Article

Exploring the Dependence of QM/MM Calculations of Enzyme Catalysis on the Size of the QM Region

Garima Jindal and Arieh Warshel*

Department of Chemistry, University of Southern California, 3620 McClintock Avenue, Los Angeles, California 90089, United States

ABSTRACT: Although QM/MM calculations are the primary current tool for modeling enzymatic reactions, the reliability of such calculations can be limited by the size of the QM region. Thus, we examine in this work the dependence of QM/MM calculations on the size of the QM region, using the reaction of catechol-O-methyl transferase (COMT) as a test case. Our study focuses on the effect of adding residues to the QM region on the activation free energy, obtained with extensive QM/MM sampling. It is found that the sensitivity of the activation barrier to the size of the QM is rather limited, while the dependence of the reaction free energy is somewhat larger. Of course, the results depend on the inclusion of the first solvation shell in the QM regions. For example, the inclusion of the Mg^{2+} ion can change the activation barrier due to charge transfer effects. However, such effects can easily be included in semiempirical approaches by proper parametrization. Overall, we establish that QM/MM calculations of activation barriers of enzymatic reactions are not highly sensitive to the size of the QM region, beyond the immediate region that describes the reacting atoms.
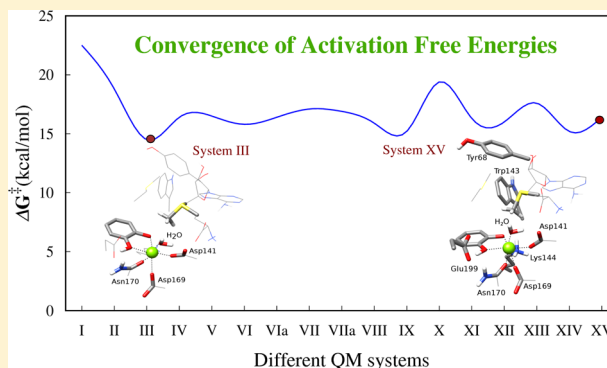
I. INTRODUCTION

QM/MM methods (for reviews, see, e.g., refs 1−7) have been used extensively in studies of enzymatic reactions. In order to obtain reliable results, it is crucial to perform extensive sampling[8] and also to use a reliable QM model for either calibrating a semiempirical model (e.g., the EVB model[9]) or actual determination of ab initio (ai) QM (ai)/MM free energy surfaces.[10] In considering the reliability of such methods, one of the obvious issues is the size of the QM region. The size problem is frequently attributed to the treatment of the boundaries between the QM and MM regions, but this issue is not so serious if one uses the same models in the reference water reaction and in the enzyme active site. Nevertheless, one would like to know the limitations imposed by the size of the QM region. Here we noted that the great advances in computer power and the use of specialized computers (e.g., GPU) allows one to use very large QM regions and the results of the corresponding calculations can be used to imply that it is essential to use large QM regions to obtain reliable results.[11,12] Thus, it is important to conduct careful analysis of the size dependence of QM/MM calculations.

The issue of the size dependence in QM clusters has been a subject of significant interest.[13−21] Similarly, the size dependence of the QM region in QM/MM models has been the subject of recent studies. Some studies explored the size dependence of QM/MM calculations of spectroscopic properties.[22,23] More relevant works (from the perspective of this paper) examined the energy difference in enzymatic reactions.[24−28] However, as much as enzymatic reactions are concerned, the most important issue is the reliability of the calculated activation free energy, that should reflect an extensive averaging, that has not been performed in most of the previous studies. One exception is the study of ref 26, which evaluated the free energy of proton transfer (PT) in lysozyme for QM regions of different sizes. This work found that for a QM with a radius of less than 6 Å we can have errors of about 3 kcal/mol. Here we note that PT reactions involve a rather small change in charge separation and one can expect a larger dependence on the size of the QM region in reactions that involve a larger change in charge distribution. Thus, we focus in this work on the dependence of the calculated activation free energy on the size of the QM region in the reaction of the enzyme COMT (catechol-O-methyl transferase) that catalyzes the transfer of a methyl group from SAM (S-adenosyl-L-methionine) to a catecholate ion via an S_N2 type reaction, as shown in Scheme 1.

The issue of the size of the QM region in COMT was highlighted recently in a study[11,12] which implied that the inclusion of large protein residues (up to 500 atoms) in the QM region is needed in order to obtain convergent results. While ref 11 only looked at the convergence of ground state properties including the donor−acceptor distance, ref 12 examined the trend in activation energy, but this was done by evaluating the single point energy change upon moving to the transition state (rather than the activation free energy). Nevertheless, such results may discourage further use of QM/

Scheme 1. Schematic Representation of the Transfer of a Methyl Group from SAM to Catecholate Catalyzed by COMT^a
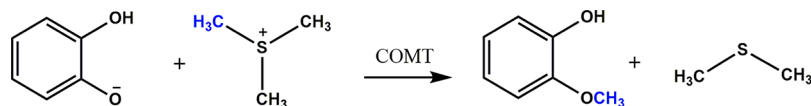
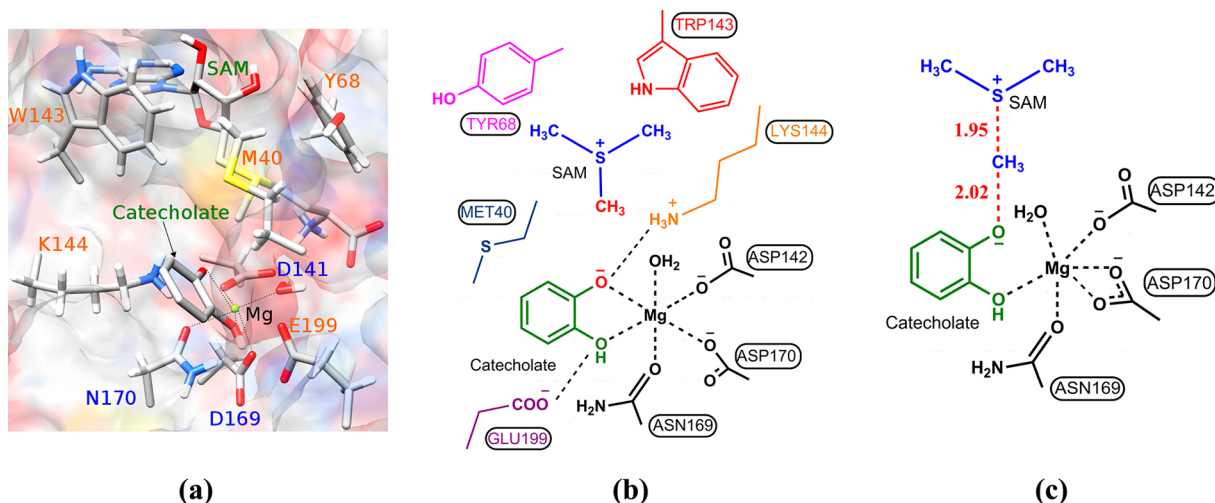^{a}The SAM unit is truncated to the methyl group for clarification.

Figure 1. (a) The active site of COMT. (b) A schematic representation of the active site of COMT enzyme with some important residues. (c) The transition state for the methyl transfer reaction between SAM and catecholate ion for a representative case (distances are in Å).

MM calculations with medium QM size and thus require further systematic studies that will be performed in this work.

In choosing COMT as our benchmark, we note that this system is an important biological system. It belongs to the general class of methyltransferases that plays an important role in human physiology and several diseases such as cancer, and genetic diseases. COMT degrades catecholamines such as dopamine and has thereby attracted a major interest in its probable role in dopamine related pathologies, in particular Parkinson's disease. V158M COMT polymorphism is speculated to be responsible for the violent behavior of schizophrenic patients.[29] It is therefore imperative to study such enzymes and understand their function at the molecular level. It should also be noted that the COMT catalyzed decomposition of dopamine follows a parallel pathway to the reaction catalyzed by MAO B, which was recently studied using the EVB method.[30] Hopefully, the present study can also be useful by shedding additional light on the results and in some cases limitations of previous QM/MM studies of COMT.[31−36]

Our study focuses on the dependence of the activation free energy of the reaction of COMT on the size of the QM region. For very small sizes where, for example, the Mg^{2+} is treated classically, we find an obvious deviation from the converging results (although this problem can be fixed by semiempirical approaches with proper parametrization of the QM/MM interactions). For larger systems, we find a reasonable convergence, with some deviations in specialized cases.

II. COMPUTATIONAL METHODS

Our study used the semiempirical PM6 method (evaluated by the MOPAC2009 package[37,38]) to treat the QM region, whereas the rest of the protein was treated classically, using the ENZYMIX force field.[39] The QM/MM interface between the QM and MM surfaces was implemented through the corresponding module of the MOLARIS package.[39] The initial coordinates for the QM/MM dynamics were taken from the crystal structure of the human, soluble form of COMT (s-COMT).[40] The first step for each system involved a relaxation by slowly heating the protein from 30 to 300 K for at least 45 ps. The subsequent MD simulations involved time steps of 1 fs at 300 K. The simulation was divided into 31 frames with each frame simulated for 10 ps. The reaction coordinate, $\xi = d(S−C) − (O−C)$, was constrained using a force constant of 100 kcal/mol Å^2 and varied from −1.0 to 2.2 Å. The PMF was obtained by combining the individual simulation windows, using the weighted histogram analysis method (WHAM) approach.[41] Further calculations for a few systems (**I**, **II**, and **III**) were done for a longer time, by dividing the simulation into 51 frames and running each window for 30 ps. This yielded similar results to the shorter runs, and therefore, for all other systems, we only carried out shorter runs.

The boundary conditions for the simulations were introduced by immersing the protein in an 18 Å sphere of water molecules using the surface-constraint all-atom solvent (SCAAS) boundary conditions.[39] The local reaction field method (LRF) which is similar to having no cutoff for electrostatic interactions is used to treat the long-range effects.[39] The geometric center of the reacting system was taken as the center of the simulation sphere. The positions of all atoms beyond 20 Å from the center of the systems were fixed as found in the initial crystal structure, and their nonbonded interactions with the atoms within the simulation sphere were turned off. To determine the protonation state, we used our coarse grained (CG) model.[42] Subsequently, all residues that lie within 12 Å from the center of the SAM and catecholate anion were explicitly ionized during the simulation. A large charge−charge dielectric constant was used to treat the effect of the more distanced ionizable residues. Three different starting configurations were generated for the simulations by

taking structures from a long trajectory at a time difference of 45 ps. The resulting structures were used in the QM/MM simulation, and the final result was taken as the average of the three runs.

Our study used the PM6 semiempirical method that allows for extensive sampling. Obviously, more correct barriers should be obtained with high level QM methods, but the issue here is the size dependence of the barrier and not obtaining an accurate barrier. Of course, we hope that the size dependence of the PM6 is related to the size dependence in more rigorous models.

III. RESULTS

Our study focused on the activation barrier for the methyl transfer step, which is depicted in Scheme 1. The different residues considered as a part of the QM region are shown in Figure 1. The active site consists of Mg^{2+} ion coordinated to the catecholate ion, two aspartate residues, asparagine residue, and a water molecule. The transition state with the Mg^{2+} ion and the ligands in the primary shell is shown in Figure 1c for a representative case.

In order to explore the size dependence of the QM region, we evaluated the activation free energies for regions of different sizes. The smallest region consisted of the substrates alone, while the largest region included the metal and seven residues, which were selected on the basis of the distance from the substrates. The first QM region investigated (system **I**) is described in Figure 2. This system includes only the two
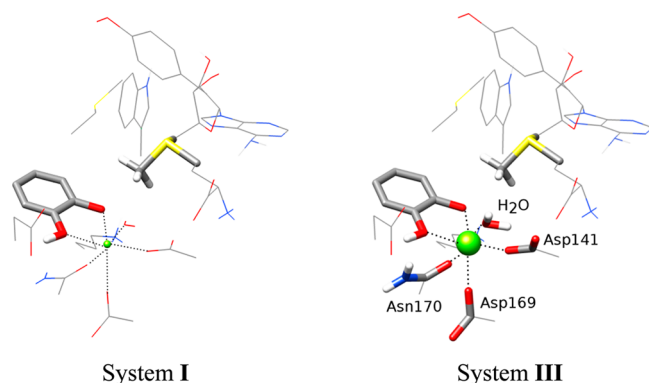

Figure 2. System **I** with only the substrates (catecholate and SAM) as the QM region and system **III** with the substrates, Mg^{2+} ion, and the surrounding ligands as the QM region. The QM part is shown as sticks and part of the MM region as wires (the rest of the protein is not shown for clarity).

substrates, SAM and the catecholate ion, and consists of 26 atoms in the QM region including the linked atoms. The Mg^{2+} metal ion was treated classically as a point charge with a repulsive van der Waals center. The activation free energy obtained was 22.5 kcal/mol.

The next system considered (system **II**) also included only the substrates in the QM region. However, the Mg^{2+} was treated using the dummy atom approach developed by Åqvist and Warshel.[43] This approach, which involves the inclusion of six dummy atoms around the metal center, as depicted in Figure 3, was shown to yield improved solvation energies. Thus, the Mg^{2+} is represented as MD_6^{2+}. The activation free energy obtained using this approach was 18.8 kcal/mol.
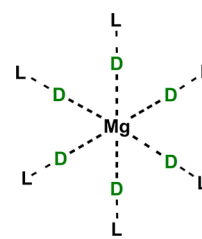

Figure 3. MD_6^{2+} model complex used instead of Mg^{2+} in the classical simulation of system **II**, where the Ds represent the dummy atoms and the Ls are the general ligands around Mg^{2+}.

Subsequently, the Mg^{2+} ion and the surrounding ligands (H_2O, Asp169, Asp141, and Asn170) were included in the QM region (system **III** shown in Figure 2). The resulting free energy barrier is 14.5 kcal/mol. It should be noted that only a part of the residues as shown in Figure 2 was included in the QM region. Inclusion of the complete residue did not alter the result significantly, and hence, in all further calculations, we used chopped residues as the ligands coordinated to the Mg^{2+} ion. The change in the barrier upon moving to system **III** reflects QM charge transfer effects, which are discussed below.

Next, we considered the effect of individual amino acid residues near the active site by including them separately in the QM region, while still retaining the residues in system **III** (Figure 4). The different residues considered are Glu199 (system **IV**), Lys144 (system **V**), Tyr68 (system **VI**), Trp143 (system **VII**), and Met40 (system **VIII**). Lys144 and Glu199 form H-bonds with the catecholate ion and result in a charged QM system. The free energy barriers obtained for systems **IV** and **V** were 16.4 and 16.5 kcal/mol, respectively. During the simulations, the hydroxyl proton at the catecholate shuttles between the O atom of the catecholate and the O atom of the Glu residue. As will be discussed below, this corresponds to another reaction path and should be considered accordingly. The inclusion of Tyr 68 (which is considered to be an important residue as its mutation decreases significantly the activity) yields a barrier of 15.8 kcal/mol. The inclusion of Met40 and Trp143 yields activation free energies of 15.9 and 17.1 kcal/mol, respectively. We also included in the QM region Tyr147 (**VIa**) and W38 (**VIIa**), which are approximately 10 Å away from the center of the substrates. This yields barriers of 16.4 and 16.9 kcal/mol, which are similar to the barriers in other systems where the residues are much closer to the active site.

After studying the effect of individual residues, we included both Lys144 and Glu199 (system **IX**) together with the Mg^{2+} metal and its ligands (Figure 4). These two residues lie closest to the substrates forming weak H-bonding interactions and are therefore included in most of the systems described hereafter. The resulting activation free energy is 15.2 kcal/mol. Further, we included three residues together keeping the Mg and its surrounding ligands, Lys144 and Glu199, constant. In system **XIII**, we studied the effect of another residue, Trp43, on the activation barrier. The different systems created (Figure 5) are (a) Lys144, Glu199, Trp143 (system **X**); (b) Lys144, Glu199, Tyr68 (system **XI**); (c) Met40, Tyr68, Trp143 (system **XII**); and (d) Trp 143, Trp43, Tyr68 (system **XIII**). In the final and largest QM systems, we included four residues together, namely, Lys144, Glu199, Trp143, and Met40, resulting in system **XIV** and residues Lys144, Glu199, Trp143, and Tyr68 that generated system **XV**.
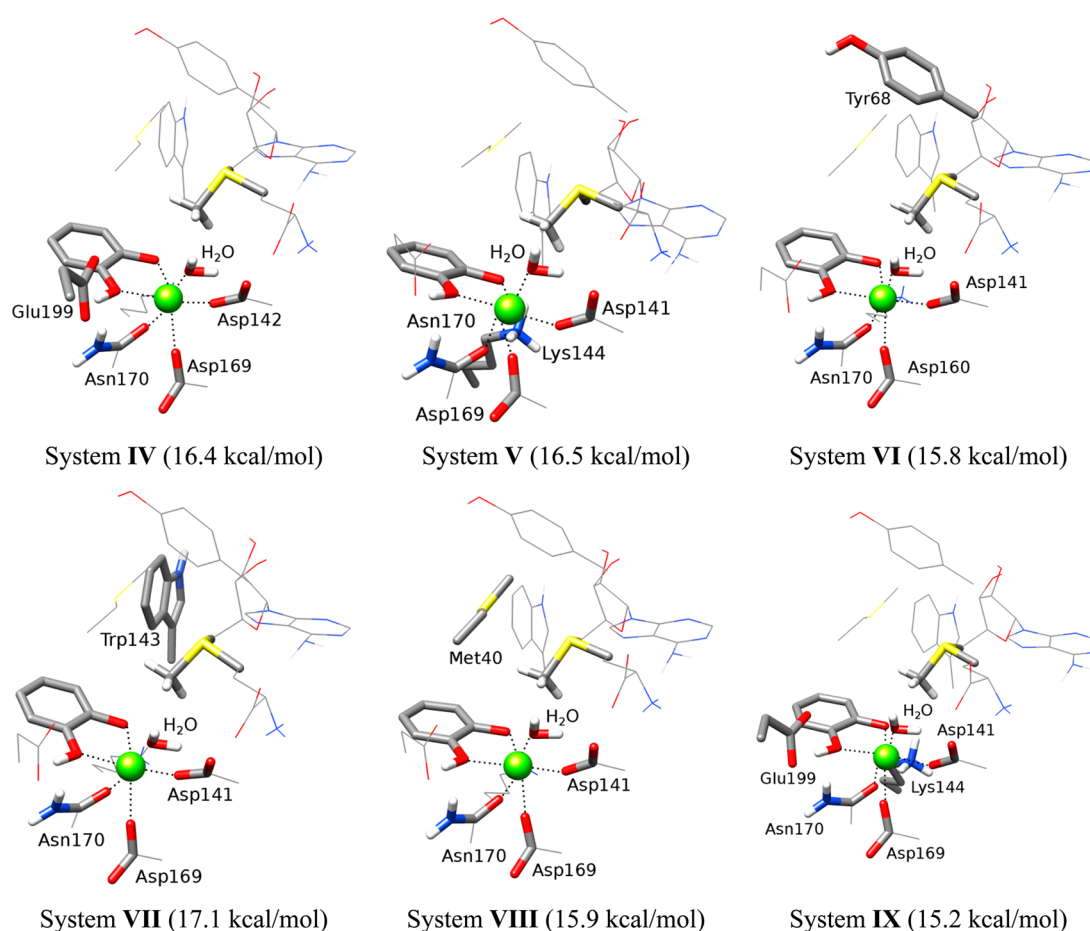
Figure 4. Systems involving different key residues in the QM region. The activation free energies in kcal/mol are provided in parentheses.

In total, 15 different QM systems have been considered and it has been found that the free energy barrier lies in the range 14.5−19.4 kcal/mol. The largest system including seven residues does not show much improvement over the system containing the Mg^{2+} and ligands alone. On the basis of our results, we propose that Mg^{2+} and its surrounding ligands (system **III**) are sufficient to be included in the QM region. Furthermore, we also carried out QM/MM studies for the Y68A mutant, where the experimental free energy barrier for the mutant is 20.0 kcal/mol whereas that of the wild type is 18.4 kcal/mol. In this mutant case, we only investigated systems **III** and **I**. It was found that, while the calculated free energy for wild type with system **I** is 22.5 kcal/mol, the corresponding barrier for Y68A is 26.8 kcal/mol (leading to an increase of 4.3 kcal/mol in the barrier). In the case of system **III**, the barriers for the wt and the mutant are 14.5 and 16.6 kcal/mol, respectively (leading to a 2.1 kcal/mol increase in barrier). It can be seen that the trend on increase in barrier upon mutation is reproduced by both systems (where system **I** gives a significant overestimate).

The results of the above study are summarized in Figures 6 and 7 as well as Table 1. The analysis of the results can start with the obvious finding that cases that do not include the ion and its ligands in the QM systems (cases **I** and **II**) present a major approximation, since the corresponding QM treatment does involve major charge transfer (CT). As we pointed out, the inclusion of the Mg^{2+} ion in the QM region requires inclusion of the ligands of this ion, and such a treatment can be very expensive when one uses a reasonable level of the QM

treatment. Fortunately, treating the Mg^{2+} ion classically and using the proper parametrization can give reasonable results.

Another case, which requires special attention, is the inclusion of Glu199 in the QM system (in addition to system **III**) without including additional residues. In this case, we find a partial proton transfer (PT) from the catecholate to Glu199. However, we actually have here a trivial case of another reaction mechanism (namely, a PT to Glu199). Of course, one should examine whether we have a real PT or an artifact of the use of the PM6 method. In fact, the problem disappears already when we add Lys144 or additional residues to the QM region. We also find that adding Lys144 alone to system **III** also leads to a deviation in the reaction free energy. Now again we have a residue that is involved in the reaction, where a PT from the catechol to the Lys can be considered as an initial step of the reaction[31] (see also below). Here a partial PT for the back reaction can appear in small QM regions, but again adding more residues leads to a disappearance of the deviations.

We like to note that, although Lys144 is the most likely candidate to accept the proton from catechol leading to the formation of catecholate ion, it is still unclear whether the proton stays on Lys144 or is transferred to some other residue or water molecules. The close proximity of this residue to Mg^{2+} (3.2 Å) in the crystal structure results in this ambiguity, as it may lead to the low pK_a of these residues. In fact, the finding that the Lys144Ala mutant leads to a minor reduction in k_{cat} indicates that the Lys does not play a major role in the rate acceleration.[44] Thus, although our previous EVB calculations gave a somewhat larger catalytic effect with neutral Lys144,[31]

System **X** (19.4 kcal/mol)

System **XI** (16.3 kcal/mol)
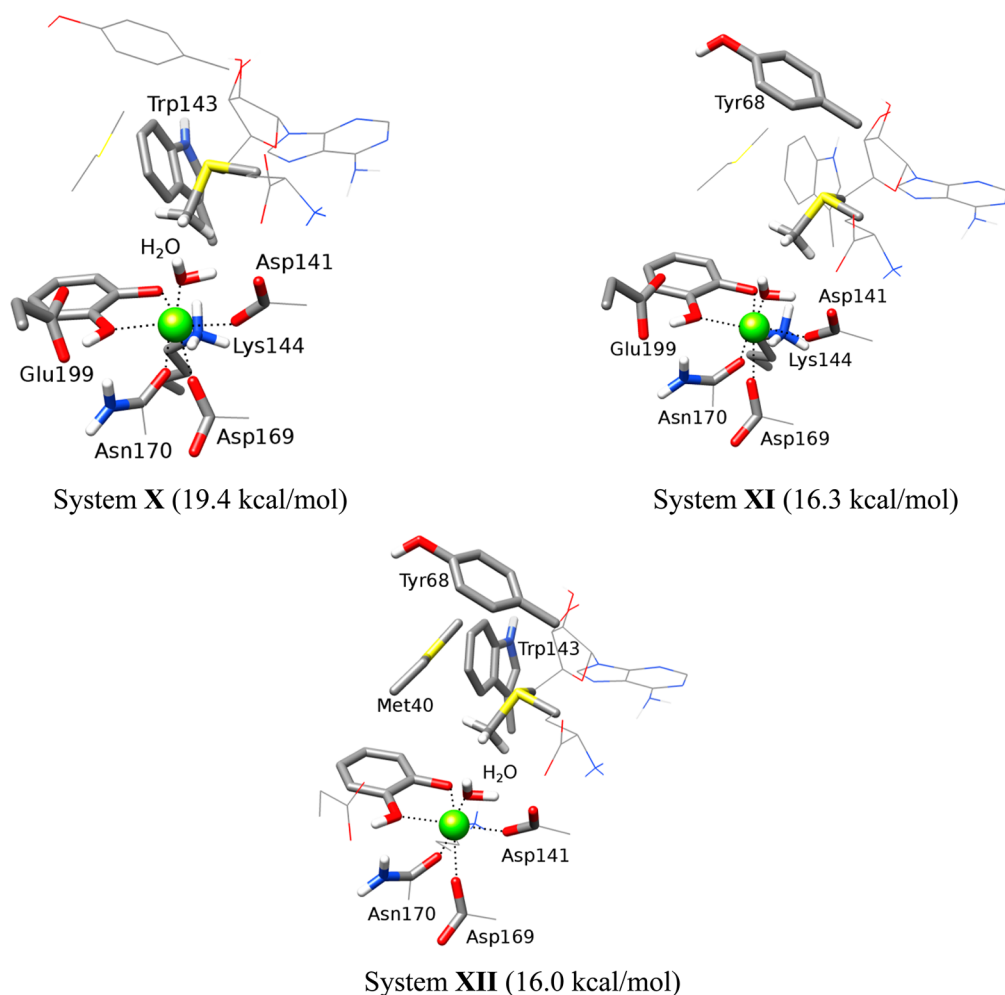
System **XII** (16.0 kcal/mol)

Figure 5. Systems that include in the QM region three residues along with the Mg^{2+} and its surrounding ligands. The activation free energies in kcal/mol are provided in parentheses.
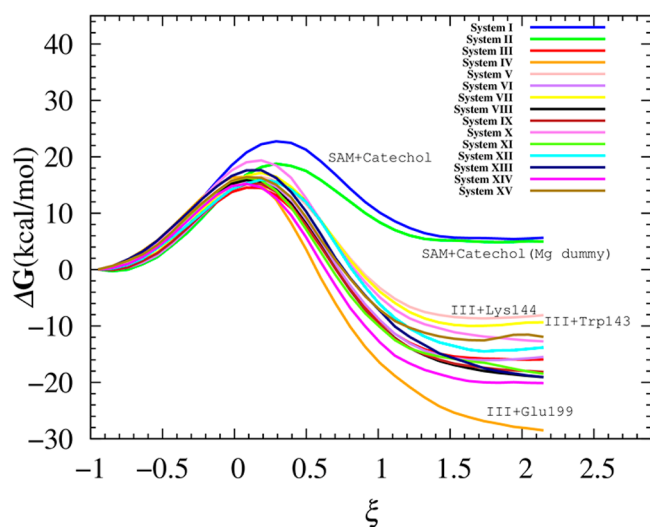

Figure 6. The PMFs obtained for different systems for the PM6/MM potential, which have been calculated using the WHAM procedure.

Figure 7. Activation free energies (kcal/mol) of different QM systems. All systems include the substrates: SAM and catechol (shown in wires). System **IV** and beyond include the residues of system **III**.

we kept this residue ionized here (in the cases when it was in the classical region). Regardless of the actual protonation state of the Lys residue during the nucleophilic attack, our analysis of the size dependence is still valid for the model system used.
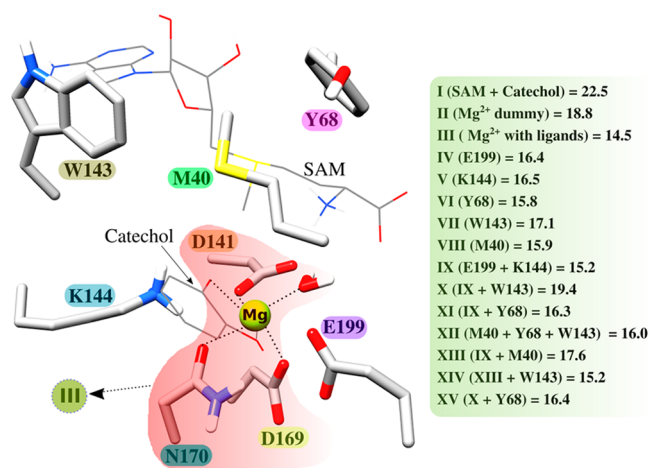
Overall, we note that, while the convergence of the activation free energies is encouraging, the convergence for the reaction free energies is not as good as that of the activation free energies (Table 1), for single residues that can be involved in side reactions like PT. Nevertheless, once other QM residues

Table 1. Activation Free Energy and Reaction Free Energy (kcal/mol) for Different Systems

QM system	$\Delta G^{\ddagger}$ (kcal/mol)	ΔG (kcal/mol)
I (substrates alone)	22.5	5.6
II (Mg^{2+} using dummy)	18.8	5.0
III (Mg^{2+} and ligands)	14.5	−15.9
IV (III + Glu199)	16.4	−28.5
V (III + Lys144)	16.5	−8.7
VI (III+ Tyr68)	15.8	−16.0
VII (III + Trp143)	17.1	−10.0
VIII (III + Met40)	15.9	−19.1
IX (III + Glu199 + Lys144)	15.2	−18.1
X (III + Lys144 + Glu199 + Trp143)	19.4	−12.7
XI (III + Lys144 + Glu199 + Tyr68)	16.3	−18.5
XII (III + Met40 + Tyr68 + Trp143)	16.0	−17.2
XIII (III + Trp43 + Trp143 + Tyr68)	17.6	−19.1
XIV (III + Lys144 + Glu199 + Trp143 + Met40)	15.2	−20.1
XV (III + Lys144 + Glu199 + Trp143 + Tyr68)	16.4	−12.6

were included with the problematic residues, we obtained converging results. This feature was found for Glu199, Lys144, and, for less clear reason, Trp143, where in all cases adding more residues leads to the disappearance of this problem.

IV. DISCUSSION

QM/MM approaches provide a major tool for evaluating the energetics of chemical reactions in complex molecular systems. However, some workers can assume that the results depend drastically on the size of the QM system and thus presumably the QM/MM idea is inherently a major approximation. Apparently, what is meant by "approximation" is not uniformly agreed upon. For example, some may assume that an accurate result can be obtained only when the protein and the solvent are represented by a high level ab initio approach. However, obtaining reliable results also requires a major sampling and satisfying the requirement of both a proper sampling and a reliable QM approach is an extremely challenging task that would be hard to accomplish in the near future. Thus, one must judge carefully which approximation is more effective. There are ways to use large QM systems, including our constraint DFT (CDFT) approach,[45,46] or the use of a fast processor or specialized computers that allow running with large QM systems.[11] However, the requirement of extensive sampling makes it crucial to exploit QM/MM approaches with a limited size of the QM region, while exploring systematically what are the limitations of such strategies. This work explored the approximations associated with the use of a relatively small size for the QM region, by a systematic analysis of this issue while modeling the free energy profile of the reaction of COMT. As seen from Figure 8, the convergence of the activation free energy is quite reasonable (except for models **I** and **II**, as discussed below, the reasons for those deviations are obvious).

The deviations between the results with different sizes of the QM region are more pronounced when we compare the calculated exothermicity (Table 1). In analyzing the corresponding results, we should consider different situations. First, we note that the cases that do not include the ion and its ligands in the QM systems present a major approximation, since the corresponding QM treatment does involve charge transfer (CT). Thus, we do have here an obvious case where one should try to include the Mg^{2+} in the QM treatment. Fortunately, the CT effect can be captured by changing the charge on the Mg^{2+} and the ligands and by proper refinement
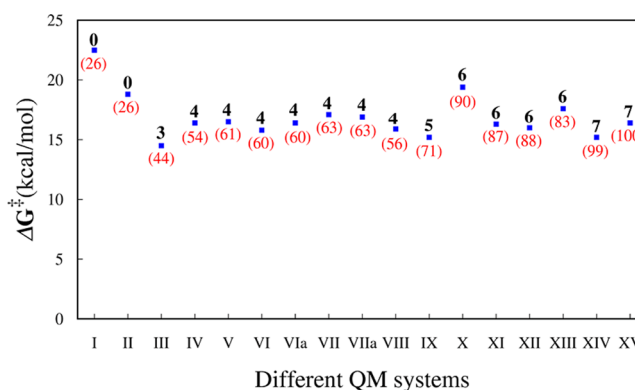

Figure 8. Convergence of the activation free energy as a function of the number of QM residues. The numbers in brackets are the number of QM atoms for the given system.

of the van der Waals parameter without the use of computationally expensive methods. The CT can also be represented by including the Mg^{2+} in an EVB treatment that can reproduce this effect. Furthermore, even without including the Mg^{2+} in the EVB region, we can easily reproduce the trend obtained with the Mg^{2+} in the QM region by a standard EVB treatment that includes parametrization of the reaction exothermicity.

As stated in section III, the convergence for the reaction free energy is not as good as the activation free energies. This is in particular true for single residues (i.e., Glu199 and Lys144) that can be involved in side reactions such as PT reactions. However, once such residues are surrounded by other QM residues, we obtain converging results. Such a behavior seems to reflect "instability" of the QM calculations, which do not follow a simple additive trend. This is reflected in the observation that the problem does not occur when the problematic residues are part of the MM system. It is also likely that the size dependence will be reduced if we use a polarizable force field.[47] It should also be noted that our reaction is not a conventional S$_N$2 reaction, as it involves two charged species that react to give a neutral product, thereby posing difficulties that arise in an S$_N$1 reaction. Thus, the solvation free energy of the reactive ionic species might result in a sensitivity of the reaction free energy.

Some workers tend to attribute the main source of the size dependence problem to the connection between the QM and

MM link atom.[48,49] However, a part of this problem is drastically reduced by using hybrid orbitals.[47,50−53] Furthermore, the problem is drastically reduced if one uses the same boundaries for reactions in water and reactions in enzyme as is done with the EVB method.

It is important to point out that even with some dependence on the QM size it is possible to explore quite reliably the catalytic and mutational effect. This issue has been demonstrated here with the wt to Y68A mutations in two different QM descriptions. Such a finding, for a mutant that is the subject of significant controversy,[31] is clearly important in view of the possible presumption that one must use large QM regions in exploring mutational effects.

Reference 12 obtained about 3 kcal/mol deviation from the final result for about 10 residues but then suggested that the convergence requires at least 30 residues (where the deviation from the converging result is about 2−3 kcal/mol). Ignoring the issue that the calculations are based on a single point energy evaluation, we like to address the implication that one needs to get around 2 kcal/mol convergence before exploring catalytic problems. Such an implication is unjustified for several reasons. First, obtaining 2 kcal/mol errors due to the size of the system still allows exploring enzyme catalysis, when the difference between the barrier in enzyme and in water is frequently on the order of 10 kcal/mol. Of course, it is also allowed (as shown here) to explore mutational effects. Second, no first principle approach can at present give errors of less than 2 kcal/mol. That is, changing the level of the quantum method can lead to several kcal/mol errors and adding a polarizable force field can change the results by a few kcal/mol. Obviously not performing sampling can lead to enormous errors.[54] Furthermore, very large errors can also be obtained from evaluation of electrostatic contributions without proper relaxation of the protein. Such calculations may produce an unphysical low dielectric that would not provide sufficient screening for charge−charge interaction.[55] Apparently, even proper free energy calculations still involve several kcal/mol errors due to incomplete sampling, which, for example, can miss water penetration effects.

The above-mentioned errors are inherent to current calculations, and once we are aware of such errors, we can better assess the effectiveness of different approximations. For example, the error can frequently be reduced by using a well-defined reference (e.g., the wt barrier in mutational studies and the reaction in water in exploring catalysis). As usual, the best way of exploring different approximations is to change the length of the run, the system size, and other factors and to explore the stability of the calculations as well as the relationship to the observed quantity.

It might also be pointed out here that there are various implications that the catalytic effect deduced from calculations with limited QM size (mainly the conclusion that the catalysis is associated with electrostatic effects) might be questionable and can be changed with larger QM sizes. Now, not only have EVB calculations seemed to give extremely stable conclusions, but also equally importantly, other proposed catalytic factors (e.g., dynamical effects, quantum effects, and more) have never been reproduced by consistent calculations regardless of the size used.[56]

V. CONCLUSION

Overall, the present study indicates that the size dependence is not as crucial as one may suspect. In particular, the size dependence is significantly smaller than the catalytic effect. Of course, the residues that are in direct contact with the reacting atoms can be involved in a major charge transfer or even in another reaction path. However, the QM description of residues that are not in direct contact do not change the activation barriers of chemical transformation in a major way.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: warshel@usc.edu. Phone: +1 (213)-740-7671.

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Warshel, A. Multiscale Modeling of Biological Functions: From Enzymes to Molecular Machines (Nobel Lecture). *Angew. Chem., Int. Ed.* **2014**, *53*, 10020−10031.

(2) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198−1229.

(3) Mulholland, A. J. Modelling Enzyme Reaction Mechanisms, Specificity and Catalysis. *Drug Discovery Today* **2005**, *10*, 1393−1402.

(4) Riccardi, D.; Schaefer, P.; Yang, Y.; Yu, H. B.; Ghosh, N.; Prat-Resina, X.; Konig, P.; Li, G. H.; Xu, D. G.; Guo, H.; Elstner, M.; Cui, Q. Development of Effective Quantum Mechanical/Molecular Mechanical (QM/MM) Methods for Complex Biological Processes. *J. Phys. Chem. B* **2006**, *110*, 6458−6469.

(5) Friesner, R. A.; Guallar, V. Ab Initio Quantum Chemical and Mixed Quantum Mechanics/Molecular Mechanics (QM/MM) Methods for Studying Enzymatic Catalysis. *Annu. Rev. Phys. Chem.* **2005**, *56*, 389−427.

(6) Marti, S.; Roca, M.; Andres, J.; Moliner, V.; Silla, E.; Tunon, I.; Bertran, J. Theoretical Insights in Enzyme Catalysis. *Chem. Soc. Rev.* **2004**, *33*, 98−107.

(7) Singh, U. C.; Kollman, P. A. A Combined Ab Initio Quantum Mechanical and Molecular Mechanical Method for Carrying out Simulations on Complex Molecular Systems: Applications to the CH$_3$Cl + Cl− Exchange Reaction and Gas Phase Protonation of Polyethers. *J. Comput. Chem.* **1986**, *7*, 718−730.

(8) Frushicheva, M. P.; Mills, M. J. L.; Schopf, P.; Singh, M. K.; Prasad, R. B.; Warshel, A. Computer Aided Enzyme Design and Catalytic Concepts. *Curr. Opin. Chem. Biol.* **2014**, *21*, 56−62.

(9) Kamerlin, S. C. L.; Warshel, A. The EVB as a Quantitative Tool for Formulating Simulations and Analyzing Biological and Chemical Reactions. *Faraday Discuss.* **2010**, *145*, 71−106.

(10) Plotnikov, N. V.; Prasad, B. R.; Chakrabarty, S.; Chu, Z. T.; Warshel, A. Quantifying the Mechanism of Phosphate Monoester Hydrolysis in Aqueous Solution by Evaluating the Relevant Ab Initio QM/MM Free-Energy Surfaces. *J. Phys. Chem. B* **2013**, *117*, 12807−12819.

(11) Zhang, J.; Kulik, H. J.; Martinez, T. J.; Klinman, J. P. Mediation of Donor-Acceptor Distance in an Enzymatic Methyl Transfer Reaction. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 7954−7959.

(12) Kulik, H.; Zhang, J.; Klinman, J.; Martinez, T. How Large Should the QM Region Be in QM/MM Calculations? The Case of Catechol O-Methyltransferase. 2015, arXiv: 1505.05730v1 [q-bio.BM].

(13) Blomberg, M. R. A.; Siegbahn, P. E. M. Quantum Chemistry as a Tool in Bioenergetics. *Biochim. Biophys. Acta, Bioenerg.* **2010**, *1797*, 129−142.

(14) Siegbahn, P. E. M.; Himo, F. The Quantum Chemical Cluster Approach for Modeling Enzyme Reactions. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2011**, *1*, 323−336.

(15) Siegbahn, P. E. M.; Borowski, T. Comparison of QM-Only and QM/MM Models for the Mechanism of Tyrosinase. *Faraday Discuss.* **2011**, *148*, 109−117.

(16) Han, W. G.; Noodleman, L. Quantum Cluster Size and Solvent Polarity Effects on the Geometries and Mossbauer Properties of the Active Site Model for Ribonucleotide Reductase Intermediate X: A Density Functional Theory Study. *Theor. Chem. Acc.* **2010**, *125*, 305−317.

(17) Liao, R. Z.; Yu, J. G.; Himo, F. Quantum Chemical Modeling of Enzymatic Reactions: The Case of Decarboxylation. *J. Chem. Theory Comput.* **2011**, *7*, 1494−1501.

(18) Hopmann, K. H.; Himo, F. Quantum Chemical Modeling of the Dehalogenation Reaction of Haloalcohol Dehalogenase. *J. Chem. Theory Comput.* **2008**, *4*, 1129−1137.

(19) Velichkova, P.; Himo, F. Methyl Transfer in Glycine N-Methyltransferase. A Theoretical Study. *J. Phys. Chem. B* **2005**, *109*, 8216−8219.

(20) Georgieva, P.; Wu, Q.; McLeish, M. J.; Himo, F. The Reaction Mechanism of Phenylethanolamine N-Methyltransferase: A Density Functional Theory Study. *Biochim. Biophys. Acta, Proteins Proteomics* **2009**, *1794*, 1831−1837.

(21) Georgieva, P.; Himo, F. Quantum Chemical Modeling of Enzymatic Reactions: The Case of Histone Lysine Methyltransferase. *J. Comput. Chem.* **2010**, *31*, 1707−1714.

(22) Flaig, D.; Beer, M.; Ochsenfeld, C. Convergence of Electronic Structure with the Size of the QM Region: Example of QM/MM NMR Shieldings. *J. Chem. Theory Comput.* **2012**, *8*, 2260−2271.

(23) Hartman, J. D.; Neubauer, T. J.; Caulkins, B. G.; Mueller, L. J.; Beran, G. J. Converging Nuclear Magnetic Shielding Calculations with Respect to Basis and System Size in Protein Systems. *J. Biomol. NMR* **2015**, *62*, 327−340.

(24) Sumowski, C. V.; Ochsenfeld, C. A Convergence Study of QM/MM Isomerization Energies with the Selected Size of the QM Region for Peptidic Systems. *J. Phys. Chem. A* **2009**, *113*, 11734−11741.

(25) Amar, D.; Berger, I.; Amara, N.; Tafa, G.; Meijler, M. M.; Aharoni, A. The Transition of Human Estrogen Sulfotransferase from Generalist to Specialist Using Directed Enzyme Evolution. *J. Mol. Biol.* **2012**, *416*, 21−32.

(26) Solt, I.; Kulhanek, P.; Simon, I.; Winfield, S.; Payne, M. C.; Csanyi, G.; Fuxreiter, M. Evaluating Boundary Dependent Errors in QM/MM Simulations. *J. Phys. Chem. B* **2009**, *113*, 5728−5735.

(27) Hu, L.; Soderhjelm, P.; Ryde, U. On the Convergence of QM/MM Energies. *J. Chem. Theory Comput.* **2011**, *7*, 761−777.

(28) Liao, R. Z.; Thiel, W. Comparison of QM-Only and QM/MM Models for the Mechanism of Tungsten-Dependent Acetylene Hydratase. *J. Chem. Theory Comput.* **2012**, *8*, 3793−3803.

(29) Singh, J. P.; Volavka, J.; Czobor, P.; Van Dorn, R. A. A Meta-Analysis of the Val158met COMT Polymorphism and Violent Behavior in Schizophrenia. *PLoS One* **2012**, *7*, e43423.

(30) Repič, M.; Vianello, R.; Purg, M.; Duarte, F.; Bauer, P.; Kamerlin, S. C. L.; Mavri, J. Empirical Valence Bond Simulations of the Hydride Transfer Step in the Monoamine Oxidase B Catalyzed Metabolism of Dopamine. *Proteins: Struct., Funct., Genet.* **2014**, *82*, 3347−3355.

(31) 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., Genet.* **2015**, *83*, 318−330.

(32) Lau, E. Y.; Bruice, T. C. Importance of Correlated Motions in Forming Highly Reactive near Attack Conformations in Catechol O-Methyltransferase. *J. Am. Chem. Soc.* **1998**, *120*, 12387−12394.

(33) 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. *J. Am. Chem. Soc.* **2000**, *122*, 2586−2596.

(34) Roca, M.; Marti, S.; Andres, J.; Moliner, V.; Tunon, I.; Bertran, J.; Williams, I. H. Theoretical Modeling of Enzyme Catalytic Power: Analysis of ″Cratic″ and Electrostatic Factors in Catechol O-Methyltransferase. *J. Am. Chem. Soc.* **2003**, *125*, 7726−7737.

(35) Soriano, A.; Castillo, R.; Christov, C.; Andres, J.; Moliner, V.; Tunon, I. Catalysis in Glycine N-Methyltransferase: Testing the Electrostatic Stabilization and Compression Hypothesis. *Biochemistry* **2006**, *45*, 14917−14925.

(36) Kanaan, N.; Pernia, J. J. R.; Williams, I. H. QM/MM Simulations for Methyl Transfer in Solution and Catalysed by Comt: Ensemble-Averaging of Kinetic Isotope Effects. *Chem. Commun.* **2008**, 6114−6116.

(37) Stewart, J. J. P. *Mopac2009*; Stewart Computational Chemistry: Colorado Springs, CO, 2008.

(38) Stewart, J. J. P. Optimization of Parameters for Semiempirical Methods V: Modification of Nddo Approximations and Application to 70 Elements. *J. Mol. Model.* **2007**, *13*, 1173−1213.

(39) Lee, F. S.; Chu, Z. T.; Warshel, A. Microscopic and Semimicroscopic Calculations of Electrostatic Energies in Proteins by the Polaris and Enzymix Programs. *J. Comput. Chem.* **1993**, *14*, 161−185.

(40) Rutherford, K.; Le Trong, I.; Stenkamp, R. E.; Parson, W. W. Crystal Structures of Human 108V and 108M Catechol O-Methyltransferase. *J. Mol. Biol.* **2008**, *380*, 120−130.

(41) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules 0.1. The Method. *J. Comput. Chem.* **1992**, *13*, 1011−1021.

(42) Vicatos, S.; Rychkova, A.; Mukherjee, S.; Warshel, A. An Effective Coarse-Grained Model for Biological Simulations: Recent Refinements and Validations. *Proteins: Struct., Funct., Genet.* **2014**, *82*, 1168−1185.

(43) Åqvist, J.; Warshel, A. Free Energy Relationships in Metalloenzyme-Catalyzed Reactions. Calculations of the Effects of Metal Ion Substitutions in Staphylococcal Nuclease. *J. Am. Chem. Soc.* **1990**, *112*, 2860−2868.

(44) Law, B. J.; Bennett, M. R.; Thompson, M. L.; Levy, C.; Shepherd, S. A.; Leys, D.; Micklefield, J. Effects of Active-Site Modification and Quaternary Structure on the Regioselectivity of Catechol-O-Methyltransferase. *Angew. Chem., Int. Ed.* **2016**, *55*, 2683−2687.

(45) Wesolowski, T.; Muller, R. P.; Warshel, A. Ab Initio Frozen Density Functional Calculations of Proton Transfer Reactions in Solution. *J. Phys. Chem.* **1996**, *100*, 15444−15449.

(46) Hong, G.; Strajbl, M.; Wesolowski, T. A.; Warshel, A. Constraining the Electron Densities in DFT Method as an Effective Way for Ab Initio Studies of Metal-Catalyzed Reactions. *J. Comput. Chem.* **2000**, *21*, 1554−1561.

(47) Warshel, A.; Levitt, M. Theoretical Studies of Enzymic Reactions - Dielectric, Electrostatic and Steric Stabilization of Carbonium-Ion in Reaction of Lysozyme. *J. Mol. Biol.* **1976**, *103*, 227−249.

(48) Reuter, N.; Dejaegere, A.; Maigret, B.; Karplus, M. Frontier Bonds in QM/MM Methods: A Comparison of Different Approaches. *J. Phys. Chem. A* **2000**, *104*, 1720−1735.

(49) Ferré, N.; Olivucci, M. The Amide Bond: Pitfalls and Drawbacks of the Link Atom Scheme. *J. Mol. Struct.: THEOCHEM* **2003**, *632*, 71−82.

(50) Assfeld, X.; Rivail, J.-L. Quantum Chemical Computations on Parts of Large Molecules: The Ab Initio Local Self Consistent Field Method. *Chem. Phys. Lett.* **1996**, *263*, 100−106.

(51) Gao, J.; Amara, P.; Alhambra, C.; Field, M. J. A Generalized Hybrid Orbital (GHO) Method for the Treatment of Boundary Atoms in Combined QM/MM Calculations. *J. Phys. Chem. A* **1998**, *102*, 4714−4721.

(52) Murphy, R. B.; Philipp, D. M.; Friesner, R. A. Frozen Orbital QM/MM Methods for Density Functional Theory. *Chem. Phys. Lett.* **2000**, *321*, 113−120.

(53) Zhang, Y.; Lee, T.-S.; Yang, W. A Pseudobond Approach to Combining Quantum Mechanical and Molecular Mechanical Methods. *J. Chem. Phys.* **1999**, *110*, 46−54.

(54) Klahn, M.; Braun-Sand, S.; Rosta, E.; Warshel, A. On Possible Pitfalls in Ab Initio Quantum Mechanics/Molecular Mechanics Minimization Approaches for Studies of Enzymatic Reactions. *J. Phys. Chem. B* **2005**, *109*, 15645−15650.

(55) Warshel, A.; Sharma, P. K.; Kato, M.; Parson, W. W. Modeling Electrostatic Effects in Proteins. *Biochim. Biophys. Acta, Proteins Proteomics* **2006**, *1764*, 1647−1676.

(56) Warshel, A.; Bora, R. P. Perspective: Defining and Quantifying the Role of Dynamics in Enzyme Catalysis. *J. Chem. Phys.* **2016**, *144*, 180901.