38 and Tyr200 form a hydrophobic pocket in the para substitution site.[6] However, in their simulations they explicitly bonded all six atoms, including the two oxygens of catechol, to the first coordination sphere of Mg^{2+}. In this way, they forced unsubstituted catechol to have the same coordination geometry as 3,5-dinitrocatechol. Since we found, in our unconstrained MD simulation of the COMT−catechol complex, a coordination change at the Mg ion,

(53) Cheng, X. D. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 293−318.

(54) Fu, Z. J.; Hu, Y. B.; Konishi, K.; Takata, Y.; Ogawa, H.; Gomi, T.; Fujioka, M.; Takusagawa, F. *Biochemistry* **1996**, *35*, 11985−11993.

(55) Swain, C. G.; Taylor, L. J. *J. Am. Chem. Soc.* **1962**, *84*, 2456−2457.

(56) Rived, F.; Roses, M.; Bosch, E. *Anal. Chim. Acta* **1998**, *374*, 309−324.

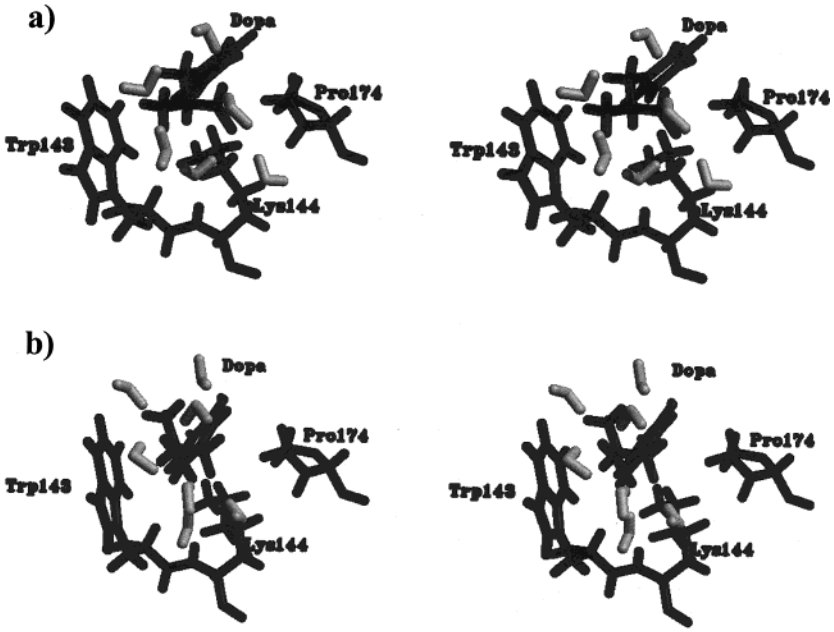


Figure 9. Stereoview of snapshots of the equilibrated L-dopa structure. The amino acid substituent is located meta (a) and para (b) to the reactive oxygen atom of catechol. All residues (black) and water molecules (gray) within 4.5 Å of C_α of L-dopa are shown.

involving a change of $\Delta\varphi(Mg,O_2,C_2) \sim 45°$ compared to the X-ray structure of the dinitro inhibitor, we performed explicit MD simulations for selected monosubstituted catechols.

We first investigated L-dopa, which has the highest yet measured meta:para ratio of 19.8. Parts a and b of Figure 9 show stereoviews of the equilibrated complex structures with the substituent located at the positions meta and para to the reactive oxygen of catechol, respectively. With the L-dopa side chain in the meta position (Figure 9a), we can identify several favorable interactions with enzymatic residues. First, the NH_3^+ group of L-dopa sits nicely above the aromatic plane of Trp143, forming an attractive cation$-\pi$ interaction. Experiments[57] and calculations[58] on model systems showed these interactions to be unusually strong, and they were found to be important for the structure of proteins[59] and in enzyme$-$ligand complexes.[60,61] In our simulation, we presumably underestimate the cation$-\pi$ interaction because of the neglect of nonadditive effects.[58] Inspecting the other substituents of the chiral carbon atom of the L-dopa side chain, we find the CO_2^- group entirely solvated and the single H atom pointing into a small pocket formed by the alkyl side chain of Lys144. In contrast to the substitution in the meta position, the side chain of the para-substituted L-dopa is mostly solvated with the single hydrogen now pointing toward the tryptophan residue (Figure 9b). The distance between the NH_3^+ group and the center of the six-membered ring of Trp143 is 3.0 and 3.8 Å for the meta and para structures, respectively. Although only from a structural point of view, we can clearly rationalize the observed product ratio in L-dopa from better interactions between the substrate and COMT when the amino acid side chain is located in the meta rather than the para position. It would be interesting to see if the same meta:para ratio would be observed for L-dopa for a W143A mutant of COMT.

While a strong cation$-\pi$ interaction, which results in binding of the substrate in a way that O-methylation is preferred, allows us to explain the product ratio for cationic substrates, different interactions must be important for anionic ligands. To this end, we performed two MD simulations for the catechol acid with $R = CH_2CO_2^-$ (meta:para ratio = 7.0) located either meta or para to the reactive catechol oxygen. Analysis of 200 ps MD trajectories showed the para side chain fully hydrated while a hydrogen bond between the side chain CO_2^- group and the NH peptide backbone is formed in the meta compound, with $r(O\cdot\cdot HN) < 3.0$ Å during 25% of the simulation time. Although an intramolecular hydrogen bond is entropically favored because of the release of bound water molecules, it is difficult to estimate the free energy effect and its impact on the product ratio.

Despite the fact that our analysis is only qualitative, we have shown in MD simulations that an alternate explanation for the preference for meta O-methylation in catechols with polar substituents can be given when a nonbonded approach for the magnesium coordination is chosen. Our earlier results for 3,5-dinitrocatechol imply that a theoretical explanation for the meta:para ratio is sensitive to the modeled protonation state of catechol.

IV. Summary and Conclusions

We have presented an application of the QM$-$FE method to study methyl transfer in the enzyme catechol O-methyltransferase and, for comparison, this reaction in aqueous solution. Our calculated $\Delta G^\ddagger$ in the enzyme (21$-$24 kcal/mol) is in reasonable agreement with experiment (18 kcal/mol), and our estimates for the relative $\Delta\Delta G^\ddagger$ of the enzyme and solution reactions (14$-$18 kcal/mol) are consistent with experiment (15 kcal/mol). Our computed energy ranges reflect the uncertainty in estimating the reaction path geometries and in determining the cratic free energy in solution, which we believe are our two most important approximations. For the first time, this study includes the effects of the full protein and water environments on the reaction free energies in COMT in an explicit manner.

Using only those fragments in the QM part where bonding is changing, the results are in reasonable agreement with experiment. In future studies, it might be interesting to increase

(57) Dougherty, D. A.; Stauffer, D. A. *Science* **1990**, *250*, 1558$-$1560.
(58) Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 4177$-$4178.
(59) Mitchell, J. B. O.; Nandi, C. L.; McDonald, I. K.; Thornton, J. M.; Price, S. L. *J. Mol. Biol.* **1994**, *239*, 315$-$331.
(60) Sussman, J. L.; Harel, M.; Frolow, F.; Oefner, C.; Goldman, A.; Toker, L.; Silman, I. *Science* **1991**, *253*, 872$-$879.
(61) Basran, J.; Mewies, M.; Mathews, F. S.; Scrutton, N. S. *Biochemistry* **1997**, *36*, 1989$-$1998.

the size of the region that is described quantum mechanically. However, the next larger QM system would have to include at least the protein residues depicted in Figure 1a, which, when suitably truncated, would increase the number of heavy atoms from currently 12 to 37. Because of the large computational burden, this is currently very expensive at the ab initio level presented.

In contrast to trypsin, previously studied by QM−FE,[9] both the enzyme and solution reactions have a significantly larger $\Delta G^{\ddagger}$ than one would find for methyl transfer in vacuo (∼5−10 kcal/mol). This is simply because, in COMT, charge is annihilated rather than created in the transition state. The enzyme has the challenging task of binding the substrate effectively while not stabilizing it electrostatically too much. It achieves this by using a Mg^{2+} to help bind and orient the catecholate anion, with a significant number of nonpolar/hydrophobic residues to bind the rest of the substrate. In this way, the enzyme increases the transition state barrier, relative to in vacuo, by ∼5 kcal/mol less than found in aqueous solution.

As found previously in trypsin, a cratic free energy difference of ∼10 kcal/mol between enzyme and solution reaction is the biggest difference between these reactions. Although our estimates of this term are approximate, this component is likely to be the single most general and important component that differentiates enzyme-catalyzed and solution reactions. This analysis rests on the *assumption* we make that the enzyme absorbs the cratic free energy in its net free energy of binding and thus binds the substrate much less tightly than it would were it not for the free energy cost of orienting the substrate for catalysis to achieve a NAC (near attack conformation, using the term of Bruice and co-workers[7]).

Nonetheless, the general importance of the cratic term enables one to rationalize significant enzyme catalysis even for radical reactions, where electrostatic stabilization of the transition state is expected to be small. It may also rationalize why biomimetic models of enzymes usually achieve only modest rate enhancements, since they do not have the extended binding sites needed to achieve a favorable (negative) ΔG_{bind} for substrate binding in a productive geometry.

This is not to underestimate the electrostatic contribution to catalysis. Nonetheless, as suggested by calculations by Honig and others,[62−64] the absolute ΔG_{bind} for ligand binding to proteins, due to electrostatic interactions, is often unfavorable or, at best, modestly favorable because of the competition with ligand electrostatic solvation, with most of the absolute driving force for binding coming from van der Waals/hydrophobic contributions (Exceptions to this, of course, may occur from metal ions in metalloenzymes). Clearly, the electrostatic interactions can be enhanced in transition states (e.g., Asp102 and the oxyanion hole in the serine proteases) and are, of course, a key element in *design*, since, even when the electrostatic contribution to transition state stabilization might be unfavorable in an absolute sense, it can be made relatively less so with selective design of key functional groups.[65,66] As noted above, COMT is able to achieve less strong and, thus less catalytically "unfavorable" electrostatic interactions of the ground state relative to water and thus, to more effectively catalyze the reaction than in solution.

We have also presented qualitative MD simulations to rationalize the significant meta OH selectivity of some catechols, most dramatically L-dopa, where the meta OH is methylated ∼20 times more frequently than the para OH. We suggest that the enzyme accomplishes this selectivity because the NH_3^+ group in L-dopa can form a particularly stable cation−π interaction with Trp143, while keeping its CO_2^- solvated, only in an enzyme−substrate binding orientation that leads to O-methylation in the meta position. Binding of catechol to COMT with the para OH pointing toward the methyl transfer unit does not allow for as favorable interactions with its ionic groups.

Our modeling of the catechol bound to COMT involves a purely noncovalent model leading to a monodentate coordination of the catechol ring oxygens with the Mg^{2+}, whereas a previous study[5] enforced a bivalent coordination as was found in the X-ray structure of COMT with a 3,5-dinitrocatechol inhibitor. Our calculations suggest that manipulating Trp143 to Ala would be an interesting test of our proposed rationale for the reactive stereoselectivity of L-dopa.

Acknowledgment. B.K. gratefully acknowledges support from the German Academic Exchange Service (DAAD) through a research scholarship. P.A.K. is pleased to acknowledge research support through the NIH (GM-29072), supercomputer support from the NSF supercomputer centers, and graphics support from the UCSF computer graphics lab, T. Ferrin, P.I. (RR-1081).

Supporting Information Available: Appendices 1 and 2 giving atomic partial charges for the QM atoms and the SAM and SAH moieties (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA992218V

(62) Miyamoto, S.; Kollman, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 8402−8406.
(63) Misra, V. K.; Hecht, J. L.; Yang, A. S.; Honig, B. *Biophys. J.* **1998**, *75*, 2262−2273.
(64) Singh, S. B.; Kollman, P. A. *J. Am. Chem. Soc.* **1999**, *121*, 3267−3271.
(65) Lee, L.-P.; Tidor, B. *J. Chem. Phys.* **1997**, *106*, 8681−8690.
(66) Kangas, E.; Tidor, B. *J. Chem. Phys.* **1998**, *109*, 7522−7545.