JCTC Journal of Chemical Theory and Computation

Case Studies of ONIOM(DFT:DFTB) and ONIOM(DFT:DFTB:MM) for Enzymes and Enzyme Mimics

Marcus Lundberg,[†,§] Yoko Sasakura,[†] Guishan Zheng,[‡,∥] and Keiji Morokuma*[,†,‡]

Fukui Institute for Fundamental Chemistry, Kyoto University, 34-4 Takano Nishihiraki-cho, Sakyo, Kyoto 606-8103, Japan, and Cherry L. Emerson Center for Scientific Computation and Department of Chemistry, Emory University, Atlanta, Georgia 30322

Received January 14, 2010

Abstract: The replacement of standard molecular mechanics force fields by inexpensive molecular orbital (QM′) methods in multiscale models has many advantages, e.g., a more straightforward description of mutual polarization and charge transfer between layers. The ONIOM(QM:QM′) scheme with mechanical embedding can combine any two methods without prior parametrization or significant coding effort. In this scheme, the environmental effect is evaluated fully at the QM′ level, and the accuracy therefore depends on how well the low-level QM′ method describes the changes in electron density of the reacting region. To examine the applicability of the QM:QM′ approach, we perform case studies with density-functional tight-binding (DFTB) as the low-level QM′ method in two-layer ONIOM(B3LYP/6-31G(d):DFTB) models. The investigated systems include simple amino acid models, one nonheme iron enzyme mimic, and the enzymatic reactions of Zn-β-lactamase and trypsin. For the last example, we also illustrate the use of a three-layer ONIOM(B3LYP/6-31G(d):DFTB:Amber96) model. The ONIOM extension, compared to the QM calculation for the small model system, improves the relative energies, but high accuracy (deviations below 1 kcal/mol) is not achieved even with relatively large QM models. Polarization effects are fairly well described using DFTB, but in some cases QM and QM′ methods converge to different electronic states. We discuss when the QM:QM′ approach is appropriate and the possibilities of estimating the quality of the ONIOM extension without having to make explicit benchmarks of the entire system.

I. Introduction

Over the three decades following their first implementation,[1] the quantum mechanics/molecular mechanics (QM/MM) methods have seen great success in a wide spectrum of applications, including biological reactions and materials science.[2−5] The success of the QM/MM methods is rooted in their multiscale nature, in which the system is partitioned into different regions treated at appropriate levels of theory. Only the core region, i.e., the QM region, is treated with a computationally expensive QM method that can describe chemical reactions, e.g., *ab initio* wave function, density functional, or semiempirical methods, while the rest of the system, i.e., the MM region, is treated with MM methods that are often thousands of times faster than QM methods. This compromise between accuracy and computational efficiency makes it possible to study systems that are computationally prohibitive to the pure QM methods.

The success of QM/MM has stimulated our development of the multiscale ONIOM method.[6−12] In two-layer ONIOM, the *high*-level method (QM) is applied only to a selected *model* system, i.e., the QM region with link atoms.[10,13] The

* Corresponding author telephone: +81-75-711-7843; fax: +81-75-781-4757; e-mail: morokuma@fukui.kyoto-u.ac.jp.

† Kyoto University.

‡ Emory University.

§ Present address: Department of Chemistry, Stanford University, Stanford, California 94305.

∥ Present address: Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138.

10.1021/ct100029p © 2010 American Chemical Society
Published on Web 03/16/2010

low-level method is applied to both the *model* and the *real* system. The latter is equivalent to the entire system and includes both the QM and MM regions of generic QM/MM models. The energy of the target *real*,*high* calculation is then approximated by an extrapolation scheme. An advantage of the ONIOM scheme is that all calculations are made on complete systems, which makes it possible to combine any methods, including molecular orbital methods, without prior parametrization or significant coding efforts. The method can also easily be extended to an arbitrary number of layers, although the present implementation is limited to three layers.[9,14−16] The ONIOM method has been successfully utilized in a wide range of applications, including transition metal catalysis, carbon nanotube chemistry, and enzymatic catalysis.[17,18]

Although the ONIOM method is different in design and implementation compared to generic QM/MM methods, there are also many similarities. When comparing these methods, we use the general terms QM region and MM region to represent different partitions of the system. One common concern is the boundary between QM and MM regions, and how to treat the interactions between the two. Some issues arising from the use of two fundamentally different physical descriptions of the system can be listed as follows: (1) the exchange interaction between the QM region and the MM region, which is partially included in the parametrized van der Waals interaction between the two regions, (2) the charge transfer between the QM and the MM region, (3) the polarization of the QM electron density induced by the MM atoms, and (4) the polarization of the MM atoms by QM and other MM atoms. Among these four issues, the last one is often considered the most severe and important.[19−25]

To properly address the polarization of MM atoms, the most straightforward approach (but by no means trivial) is to develop a polarizable force field.[19,20,26] In recent years, several approaches have been adopted to develop such force fields, e.g., the fluctuation charge model[19,20] and explicit polarization potential.[27] These models can partially alleviate the issues connected with the neglect of polarization effects, but they are not without problems.[19,20] For example, dissociated diatomic molecules bear finite charges on both atoms. This has been explicitly addressed by the Martinez group by using the QTPIE model, a generalization of the flucq model.[21,23,24] Although there is much progress, the performance of currently proposed polarizable force field models remains to be seen.

Development of accurate force fields for biological systems is further complicated by the large number of parameters.[22,28] QM/MM calculations have also been extended to include charge transfer between QM and MM regions using the principle of chemical potential equilibration.[29] A drawback of these methods is that they require multiple QM iterations to reach consistency between QM and MM charges, which leads to significant increases in computational cost.

An alternative approach is to describe the entire system by quantum mechanics. Large systems can be treated entirely by high-level QM methods by dividing them into fragments,[30−32] but these methods invariably spend the largest amount of time calculating the nonreacting part of the system. Savings in computational time can be achieved by a QM/QM′ approach, where a less expensive molecular orbital method, QM′, is used to describe the environment. In early work, Cortona proposed to combine different density functionals based on superposition of atomic densities.[33] Other groups have used the QM/QM′ approach to embed an expensive correlated wave function calculation in the environment of a relatively fast DFT method.[34−36] To be able to treat very large systems, Gogonea et al. constructed a hybrid DFT/semiempirical Hamiltonian based on the idea of equilibration of chemical potentials combined with the divide and conquer method.[30,37] The approach allows for mutual polarization and charge transfer but requires an iterative approach to equilibrate the chemical potential. Cui et al. have also coupled DFT with semiempirical methods, both in an iterative fashion and in the ONIOM formalism.[38]

The properties of the ONIOM method make it very straightforward to design QM:QM′ models.[7,10,11,39] As no parametrization is required to combine different methods, the selection of quantum mechanical methods can be made to fit each specific chemical process.[14,16,40,41] We are interested in reactions in complex systems, e.g., transition metal enzymes. For these systems, the compromise is often to use the best possible method to calculate the reaction energy and use a relatively cheap method to describe the environmental effects. In the present study, QM is a density functional method, while QM′ is the density functional tight binding method (DFTB). In the ONIOM(QM:QM′) scheme with mechanical embedding (ME), the *model*,*QM* calculations are independent of the QM′ electron density, and there is no need for multiple iterations of the QM region. This makes the QM:QM′−ME approach cost-efficient for systems that require expensive QM schemes, as it minimizes the number of QM iterations.

A limitation in the mechanical embedding scheme is that the environmental effects are evaluated only at the QM′ (*low*) level, as the difference between *real*,*QM′* and *model*,*QM′* electron densities. The QM′ method must therefore be able to describe the changes in electron density of the reacting region; e.g., if an electron transfer reaction occurs in the *model*,*QM* system, the same process should occur also in the layers described by QM′. The applicability of the QM:QM′ scheme thus depends heavily on the QM′ description of the environmental effect on a specific reaction, and it is difficult to perform comprehensive benchmark tests. The purpose of this investigation is therefore to demonstrate the advantages and limitations of the QM:QM′ method by applying it to a few illustrative examples. The resultant understanding of the principles of the ONIOM scheme should help in the design of more reliable QM:QM′ methods for a wide range of systems than before.

Another important aspect of the QM:QM′ scheme is that QM′ calculation is performed for the entire *real* system, which means that parametrized methods used as QM′ must have parameters applicable also for the reactive region. In the present paper, for reactions including transition metals, we propose to use a density functional method as the high-level QM method and the density-functional tight-binding (DFTB) method[42−44] as the low-level QM′ method. The

DFTB method is a second-order approximation of DFT;[45] this combination can be expected to give a similar performance to pure DFT methods.[46,47] Since DFTB is 100−1000 times faster than DFT, the ONIOM combination can be applied to a much larger system than a pure DFT method can be applied to. We have developed DFTB parameters for many of the first-row transition metal elements[48] and have also implemented a DFTB code in the Gaussian program, which enables us to use the full functionalities of this program for DFTB calculations.[47]

We have previously illustrated the applicability of the DFT:DFTB approach for transition metal systems by applying it to a nonheme enzyme (isopenicillin N synthase).[48] DFTB has also been used as a *low* layer in B3LYP:DFTB and MP2:DFTB combinations for copper- and titanium-containing systems, with errors typically less than 2 kcal/mol.[49] The choice of DFTB as the QM′ method for proteins is supported by the good performance for geometries and relative energies in biological systems.[50−52]

The QM:QM′−ME approach is efficient only when the cost of the QM calculation for the *model* system is higher than the cost of the QM′ calculation for the *real* system. For very large systems, it is therefore useful to make a three-layer partition of the system, i.e., QM:QM′:MM.[9] As a computationally cheap QM′ method can be utilized economically with thousands of atoms in many applications, this method pushes the boundary between QM and MM away from the reaction center, which is expected to largely reduce the issues caused by the QM:MM boundary mentioned above. The performance of three-layer ONIOM has been tested by our group before.[14−16] Other three-layer approaches have been used to describe, e.g., spectral tuning in protein environments.[53,54]

In this study, we test both two-layer ONIOM(B3LYP:DFTB) and three-layer ONIOM(B3LYP:DFTB:MM) models. In section II, we discuss how to evaluate ONIOM in terms of its performance compared to the *high*-level calculation of the *real* system. In the following section, several carefully chosen examples are presented, which range from proton affinity calculations of titratable amino acids to an active site model of Zn-β-lactamase, a nonheme iron catalase mimic, and finally to the acylation process in trypsin. A discussion of the advantages and limitations of the ONIOM(QM:QM′) scheme is presented in section IV before the conclusions in section V.

II. Computational Details

A. ONIOM Calculation Evaluation. ONIOM uses an extrapolation scheme to approximate the costly target calculation, a *high* (QM)-level treatment of the entire *real* system:

$$\Delta E^{\mathrm{ONIOM}} = \Delta E^{\mathrm{model,high}} + \Delta E^{\mathrm{real,low}} - \Delta E^{\mathrm{model,low}} \tag{1}$$

where ΔE refers to the energy of reaction, E(product) − E(reactant), or the activation energy, E(transition state) − E(reactant).

The error ΔD in ΔE in the ONIOM approximation, relative to the *real*,*high* target calculation, can be written as

$$\Delta D = \Delta E^{\mathrm{real,high}} - \Delta E^{\mathrm{ONIOM}} \tag{2}$$

A convenient measure for evaluating ONIOM calculations is the ΔS value that describes the environmental effect at each computational level:[11]

$$\Delta S^{\mathrm{high}} = \Delta E^{\mathrm{real,high}} - \Delta E^{\mathrm{model,high}} \tag{3}$$

$$\Delta S^{\mathrm{low}} = \Delta E^{\mathrm{real,low}} - \Delta E^{\mathrm{model,low}} \tag{4}$$

ΔS^{high} is "the substituent effect" on the energetics when the *real* system is approximated by the *model* system at the given *high*-level method, i.e., a measure of environmental effect; ΔS^{low} is the corresponding value for the *low*-level method. Using eqs 1−4, the error in the ONIOM approximation can be conveniently expressed in the form of ΔS values:

$$\begin{aligned}\Delta D &= \Delta E^{\mathrm{real,high}} - \Delta E^{\mathrm{ONIOM}} \\ &= \Delta E^{\mathrm{real,high}} - (\Delta E^{\mathrm{model,high}} + \Delta E^{\mathrm{real,low}} - \Delta E^{\mathrm{model,low}}) \\ &= \Delta S^{\mathrm{high}} - \Delta S^{\mathrm{low}}\end{aligned} \tag{5}$$

If the low-level method describes the environmental effect in the same way as the high-level method ($\Delta S^{\mathrm{low}} = \Delta S^{\mathrm{high}}$), the ONIOM energy is exact, i.e., becomes the same as the target calculation, $\Delta E^{\mathrm{real,high}}$, with $\Delta D = 0$. The error can also be small if the total effect of the environment is very small ($\Delta S \approx 0$) at each level; in this case, a *high*-level calculation of the *model* system is enough, and there is no need for an ONIOM extension. To distinguish the situation where the error is small due to small environmental effects from the situation where the ONIOM extrapolation scheme is accurate, we introduce the ONIOM error score (OES):

$$\mathrm{OES} = \Delta D/\Delta S^{\mathrm{high}} \tag{6}$$

A small absolute value of the error score means that the ONIOM setup is appropriate. If the absolute value of the error score is <1, the ONIOM calculation improves the result compared to the *model* calculation. If the absolute value is >1, the addition of the *real* system makes the relative energies even worse. It is possible that the *model*,*high* calculation gives the right result by neglecting two different environmental effects with opposite sign at the two different levels. In these cases, an ONIOM model could show a larger error but still give a qualitatively more correct description of the system. The ONIOM error score is therefore mainly of use when evaluating isolated or incremental effects.

B. Computational Details. Calculations have been performed using a development version of the Gaussian 03 package[55] in which the DFTB method has been implemented.[47] Tests of the accuracy of the ONIOM model are made against the target *real*,*high* calculation for all systems, and to simplify the analysis, all comparisons between methods are made at the same geometry. A previous QM:DFTB test by Iordanov showed that geometry errors were in general smaller than energetic errors.[49]

In most calculations, the high-level method is B3LYP/6-31G(d), mainly because DFTB is parametrized against DFT with a double-ζ basis set. The contribution from the DFTB layer only reflects the environmental effect, and assuming

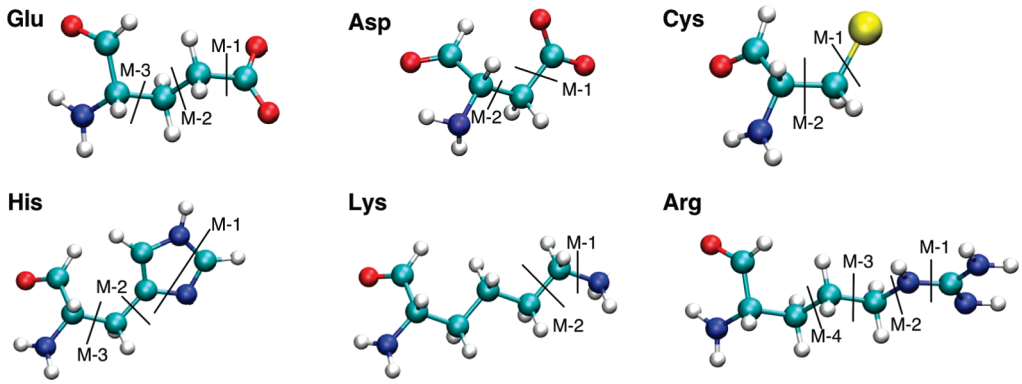


Figure 1. ONIOM(QM:QM′) partitions in amino acid side chains. Lines labeled M-1 to M-4 illustrate model cuts, with the part to the right being part of the *model* system.

the selected high-level method gives a good description of that effect, the ONIOM error should not be very sensitive to the choice of high-level methods. In cases where the environmental effects are basis-set-dependent, e.g., in the calculation of proton affinities, calculations have also been made with larger basis sets.

III. Results

The selected systems represent a wide range of applications, from simple amino acid models to medium-sized models of transition metal systems and a full protein. The purpose of calculating proton affinities for amino acids is to illustrate the general principles of the QM:QM′ approach for reactions with clear environmental effects. Tests of medium-sized models include transition metal systems, one enzymatic reaction, Zn-β-lactamase, and an iron catalase mimic. The latter example shows how the presence of multiple electronic states affects the applicability of the QM:QM′ scheme. Finally, we apply the three-layer QM:QM′:MM model to trypsin to illustrate that the method can be used for full enzymatic systems.

A. Proton Affinities of Amino Acids and Peptides—Effect of Improved Treatment of Side Chains. In active-site QM models of enzymatic reactions, amino acid residues are often represented by truncated models of their side chains; e.g., acetate is used as a model for the negative residue Asp.[56] Here, we evaluate the benefit of an additional QM′ layer to represent the left-out part of the side chain when calculating the gas phase proton affinity of six titratable side chains: arginine, lysine, histidine, aspartate, glutamate, and cysteine. Proton affinities are very sensitive to environmental effects and good candidates for evaluating different methods.[57]

Amino Acid Monomers. In the first set of calculations, the *real* systems are the side chains, with truncated backbone bonds capped by hydrogen atoms. Several *model* systems are designed by making different cuts in each side chain, see Figure 1. Proton affinities with pure DFTB have been calculated using an energy of H^+ of 141.9 kcal/mol.[58] ONIOM results are not affected by this value as it cancels in the ($E^{\text{real,low}} - E^{\text{model,low}}$) contribution. Results do not include corrections for basis set superposition error (BSSE). The ΔS^{high} values are not likely to be much affected by BSSE, as the substituents are made in sites away from the proton. Geometries are optimized using ONIOM separately for each size of the *model* system, and using B3LYP/6-31G(d) for the reference calculation. The energies from the respective *real*,*high* calculations for each ONIOM model can therefore be used to assess the quality of the ONIOM geometry optimization.

Results are summarized in Table 1. The ONIOM results are significant improvements compared to the *model*,*high* calculations (using the *real*,*high* calculations as a reference). As an example, the error for an acetate model (M-2) of glutamate is 10.8 kcal/mol at the *model*,*B3LYP* level and decreases to 2.2 kcal/mol with the additional DFTB layer (ONIOM error score of 0.21). For residues where the base is negative, i.e., Asp, Glu, and Cys, the ONIOM errors do not consistently decrease as the size of the *model* system increases. The remaining ONIOM error for the largest *model* systems comes from an underestimation of the effect of the backbone of 2−4 kcal/mol in DFTB compared to DFT. In the smallest model system, this error is partly canceled by a difference between DFTB and DFT in describing the effect of adding the first methyl group. This error cancellation gives a smaller apparent ONIOM error for the smallest *model* systems compared to the larger *model* systems.

In contrast, residues whose base is neutral, i.e., His, Lys, and Arg, show lower errors when the size of the *model* system increases, and errors for cuts three or more bonds away from the proton are ∼1 kcal/mol or less. ONIOM optimization also gives overall good performance for geometries with typical deviations in *real*,*high* energy between different geometries of <1 kcal/mol. However, the artificially truncated backbone is not restricted and can adapt different conformations, which may lead to jumps in the calculated proton affinity. This explains the differences in the calculated proton affinity of the *real*,*B3LYP* system between Lys model M-2 and the full model, as well as differences between Arg model M-4 and the full model, see Table 1.

Tripeptides. To better include the effect of the peptide backbone, two tests were made for tripeptides. Histidine was chosen as a representative of the "neutral base" group, and glutamate was chosen as a representative of the "negative base" group. For the Gly−Glu−Gly (GEG) tripepeptide, the error when using acetate as a model (M-2) is 28.2 kcal/mol at the B3LYP level, and adding the DFTB layer reduces the error to 0.7 kcal/mol. Results are similar also for other cut positions, see Table 1. For the Gly−His−Gly (GHG)

Table 1. Proton Affinities (kcal/mol) of Selected Amino Acids Calculated Using ONIOM B3LYP/6-31G(d):DFTB[a]

real system	model system	B3LYP (real)	ONIOM	B3LYP (model)	DFTB (real)	DFTB (model)	ΔS^{high}	ΔD	OES
Glu	M-1	355.6	355.2	362.5	358.4	365.7	−6.9	0.4	−0.05
	M-2	355.6	357.9	366.5	358.0	366.6	−10.8	−2.2	0.21
	M-3	355.3	357.5	364.8	358.4	365.7	−9.5	−2.2	0.23
	full	**354.7**			358.0				
Asp	M-1	352.3	352.8	362.5	356.0	365.7	−10.2	−0.5	0.05
	M-2	352.1	355.5	366.3	355.8	366.6	−14.1	−3.3	0.24
	full	**352.0**			355.7				
Cys	M-1	353.2	350.4	363.0	343.9	356.5	−9.7	2.9	−0.29
	M-2	353.4	357.1	369.1	344.1	356.1	−15.7	−3.7	0.23
	full	**352.9**			344.3				
His	M-1	247.2	252.6	218.3	239.5	205.2	28.9	−5.5	−0.19
	M-2	246.2	244.4	236.6	239.5	231.7	9.6	1.9	0.19
	M-3	246.4	245.2	240.6	239.5	234.9	5.8	1.2	0.20
	full	**246.5**			239.2				
Lys	M-1	233.5	231.7	217.2	213.0	198.5	16.3	1.8	0.11
	M-2	233.8	233.1	226.8	212.3	206.0	7.0	0.7	0.10
	full	**230.5**			212.5				
Arg	M-1	258.0	259.3	239.8	251.4	231.8	18.2	−1.4	−0.08
	M-2	256.8	255.1	249.2	248.9	243.0	7.6	1.7	0.22
	M-3	254.1	253.7	253.9	246.9	247.1	0.2	0.5	2.38
	M-4	257.0	256.9	255.5	250.2	248.7	1.5	0.0	0.03
	full	**253.9**			250.5				
				Tripeptides					
GEG	M-1	338.9	336.3	363.3	339.5	366.5	−24.4	2.6	−0.11
	M-2	338.7	339.4	366.8	339.2	366.7	−28.2	−0.7	0.03
	M-3	338.8	339.6	367.0	339.5	366.9	−28.2	−0.8	0.03
	full	**338.2**			340.2				
GHG	M-2	246.5	247.8	235.9	243.3	231.4	10.6	−1.3	−0.12
	full	**248.5**			242.1				

[a] Separate values for B3LYP/6-31G(d) and DFTB applied to *real* and *model* systems are also listed. For each model, all energy calculations are performed at the ONIOM optimized geometry. The *real,B3LYP* and *real,DFTB* results therefore also vary with the ONIOM partition. For descriptions of the different computational models, see Figure 1. ΔS^{high}, ΔD, and OES (ONIOM error score) are defined in section II.A. One-letter amino acid codes are used for the tripepeptides.

Table 2. Proton Affinities (kcal/mol) for Two Tripeptides Using ONIOM(B3LYP:DFTB) with Different Size of the Basis Set in the B3LYP Calculation[a]

real system	QM (high) basis	model system	B3LYP (real)	ONIOM	B3LYP (model)	DFTB (real)	DFTB (model)	ΔS^{high}	ΔD	OES
GEG	6-31G(d)	M-1	338.2	337.7	363.6	340.2	366.1	−25.3	0.6	−0.02
	6-31+G(d)	M-1	326.8	319.2	345.1	340.2	366.1	−18.4	7.5	−0.41
	6-311++G (2df, 2pd)	M-1	331.8	324.2	350.1	340.2	366.1	−18.4	7.6	−0.41
GHG	6-31G(d)	M-2	248.5	246.3	234.9	242.1	230.8	13.6	2.2	0.16
	6-31+G(d)	M-2	242.6	241.1	229.8	242.1	230.8	12.9	1.6	0.12
	6-311++G (2df, 2pd)	M-2	245.8	242.9	231.5	242.1	230.8	14.3	3.0	0.21

[a] The separate values for B3LYP and DFTB applied to *real* and *model* systems are also listed. Calculations are performed as single point on B3LYP/6-31G(d) optimized geometries. ΔS^{high}, ΔD, and OES (ONIOM error score) are defined in section II.A. One-letter amino acid codes are used for the tripepeptides, and ONIOM cuts are shown in Figure 2.

tripeptide, the error of an imidazole model (M-2) is 10.6 kcal/mol, and the error is reduced to 1.3 kcal/mol in ONIOM. These results indicate that DFTB fairly well describes the effects of a DFT calculation with a double-ζ basis.

Extending the basis set from 6-31G(d) to 6-31+G(d), which includes diffuse functions, has large effects on the ΔS value (environmental effect) for the negative Gly−Glu−Gly system (from −25.3 to −18.4 kcal/mol), see Table 2. The DFTB ΔS value is closer to the B3LYP value without diffuse functions, and consequently the ONIOM error increases by 7 kcal/mol when going from the smallest to the largest basis set. Further increasing the basis to triple-ζ quality (6-311++(2df,2pd)) has a very limited effect.

The effect of diffuse functions is particularly large in the present case because the negatively charged *model* system is not well described with the small 6-31G(d) basis set; e.g., an acetate *model* system has occupied orbitals with positive eigenvalues. Adding diffuse functions thus stabilizes the acetate group relative to its base, leading to a significant basis set effect on the proton affinity. For better balanced systems, the effect of the environment is less basis-set-dependent, as seen by the results for the Gly−His−Gly (GHG) tripeptide in Table 2.

Most QM calculations of enzyme active sites use truncated amino acid models, and considering the large effects of the backbone on the calculated proton affinities (−18.4 kcal/

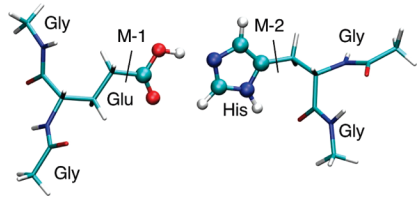


Figure 2. Proton transfer between two tripeptides, Gly−Glu−Gly and Gly−His−Gly. Lines M-1 and M-2 illustrate the used *model* cuts (labels match the cuts made for the single amino acid systems in Figure 1). Atoms in ball-and-stick are included in the *model* system, while atoms in stick representation are part of the *real* system.

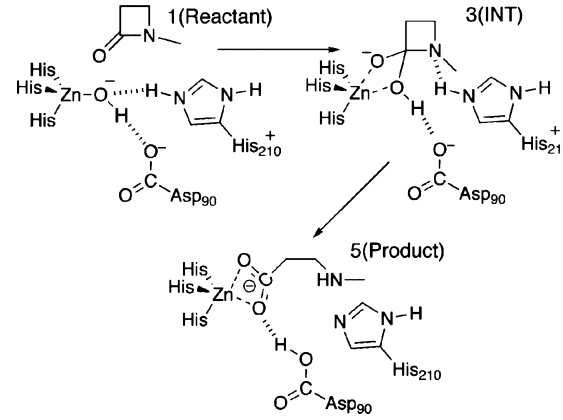
mol for Gly−Glu−Gly (GEG) with the largest basis, see Table 2) raises the question whether there is a significant error in the relative energies of proton transfer processes in active-site models. Comparing only the relative proton affinities of the *model*,*high* systems to those of the *real*,*high* systems for the tripeptides Gly−Glu−Gly and Gly−His−Gly (GHG) would give a difference in proton transfer energy of more than 30 kcal/mol. However, in the combined system, see Figure 2, the deviation between the *model*,*high* and the *real*,*high* calculation is only 1.8 kcal/mol, because that reaction does not include changes in the charge of the system. For this system, the use of ONIOM does not improve the result (deviation of −1.7 kcal/mol compared to the *model*,*high* value).

B. Zn-β-lactamase−Various Active Site Models. Enzyme catalysis is an attractive area for multiscale models, and the B3LYP/6-31G(d):DFTB combination is therefore evaluated for an enzymatic reaction, the hydrolysis of N-methylazetidinone in a mononuclear Zn-β-lactamase. Production of β-lactamase is considered the primary route in which bacteria acquire resistance to the common β-lactam antibiotics such as penicillins and cephalosporins. DFTB calculations of Zn active sites give relatively good agreement with B3LYP for distances and most reaction energies.[59] Modeling of substrate binding in a dinuclear Zn-β-lactamase shows relatively similar Zn−ligand distances for DFTB/MM simulations compared to B3LYP calculations (usually within 0.1 Å for reactants and intermediates).[60,61]

The reaction pathway and the initial coordinates were obtained from an active-site study by Diaz et al.[62] In their HF/6-31G(d) calculations, the reaction goes through five stationary points. As seen in Scheme 1, from 1(Reactant), a hydroxyl group bound to Zn performs a nucleophilic attack on the carbonyl of the four-membered β-lactam ring 2(TS). The ring is still intact in the tetrahedral intermediate 3(INT), and cleavage of the β-lactam C−N bond 4(TS) is initiated by proton transfer from His210 to the substrate nitrogen. The reaction also leads to a proton transfer from the Zn-coordinated hydroxo group to Asp90, leading to the final state 5(Product). At the B3LYP/6-31G(d) level, the reaction goes directly from 2(TS) to 5(Product) without any intermediates.[62]

To design a stringent test, a minimum-sized *model* system (model Z-1) was used, as shown in Figure 3. All ONIOM cuts have to be made through formal single bonds, and Asp90 is modeled as formic acid and the histidine residues as

Scheme 1. Reaction Mechanism for a Mononuclear Zn-β-lactamase Adapted from Ref 62



methylamine. This leaves 38 atoms in the *model* system (including link hydrogens). Calculations were initially performed at the HF optimized geometries because they cover a larger part of the reaction coordinate, see Figure 4 and Table S1 in Supporting Information.

The ONIOM energies for model Z-1 are improvements over the DFTB results (mean average deviation (MAD) for the 4 states in Figure 4 decreases to 4.3 kcal/mol from 10.6 kcal/mol), but only slightly better than the *model*,*B3LYP* calculations (MAD of 5.5 kcal/mol). The relatively large errors probably come from an inappropriate model partition, especially the use of methylamines as histidines in the *model* system. The 5-membered imidazole ring is conjugated while the *model* methylamine is not. Cuts in conjugated systems can be done in ONIOM, but they are in general more challenging. Similar partitions gave significant errors in the calculation of proton affinities, see Table 2. Errors can be systematically decreased by increasing the size of the *model* system. If the histidines that coordinate Zn are fully included in the *model* system (Z-2), the ONIOM error is reduced to 3.4 kcal/mol. Moving the proton-donating His210 into the *model* system (Z-3) further decreases the ONIOM deviation to 1.4 kcal/mol.

ONIOM deviations do not correlate with errors in the DFTB method itself. For 3(INT), the DFTB result is within 3 kcal/mol of the B3LYP value, but the ONIOM deviation in the small Z-1 model is 5.2 kcal/mol. For 5(Product), the DFTB deviation is 28 kcal/mol, but the ONIOM error is only 0.3 kcal/mol.

In the next step, we optimize the three stationary points on the B3LYP potential energy surface, 1(Reactant), 2(TS), and 5(Product), with both B3LYP and ONIOM(B3LYP:DFTB) using model Z-1; the results are shown in Figure 5. As shown in Table 3, two optimized geometries give only small energy difference up to 0.7 kcal/mol for either state at either level of calculations, indicating that the ONIOM geometry optimization is quite reliable for the reaction energetics.

C. Reaction between Hydrogen Peroxide and a Nonheme Iron Catalase Mimic. Redox-active transition-metal centers present special challenges in modeling due to the presence of nearly degenerate electronic states, which affects both the accuracy and the convergence properties of

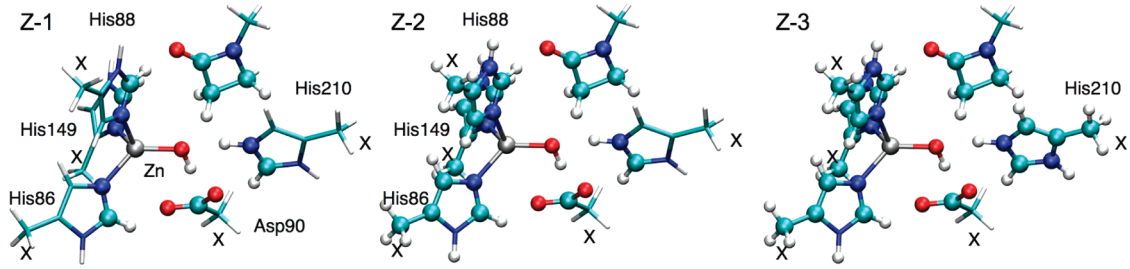


Figure 3. ONIOM models (Z-1 to Z-3) of the active site in a mononuclear Zn-β-lactamase. Atoms in the *model* system are shown in ball-and-stick representation, while atoms in the *real* system are shown in stick representation. Atoms with coordinates frozen during optimizations are marked with X.

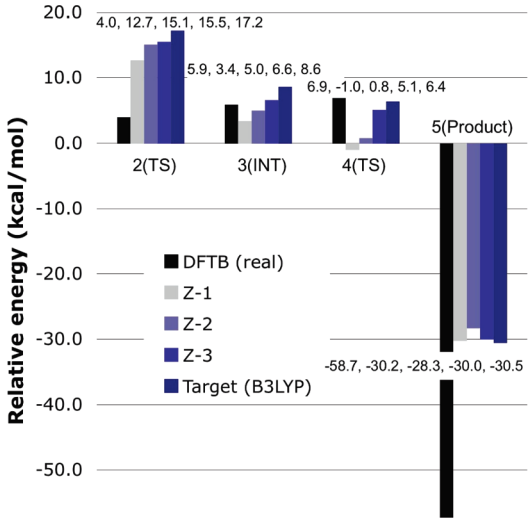


Figure 4. DFTB, ONIOM(B3LYP/6-31G(d):DFTB), and B3LYP/6-31G(d) energies for hydrolysis of N-methylazetidinone in mononuclear Zn-β-lactamase in Scheme 1. Illustrations of ONIOM models Z-1 to Z-3 are given in Figure 3.

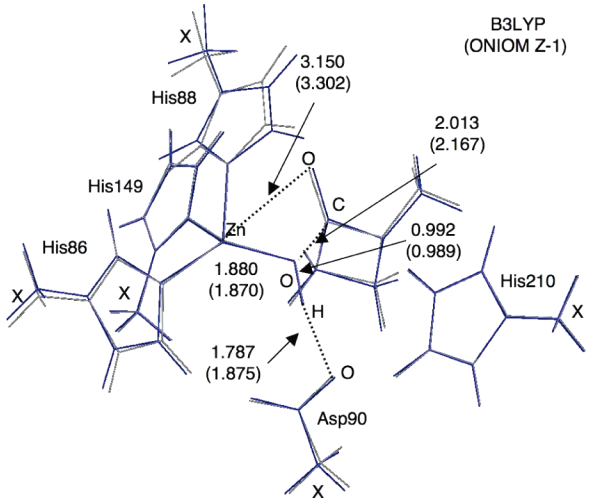


Figure 5. Optimized structure of 2(TS) with B3LYP and ONIOM(B3LYP:DFTB) for model Z-1. Selected distances (in Å) are given at B3LYP/6-31G(d) level with ONIOM(B3LYP/6-31G(d):DFTB) results in parentheses. The total RMSD between the two structures is 0.124 Å.

the electronic structure calculations. To extend the tests of the ONIOM(B3LYP:DFTB) approach to redox-active systems, we study the reaction between hydrogen peroxide and an inorganic catalase mimic, the dibenzotetraaza[14]-annulene−Fe^{III} complex ([Fe($C_{24}H_{22}N_4O_4$)]$^+$),[63] see Figure 6. The ONIOM model is formed by truncating the conjugated ligand at bonds that can formally be assigned as single bonds.

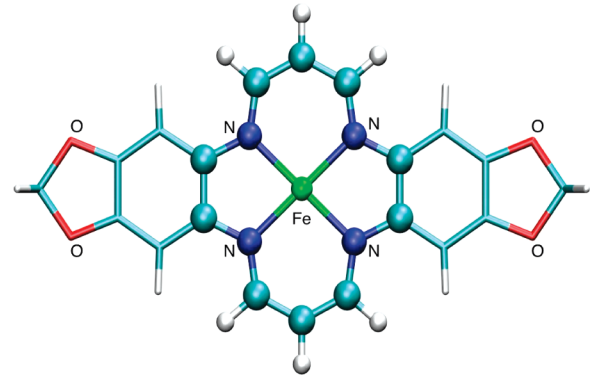


Figure 6. ONIOM division for the dibenzotetraaza-[14]annulene−Fe^{III} complex. Atoms in the *model* system are shown in ball-and-stick representation, while atoms in the *real* system are shown in stick representation.

QM:QM′ calculations of transition metal systems are challenging as three separate calculations of the transition metal system have to be performed. The three calculations must all converge to the same electronic state. Otherwise, the low-level method describes a different state from the high-level method, and the environmental effect in the *real*,*low* calculation becomes qualitatively incorrect. Despite these challenges, a previous test of a redox reaction in a nonheme transition metal enzyme gave reasonable results for the B3LYP:DFTB method.[48] With this in mind, it is of great interest to understand under what circumstances that QM:QM′ methods can be applied to transition metal systems.

The reaction mechanism of the catalase mimic has been previously studied by DFT methods.[64,65] The present investigation follows the reaction pathway in ref 64. A flaw in that study is that the additional axial (proximal) ligand was not taken into account.[65] For the present purpose, this is not critical because the comparison between methods should be valid even if the test system is not an ideal representation of an experimental situation.

According to ref 64, the reaction between hydrogen peroxide and the iron complex goes through nine stationary points and eventually leads to formation of water and an Fe(IV)−oxo species. The potential energy profiles from the full B3LYP calculation as well as ONIOM and *model*,*B3LYP* calculations are shown in Figure 7. Full results are given in Table 4. All calculations are performed on the quartet surface,

Table 3. B3LYP/6-31G(d) and ONIOM (B3LYP/6-31G(d):DFTB) Energies (in kcal/mol) for the Hydrolysis of N-Methylazetidinone in Mononuclear Zn-β-lactamase at B3LYP and ONIOM Optimized Geometries

model	state	B3LYP (real)	ONIOM	B3LYP (model)	DFTB (real)	DFT (model)	ΔS^{high}	ΔD	OES
				Geometries Obtained at the B3LYP/6-31G(d) Level					
Z-1	1	0.0	0.0	0.0	0.0	0.0	NA	NA	NA
	2(TS)	16.2	10.1	12.3	2.2	4.4	3.9	6.1	1.57
	5(Prod)	−30.0	−29.7	−36.5	−55.0	−61.8	6.5	−0.3	−0.05
				Geometries Obtained at the ONIOM(B3LYP/6-31G(d):DFTB) Level					
Z-1	1	0.0	0.0	0.0	0.0	0.0	NA	NA	NA
	2(TS)	15.6	11.4	10.0	2.4	1.0	5.6	4.2	0.76
	5(Prod)	−30.2	−31.1	−35.7	−55.0	−59.6	5.6	0.9	0.16

and results are evaluated at the B3LYP/6-31G(d) geometries. Transition states have not been fully optimized; instead, the Fe−O, O−O, and O−H distances were taken from the study of Wang et al.[64] and kept frozen during the optimization. There are differences in the reported B3LYP potential energy surface in Figure 7 and the corresponding values in ref 64, and they come from differences in basis set between the two studies, as well as the neglect of zero-point energy in the present study.

The mean absolute deviation between B3LYP:DFTB and the target B3LYP results are 6.3 kcal/mol, and the maximum deviation is 13.2 kcal/mol, see Table 4. The performance of the B3LYP:DFTB method is much better than that of the stand-alone DFTB method, but the deviation is still too large to be acceptable in mechanistic studies. Adding the low-level correction in stationary points gives worse results than obtained by the *model*,*B3LYP* calculation, as seen from the large ONIOM error scores in Table 4.

To understand the ONIOM deviations, we compare the electronic structures obtained in the four separate calculations: *real*,*high*; *model*,*high*; *real*,*low*; and *model*,*low,* as shown in Table 5. At the first stationary point **1**, which is the catalyst before addition of hydrogen peroxide, the full B3LYP calculation (*real*,*high*) corresponds to an intermediate spin triplet Fe(II) which resulted from reduction of its formal Fe(III) state by an electron transfer from the annulene ligand. The reason for this electron transfer is probably the lack of the distal ligand.[65] The spin on iron couples ferromagnetically with the spin of the unpaired electron of the ligand to form a quartet state. The present assignment agrees with the results in ref 64 (unpaired electrons in d_{z^2}, d_{xz}, and the ligand b_{1u} orbitals). The corresponding porphyrin−Fe^{III} complex does not oxidize the ligand,[66] probably because the electron affinity is lower for the annulene ligand than for the porphyrin.[64] In B3LYP, the electronic structure of the *model* system is rather similar to that of the *real* system, see Table 5. The DFTB calculation of the *real* system gives a description more similar to an intermediate-spin quartet Fe(III) system; i.e., the oxidation of the ligand does not occur fully in DFTB. On the other hand, DFTB calculation for the *model* system converges to a high-spin quintet Fe(II) state that couples antiferromagnetically with the spin on the ligand, instead of an intermediate-spin Fe(II) state with ferromagnetic coupling to the ligand spin as in the B3LYP calculation, see Table 5. Thus, the ONIOM subtraction scheme provides the triplet Fe(II) state for the *model* system, but with the substituent effect evaluated using DFTB between the quartet Fe(III) *real* system and the quintet Fe(II) + ligand radical *model* system. This does not correspond exactly to the desired triplet Fe(II) state in the *real*,*B3LYP* calculations, and the QM:QM′ extrapolation scheme does not work as intended for this system.

The situation is similar for the stationary point **2**. Table 5 shows that *real*,*DFTB* and *model*,*B3LYP* give the quartet Fe(III) state, while *model*,*DFTB* gives the high-spin quintet Fe(II) state. Again the ONIOM scheme does not result in the correct description of the triplet Fe(II) state in the *real*,*B3LYP* calculations.

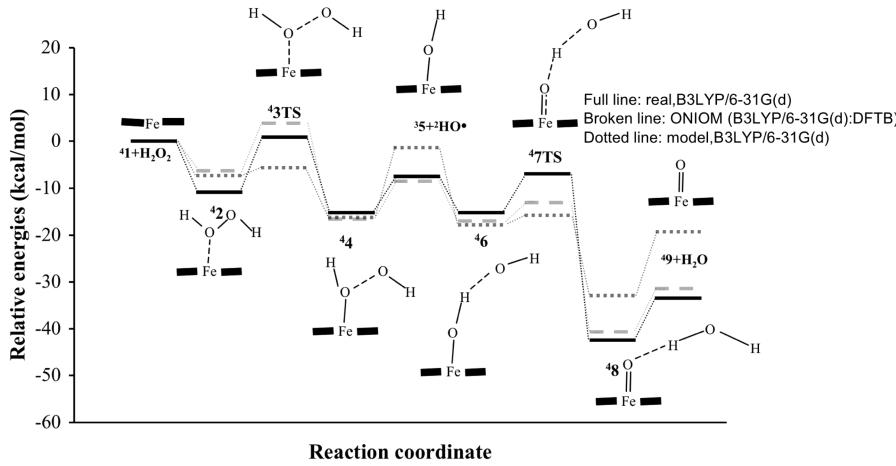


Figure 7. B3LYP/6-31G(d) and ONIOM (B3LYP/6-31G(d):DFTB) potential energy profiles for the formation of a high-valent ferryl-oxo species in a catalase mimic. The ONIOM system is shown in Figure 6.

Table 4. Relative Energies (in kcal/mol) for the Reaction between Hydrogen Peroxide and the Iron Complex Dibenzotetraaza[14]Annulene−Fe^{III} Calculated Using ONIOM(B3LYP/6-31G(d):DFTB)[a]

stationary point	B3LYP (real)	ONIOM	B3LYP (model)	DFTB (real)	DFTB (model)	ΔS^{high}	ΔD	OES
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
2	−10.8	−6.8	−5.7	−24.5	−23.4	−5.2	−4.1	0.79
3	0.8	−5.6	1.9	−33.6	−26.2	−1.1	6.4	−5.93
4	−15.4	−16.1	−16.1	−55.9	−55.9	0.8	0.8	0.98
5	−7.6	−1.7	−8.5	−29.2	−36.0	0.9	−5.9	−6.89
6	−15.2	−17.9	−17.3	−57.5	−56.9	2.1	2.7	1.27
7	−6.9	−15.4	−12.8	−75.5	−72.9	5.8	8.5	1.45
8	−42.4	−33.3	−40.8	−91.2	−98.7	−1.6	−9.1	5.76
9	−33.3	−20.1	−31.5	−86.3	−97.7	−1.8	−13.2	7.52
mean absolute error						2.4	6.3	

[a] The separate values for B3LYP/6-31G(d) and DFTB applied to *real* and *model* systems are also listed. The different stationary points are shown in Figure 7. All calculations are performed at B3LYP/6-31G(d) optimized geometries.

Table 5. Mulliken Spin Populations and Assigned Charges for the Formal Annulene−Fe^{III} Complex[a]

		Mulliken spin population/assigned state label			
point	group	*real,B3LYP*	*real,DFTB*	*model,B3LYP*	*model,DFTB*
1	Fe	2.16 /triplet Fe(II)	3.39 /quartet Fe(III)	2.00 /triplet Fe(II)	4.10 /quintet Fe(II)
	ligand	$0.85/L^{-1}$	$-0.39/L^{-2}$	$1.00/L^{-1}$	$-1.10/L^{-1}$
2	Fe	2.18/triplet Fe(II)	3.33/quartet Fe(III)	2.54/quartet Fe(III)	4.01 /quintet Fe(II)
	ligand	$0.80/L^{-1}$	$-0.35/L^{-2}$	$0.42/L^{-2}$	$-1.03/L^{-1}$
	H_2O_2	0.03/0	0.03/0	0.04/0	0.02/0

[a] Only the first two stationary points in the reaction with hydrogen peroxide are listed. The assigned states are based on the number of unpaired electrons and should be taken as labels rather than exact assignment of states.

D. The Acylation Process in Trypsin—Two- and Three-Layer Calculations. The target of the ONIOM(QM:QM′) approach is to describe full enzymatic systems, but for systems with tens of thousands of atoms, the time required for the QM′ calculation would be very high. As an intermediate stage, the three-layer ONIOM(QM:QM′:MM) approach offers an efficient alternative. The three-layer QM:QM′:MM energy is obtained from five subcalculations:

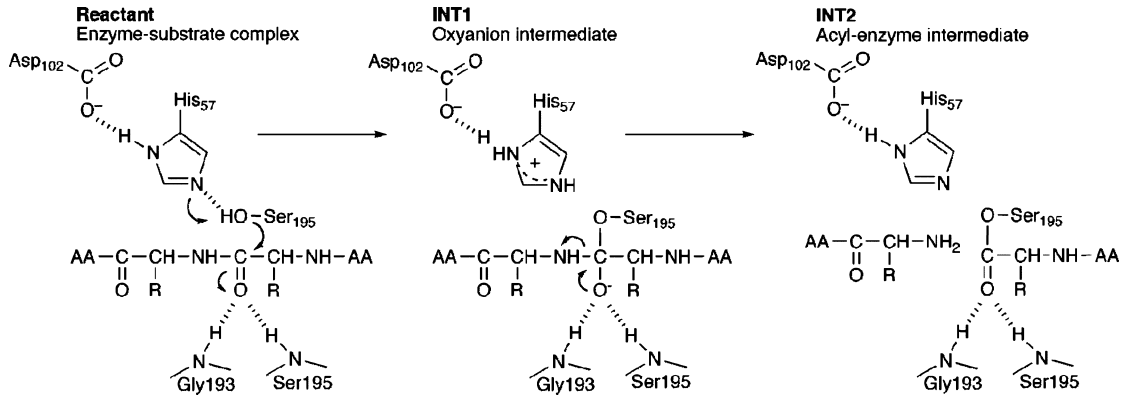
$$E^{\mathrm{ONIOM}} = E^{\mathrm{model,QM}} + E^{\mathrm{Intermediate,QM'}} - E^{\mathrm{model,QM'}} + E^{\mathrm{Real,MM}} - E^{\mathrm{Intermediate,MM}} \quad (7)$$

We illustrate the use of the three-layer approach for peptide cleavage in serine proteases. This reaction is one of the most well-known enzymatic reactions and appears in many biochemistry textbooks. Scheme 2 shows the proposed mechanism for the first half of the reaction, the acylation step.

An important motif in these enzymes is a conserved Ser−His−Asp triad, known as the catalytic triad. Ser195 performs a nucleophilic attack on the substrate peptide and transfers a proton to His57 (**INT1**). The negative Asp102 significantly stabilizes the proton transfer reaction by polarizing His57. The nucleophilic attack of serine leads to the formation of an oxyanion in the substrate peptide chain, and this species is stabilized by a second important catalytic motif, the "oxyanion hole". This motif provides hydrogen bonds to the peptide carbonyl, interactions that increase in strength as the charge of the oxygen increases. In the next step, the peptide C−N bond breaks, and the newly formed N-terminal group accepts the proton from His57 (**INT2**). This completes the acylation part of the reaction.

Active-Site QM:QM′ Models. As the full enzyme cannot be benchmarked by QM calculations, we start with active-site models. To identify the effects of different catalytic motifs, we use three two-layer ONIOM systems (T-1 to T-3),

Scheme 2. Textbook Mechanism for the Acylation Step in Serine Proteases



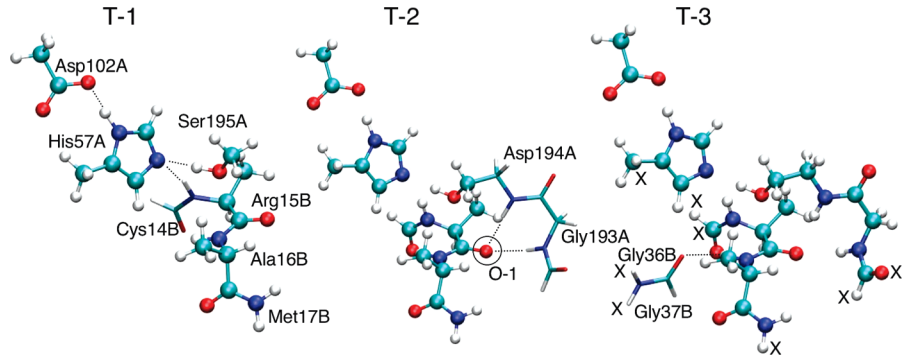


Figure 8. ONIOM divisions for an active-site model of trypsin. A and B in the amino acid labels refer to different peptide chains. Atoms in the *model* system are shown in ball-and-stick representation, while atoms in the *real* system are shown in stick representation. Hydrogen bonds are marked with dashed lines as they appear in the extended systems. Atoms with coordinates frozen at the X-ray structure are marked with **X** in model T-3.

Table 6. Reaction Energies (in kcal/mol, Relative to the Reactant) for the Acylation Process in Trypsin Calculated Using ONIOM B3LYP/6-31G(d):DFTB[a]

model	state	B3LYP (real)	ONIOM	B3LYP (model)	DFTB (real)	DFTB (model)	ΔS^{high}	ΔD	OES
T-1	**INT1**	30.5	27.6	28.0	29.9	30.4	2.4	2.9	1.17
	INT2	2.8	4.6	8.1	2.0	5.6	−5.3	−1.7	0.32
T-2	**INT1**	23.3	18.6	30.5	18.1	29.9	−7.2	4.7	−0.66
	INT2	11.6	8.4	2.8	7.6	2.0	8.8	3.3	0.37
T-3	**INT1**	27.3	24.8	23.3	19.5	18.1	4.0	2.5	0.63
	INT2	12.6	11.7	11.6	7.6	7.6	0.9	0.9	0.92

[a] The separate values for B3LYP/6-31G(d) and DFTB applied to *real* and *model* systems are also listed. The different models are shown in Figure 8.

with different *model* and *real* selections, as shown in Figure 8. T-1 (47 atoms in the *model* system, 51 atoms in the *real* system) includes the catalytic triad and four amino acids of the substrate peptide, but one of the peptide units, Cys14B, is only part of the *real* system to test how the QM′ layer handles through-bond interactions. T-2 (51 atoms in the *model* system, 65 atoms in the *real* system) is created by adding the groups that stabilize the oxyanion hole. The environmental effect comes from two hydrogen bonds from the protein backbone. In DFTB, hydrogen bond distances are underestimated by about 0.1 Å on average, and hydrogen bonding energies are systematically underestimated by 1−2 kcal/mol.[51] To isolate the effect of this new motif, all atoms in the T-1 system are placed in the *model* system. ONIOM T-3 (65 atoms in the *model* system, 71 atoms in the *real* system) is created by adding a part of the backbone that forms a hydrogen bond with the amide NH group of the substrate peptide. Again, the new group is treated by QM′, while all the atoms from system T-2 are placed in the *model* system, see Figure 8.

Neglect of the surrounding protein makes the secondary structures unstable, and to avoid comparing different local minima, all calculations have been performed at the B3LYP/6-31G(d) optimized geometries of the full T-3 system. In this system, the relative energies of **INT1** and **INT2** are 27.3 and 12.6 kcal/mol, respectively. **INT1** is not a stationary point for the B3LYP/6-31G(d) optimization, and the structure is obtained by freezing the newly formed O−C distance at 1.513 Å (from ref 67). The relative energies are higher than expected for an enzymatic pathway and do not change significantly when applying a larger basis or a PCM solvent correction. It is possible that the reaction pathway is different in the present model than in the QM/MM free-energy perturbation calculations by Ishida and Kato,[67] but the exact reaction pathway is not critical for the comparison of QM and ONIOM(QM:QM′) results.

Results for the different B3LYP:DFTB models are given in Table 6. Results for AM1 and MM (Amber96) are given in the Supporting Information (Table S2). For **INT1**, the largest error (4.7 kcal/mol) comes from the treatment of the oxyanion hole with DFTB (T-2). At the B3LYP/6-31G(d) level, the addition of the oxyanion hole stabilizes **INT1** by 7.2 kcal/mol ($\Delta S^{high} = -7.2$ kcal/mol), which reflects one of the important enzymatic effects in the serine proteases. However, the DFTB layer gives a larger stabilization ($\Delta S^{low} = -11.9$ kcal/mol), leading to an ONIOM error of 4.7 kcal/mol.

The errors for the serine protease are much larger than those observed for a simple proton transfer reaction in which DFTB can well account for environmental effects of 23 kcal/mol with an error of only 1.2 kcal/mol.[47] The results illustrate that, as the environmental effect is calculated at the low-level only, not only must the low-level method be able to describe the electronic polarization effect but it must also be able to properly describe the changes in electronic structure of the reacting region. In DFTB, the electrostatic interactions are calculated using Mulliken charges,[43] and we therefore use Mulliken charges to discuss changes in charge distribution. In the reactant, DFTB assigns a Mulliken charge of −0.593 to the oxygen (O-1 in Figure 8), close to the value from B3LYP (−0.591), see Table 7. In **INT1**, DFTB assigns a much higher negative charge to the oxygen (−0.910) than the B3LYP calculation (−0.732), and it is not surprising that

Table 7. Mulliken Charges for the Backbone Oxo Group of Arg15B That Becomes an Oxyanion in **INT1**[a]

ONIOM	state	B3LYP (real)	DFTB (real)	B3LYP (model)	DFTB (model)
T-2	Reac	−0.591	−0.593	−0.575	−0.558
	INT1	−0.732	−0.910	−0.730	−0.915
	INT2	−0.529	−0.569	−0.510	−0.510

[a] The oxygen is labeled O-1 in Figure 8.

the stabilizing effect of the oxyanion hole is overestimated in the DFTB calculation. In **INT2**, the charge distributions are similar for B3LYP and DFTB, and ONIOM also gives smaller errors.

Full Protein QM:QM′:MM Models. To illustrate the applicability of the three-layer ONIOM (DFT:DFTB:MM) models for realistic system sizes, we designed an ONIOM model of trypsin that includes the entire protein and parts of the solvent shell. The initial structure for trypsin was taken from the PDB file 1TAW.[68] On the basis of a check of the structure with WHAT IF,[69] the terminal part of the Gln30 side chain was rotated. pK_a values from PropKa suggested all His residues were singly protonated.[70] His40 and His91 were assigned as His$_\varepsilon$, and His57 was assigned as His$_\delta$. The protein was solvated in a flexible water box of approximate dimension 60 × 57 × 74 Å^3 using periodic boundary conditions. Seven chloride ions and four sodium ions were added in random positions in the solvent to achieve neutrality. With the pure MM method using the Amber96 force field, the system was minimized using a conjugate gradient method for 5000 steps, followed by 0.5 ns equilibration at 298 K with a constrained backbone and 1 ns without constraints, using the program NAMD 2.6.[71] After equilibration, an initial QM:MM *real* system was selected by including all protein atoms and all solvent molecules with at least one atom within 5 Å of the protein, totaling 6327 atoms. The selection includes one charged sodium atom from the solvation shell located in the vicinity of residue Phe34B at a position 16.3 Å from the peptide bond that is cleaved. Atoms more than 15 Å from any atom in the initial 88-atom *model* system were kept frozen in the optimizations. The 88-atom *model* system includes selected parts of residues Ala56, His57, Asp102, Cys191, Gly192, Gly193, Asp194, and Ser195 in chain A and residues Arg15, Ala16, Met17, Gly36, and Gly37 in chain B. Three stationary points are included in the test: the reactant, the oxyanion structure (**INT1**), and the intermediate with a cleaved C−N bond (**INT2**).

The 6327 atom protein structure was initially optimized at the B3LYP:MM level with a *model* QM system size of 88 atoms, see Figure 9. The QM selection is similar to the 71-atom active-site model (T-3 *real* system) but includes hydrogen-bonding groups whose mobility caused problems with geometry optimization in the active-site model. From the optimized reactant structure, the serine residue is moved toward the carbonyl of the peptide, until the O−C distance is 1.513 Å. This distance is frozen when the structure is reoptimized to form **INT1**. **INT2** is formed from **INT1** by elongating the peptide C−N bond and then freely optimizing all reaction coordinates. In studies of reaction mechanisms, the different stationary points should connect along the same potential energy surface to avoid effects from artificial

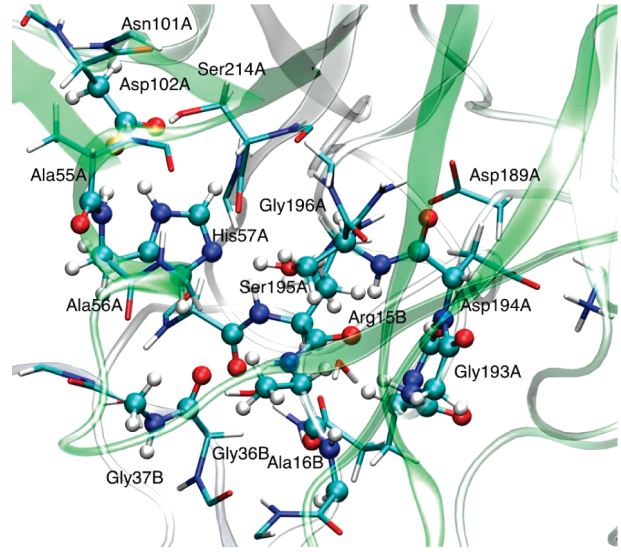


Figure 9. ONIOM protein model of trypsin with 88 atoms in the *model* system (ball-and-stick representation), 215 atoms in the *intermediate* system (ball-and-stick and stick representations), and the *real* system (full protein) in cartoon representation.

changes in protein geometry during the optimizations. This typically requires several iterations between reactant and product until there no longer are any conformational changes in the surrounding protein. However, in the present study, we are mainly interested in the performance of different methods applied to the same stationary state, and going from reactant to product in a single step should still give relevant results.

The results of the new 88-atom *model* system are similar to those of the 71-atom active-site model. As an example, the relative energy of **INT1** is 27.0 and 27.3 kcal/mol, respectively, see Table 8. In all the calculations, the geometries optimized at the two-layer B3LYP/6-31G(d):MM level with 88 QM atoms are used.

First, as shown in Table 8, the QM system is extended to 215 atoms, and the performance of DFTB(215) and B3LYP(88):DFTB(215) (division D in Table 8) is compared to B3LYP(215) (division B in Table 8). The performance of the DFTB layer can be assessed by comparing the ΔS values for the 215-atom subsystem minus the 88-atom subsystem (column 7 in Table 8). The deviations are +2.4 kcal/mol for **INT1** and −4.8 kcal/mol for **INT2**.

Next, we compare the effects of using DFTB in the 795-atom subsystem instead of MM. In the present example, this DFTB layer is made up of the residues closest to the active site, but they can also be chosen on the basis of an evaluation of the protein effects in a previous QM:MM calculation. A comparison of the B3LYP(215):DFTB(793):MM (division A) and B3LYP(215):MM (division B) models shows that in **INT1** there is a relatively small effect, and a small difference (1.5 kcal/mol) between the two methods, see column 8 in Table 8. However, for **INT2**, the difference between the DFTB and the MM description of the environmental effect is big ∼12 kcal/mol. The situation for B3LYP(88):DFTB(793):MM (division C) and B3LYP(88):DFTB(215):MM (division D) models as well as for DFTB(793):MM (division E) and

Table 8. The ONIOM Relative Energies and Their Components ΔE and ΔS (in kcal/mol, relative to the Reactant) for Different Two- and Three-Layer Models of Structures **INT1** and **INT2** of Trypsin[a]

	systems and methods[b]				ΔE[c]	ΔS[d]			
division	88	215	793	full	88	215−88	793−215	full−793	ONIOM[e]
					INT1				
A	B3LYP	B3LYP	DFTB	Amber	27.0	−4.3	1.2	5.0	29.0
B	B3LYP	B3LYP	Amber	Amber	27.0	−4.3	−0.3	5.0	27.5
C	B3LYP	DFTB	DFTB	Amber	27.0	−1.8	1.2	5.0	31.4
D	B3LYP	DFTB	Amber	Amber	27.0	−1.8	−0.3	5.0	29.9
E	DFTB	DFTB	DFTB	Amber	31.3	−1.8	1.2	5.0	35.7
F	DFTB	DFTB	Amber	Amber	31.3	−1.8	−0.3	5.0	34.2
					INT2				
A	B3LYP	B3LYP	DFTB	Amber	14.2	0.6	−0.8	2.7	16.7
B	B3LYP	B3LYP	Amber	Amber	14.2	0.6	−12.4	2.7	5.1
C	B3LYP	DFTB	DFTB	Amber	14.2	−4.3	−0.8	2.7	11.8
D	B3LYP	DFTB	Amber	Amber	14.2	−4.3	−12.4	2.7	0.3
E	DFTB	DFTB	DFTB	Amber	12.5	−4.3	−0.8	2.7	10.1
F	DFTB	DFTB	Amber	Amber	12.5	−4.3	−12.4	2.7	−1.4

[a] The full systems are formally divided into 88, 215, and 793 subsystems in Figure 9, and the method used for each subsystem is identified. ΔE is for the 88 atom system at the highest level, and ΔS is the difference between the two subsystems at the given level. [b] B3LYP is B3LYP/6-31G(d), and Amber is Amber96. [c] ΔE is the energy for the smallest 88 system at the highest level of the given division. [d] For instance, 1.2 in the row 4, column 8 represents $\Delta S^{793-215,\mathrm{DFTB}} = \Delta E^{793,\mathrm{DFTB}} - \Delta E^{215,\mathrm{DFTB}}$. [e] ONIOM energy is the sum of five contributions from column 6 to 9.

DFTB(215):MM (division F) is exactly the same as this effective is simply additive.

The large environmental effect at the MM level is partly due to changes in the structure, e.g., the orientation of a distant water (Wat39 in 1TAW numbering). These are simple artifacts of the optimization approach; i.e., different local MM minima are found for different intermediates. This would not be acceptable in a calculation of reaction energies, but as this is an ONIOM evaluation, no additional efforts were made to properly explore the MM energy landscape. From an ONIOM perspective, the most interesting data are the large differences between the DFTB and the MM description. Part of this difference probably comes from different relative energies for the artificial changes in protein geometry. Another effect is that the 793-atom DFTB subsystem allows transfer of charges and mutual polarization between the 215-atom subsystem and the 793-minus-215-atom layer, effects that are not included in the mechanical embedding MM method.[72]

IV. Discussion

The simple calculations of proton affinities for amino acids illustrate important points about the B3LYP:DFTB models. The DFTB layer significantly improves the results compared to the *model* systems and gives good results also when the environmental effects are large. However, errors are not systematic; i.e., they do not decrease as the size of the *model* system increases. For the tripeptide systems, errors are 1−2 kcal/mol even for small *model* systems. As a comparison, errors of a carefully parametrized frozen orbital scheme are ∼1 kcal/mol for cuts in the backbone of the peptide chain.[57] For reactions where the environmental effect is basis-set-dependent, e.g., proton affinities of negative residues, the DFTB layer gives better results when combined with a double-ζ calculation, at which level DFTB is parametrized, compared to the triple-ζ calculation. These large effects are due to well-known problems of describing negative ions with an insufficient basis set., e.g., positive eigenvalues of occupied orbitals. However, for systems where the environmental effect is relatively independent of the basis set, e.g., proton affinities of positive residues, the DFTB layer works well independent of the basis set used in the QM calculation.

The QM:QM′ model of the nonheme catalase mimic illustrates an important point. The ONIOM method fails when QM and QM′ give different descriptions of the electronic state of the transition metal. This is unrelated to the *model* selection and can only be fixed by choosing more appropriate computational methods. Prior to the use of a QM:QM′ model, it is thus important to compare the electronic structure of the proposed *model* system with QM and QM′. Although B3LYP/6-31G(d) and DFTB give different energies for spin splitting energies of transition metal compounds, the two methods will often describe the same spin state as the multiplicity of the QM′ calculation is assigned at the start of the calculation. Problems with different electronic states should therefore only occur in systems with several electronic states for a given multiplicity. In the present example, the ligand radical could couple ferromagnetically with intermediate-spin iron or antiferromagnetically to high-spin iron. In our parametrization study, spin splittings of 13 Fe compounds had a mean average deviation of 15.2 kcal/mol and a maximum deviation of 34.3 kcal/mol when DFTB results were compared to B3LYP/SDD+6/31G(d).[48] This error is in the range of the difference between GGA and hybrid DFT functionals. However, in ONIOM, the requirement is only that the low-level method gives the same state as the high-level method, and out of the 67 complexes in ref 48, B3LYP and DFTB predict the same spin state in 52 cases.

Even in systems without open-shell species, the active-site B3LYP:DFTB models show some significant errors. In these tests, the reference *real*,*high* calculations must be affordable, and the tested systems are therefore always smaller, and the errors larger, than expected for normal applications. However, the trypsin calculations show that

chemical accuracy (~1 kcal/mol) is not achieved even with relatively large QM models. For trypsin, these errors come from different descriptions of the charge distribution of the *model* system, e.g., of the oxyanion intermediate. This weakness of the mechanical embedding approach should be balanced against the savings in computational time.

When testing the applicability of a QM:QM′ model prior to its use, the first priority is to investigate if the reaction is qualitatively correctly described by the QM′ method. This can be done by comparing the electronic structure of the two methods, e.g., through population analysis, or by comparing changes in dipole moment. Another possibility is to make small ONIOM models and investigate the environmental effects of a single polarizing residue. If the effect of this residue is not qualitatively correct, the ONIOM model is not likely to give good results for any system selection.

More accurate results can be achieved by primarily using the QM:QM′ approach for optimizations. An important part of this paper is to show the possibility to fully optimize transition states with the QM:QM′ method. The energy can then be evaluated in a single-point calculation with a high-level QM method. For these calculations, the flexibility of the ONIOM method makes it possible to combine a larger basis set for the *model* system with a medium-sized basis set or low-cost method for the surrounding.

An interesting alternative to improve the description of electrostatic effects in QM:QM′ is the newly developed ONIOM(QM:QM′)−EE scheme.[73] In this scheme, the environmental effect at *real*,*low* is adjusted by the difference in response between QM and QM′ methods to a point charge environment obtained at the *real*,*QM′* level. QM:QM′−EE improves the DFT:HF results for several tested reactions.[74] As parts of the problem with the DFT:DFTB approach come from a different description of the reactive region, ONIOM−EE should improve the accuracy also with DFTB as the low-level method. However, the calculated environmental charges are geometry-dependent, and evaluation of QM:QM′−EE forces requires solutions of iterative coupled-perturbed equations, similar to coupled-perturbed Hartree−Fock.[73] This increases the cost of the *real*,*low* calculation approximately by a factor of 2, but it also means that not all methods can be directly used as low-layer methods. At the moment, QM:QM′−EE is only available with HF and DFT in the low layer.

We have also illustrated the use of a three-layer QM:QM′:MM model of an enzymatic system. Compared to standard QM/MM models, the scheme can be used in several different ways. In the trypsin model, parts of the MM system were replaced by DFTB, which in principle should lead to a better description of the environmental effect. It is also possible to reduce the size of the *model* system to allow for an improved description of the reactive region by correlated ab initio methods, e.g., CCSD(T) or MRCI.

Optimization of large three-layer models is still difficult. QM:MM optimizations are performed with microiterations where the MM system is optimized using a first-order algorithm. If a DFTB layer of ~1000 atoms is optimized together with the MM part, several hundred QM′ energy and gradient evaluations will be performed for each QM iteration. The computational time for the QM′ calculations will then become higher than the time required for the QM calculation. A second alternative is to optimize the QM′ layer together with the QM region in a second-order algorithm. The drawbacks are that the optimization algorithm becomes more costly, and more importantly, the number of macroiterations increases. This leads to an increase in the number of QM calculations required to reach geometry convergence. Higher efficiency might be reached with a hybrid technique that uses three different optimization levels, but such a scheme has yet to be implemented.

V. Conclusions

The ONIOM(QM:QM′) scheme with mechanical embedding is cost-efficient as it only requires a single QM evaluation at each geometry. The drawback is that all environmental effects are evaluated at the low (QM′) level, and the accuracy of the scheme depends on how well the low-level QM′ method describes the environmental effects and the changes in electron density during the reaction.

To illustrate the advantages and limitations of this method, we have applied ONIOM(B3LYP:DFTB) and three-layer ONIOM(B3LYP:DFT:MM) combinations to models of enzymes and enzyme mimics. Although the DFTB layer reduces a large part of the error in the underlying *model* calculations, remaining errors of several kilocalories per mole are not uncommon. The polarization effects are fairly well described using DFTB, but the QM and QM′ methods do not always describe the same electronic state throughout the reaction, causing some difficulties.

Use of the QM:QM′ model requires an in-depth understanding of both the QM and the QM′ method, and the applicability of the QM:QM′ scheme must be carefully investigated in each application. Separate calculations of the electronic structure using both QM and QM′ methods as well as a test of small ONIOM systems offer a reasonable way to test a QM:QM′ scheme without performing extensive benchmark tests.

The good performance of the DFTB method for geometries, together with the possibility to optimize transition states with the QM:QM′ scheme, makes it an excellent tool for exploration of geometries. Accurate energies can then be obtained by single-point calculations using a high-level method.

Acknowledgment. One of the authors (M.L.) acknowledges a Fukui Institute for Fundamental Chemistry Fellowship. The present work was in part supported by a CREST (Core Research for Evolutional Science and Technology) grant in the Area of High Performance Computing for Multiscale and Multiphysics Phenomena from the Japan Science and Technology Agency (JST). The computational resources at the Research Center for Computational Science at Institute for Molecular Science are gratefully acknowledged.

Supporting Information Available: Table S1 listing ONIOM energies for hydrolysis of N-methylazetidinone in mononuclear Zn-β-lactamase. Table S2 listing ONIOM energies for active-site model of peptide hydrolysis in trypsin.

PDB files for trypsin stationary points. Gaussian input files with ONIOM layer selections (88, 215, and 793 atoms) for trypsin. This information is available free of charge via the Internet at http://pubs.acs.org/.

References

(1) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227–249.

(2) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1986**, *7*, 718–730.

(3) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700–733.

(4) Lin, H.; Truhlar, D. G. *Theor. Chem. Acc.* **2007**, *117*, 185–199.

(5) Senn, H. M.; Thiel, W. In *Atomistic Approaches in Modern Biology: From Quantum Chemistry to Molecular Simulations*; Springer: New York, 2007; Vol. 268, pp 173−290.

(6) Maseras, F.; Morokuma, K. *J. Comput. Chem.* **1995**, *16*, 1170–1179.

(7) Humbel, S.; Sieber, S.; Morokuma, K. *J. Chem. Phys.* **1996**, *105*, 1959–1967.

(8) Matsubara, T.; Maseras, F.; Koga, N.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 2573–2580.

(9) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19357–19363.

(10) Dapprich, S.; Komaromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J. *J. Mol. Struct. Theochem* **1999**, *462*, 1–21.

(11) Vreven, T.; Morokuma, K. *J. Comput. Chem.* **2000**, *21*, 1419–1432.

(12) Vreven, T.; Byun, K. S.; Komaromi, I.; Dapprich, S.; Montgomery, J. A.; Morokuma, K.; Frisch, M. J. *J. Chem. Theory Comput.* **2006**, *2*, 815–826.

(13) Derat, E.; Bouquant, J.; Humbel, S. *J. Mol. Struct. Theochem* **2003**, *632*, 61–69.

(14) Morokuma, K.; Wang, Q. F.; Vreven, T. *J. Chem. Theory Comput.* **2006**, *2*, 1317–1324.

(15) Froese, R. D. J.; Morokuma, K. *Chem. Phys. Lett.* **1999**, *305*, 419–424.

(16) Vreven, T.; Morokuma, K. *J. Phys. Chem. A* **2002**, *106*, 6167–6170.

(17) Morokuma, K. *Bull. Chem. Soc. Jpn.* **2007**, *80*, 2247–2261.

(18) Lundberg, M.; Morokuma, K. *In Multi-scale Quantum Models for Biocatalysis: Modern Techniques and Applications*, Eds. Lee, T.-S. York, D. M., Springer Verlag 2009.

(19) Stern, H. A.; Kaminski, G. A.; Banks, J. L.; Zhou, R. H.; Berne, B. J.; Friesner, R. A. *J. Phys. Chem. B* **1999**, *103*, 4730–4737.

(20) Stern, H. A.; Rittner, F.; Berne, B. J.; Friesner, R. A. *J. Chem. Phys.* **2001**, *115*, 2237–2251.

(21) Morales, J.; Martinez, T. J. *J. Phys. Chem. A* **2004**, *108*, 3076–3084.

(22) Gascon, J. A.; Leung, S. S. F.; Batista, E. R.; Batista, V. S. *J. Chem. Theory Comput.* **2006**, *2*, 175–186.

(23) Chen, J. H.; Martinez, T. J. *Chem. Phys. Lett.* **2007**, *438*, 315–320.

(24) Chen, J. H.; Hundertmark, D.; Martinez, T. J. *J. Chem. Phys.* **2008**, *129*, 214113.

(25) Xie, W. S.; Song, L. C.; Truhlar, D. G.; Gao, J. L. *J. Chem. Phys.* **2008**, *128*, 234108.

(26) Ren, P. Y.; Ponder, J. W. *J. Phys. Chem. B* **2003**, *107*, 5933–5947.

(27) Xie, W. S.; Song, L. C.; Truhlar, D. G.; Gao, J. L. *J. Phys. Chem. B* **2008**, *112*, 14124–14131.

(28) Xie, W. S.; Gao, J. L. *J. Chem. Theory Comput.* **2007**, *3*, 1890–1900.

(29) Zhang, Y.; Lin, H. *J. Chem. Theory Comput.* **2008**, *4*, 414–425.

(30) Yang, W. T.; Lee, T. S. *J. Chem. Phys.* **1995**, *103*, 5674–5678.

(31) Wesolowski, T. A.; Weber, J. *Chem. Phys. Lett.* **1996**, *248*, 71–76.

(32) Nakano, T.; Kaminuma, T.; Sato, T.; Akiyama, Y.; Uebayasi, M.; Kitaura, K. *Chem. Phys. Lett.* **2000**, *318*, 614–618.

(33) Cortona, P. *Phys. Rev. B* **1991**, *44*, 8454–8458.

(34) Henderson, T. M. *J. Chem. Phys.* **2006**, *125*, 014105.

(35) Huang, P.; Carter, E. A. *J. Chem. Phys.* **2006**, *125*, 084102.

(36) Huang, P.; Carter, E. A. *Annu. Rev. Phys. Chem.* **2008**, *59*, 261–290.

(37) Gogonea, V.; Westerhoff, L. M.; Merz, K. M. *J. Chem. Phys.* **2000**, *113*, 5604–5613.

(38) Cui, Q.; Guo, H.; Karplus, M. *J. Chem. Phys.* **2002**, *117*, 5617–5631.

(39) Svensson, M.; Humbel, S.; Morokuma, K. *J. Chem. Phys.* **1996**, *105*, 3654–3661.

(40) Vreven, T.; Morokuma, K. *J. Chem. Phys.* **1999**, *111*, 8799–8803.

(41) Tschumper, G. S.; Morokuma, K. *J. Mol. Struct. Theochem* **2002**, *592*, 137–147.

(42) Porezag, D.; Frauenheim, T.; Kohler, T.; Seifert, G.; Kaschner, R. *Phys. Rev. B* **1995**, *51*, 12947–12957.

(43) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. *Phys. Rev. B* **1998**, *58*, 7260–7268.

(44) Kohler, C.; Seifert, G.; Gerstmann, U.; Elstner, M.; Overhof, H.; Frauenheim, T. *Phys. Chem. Chem. Phys.* **2001**, *3*, 5109–5114.

(45) Frauenheim, T.; Seifert, G.; Elstner, M.; Hajnal, Z.; Jungnickel, G.; Porezag, D.; Suhai, S.; Scholz, R. *Phys. Status Solidi B* **2000**, *217*, 41–62.

(46) Zheng, G. S.; Irle, S.; Morokuma, K. *Chem. Phys. Lett.* **2005**, *412*, 210–216.

(47) Zheng, G. S.; Lundberg, M.; Jakowski, J.; Vreven, T.; Frisch, M. J.; Morokuma, K. *Int. J. Quantum Chem.* **2009**, *184*, 1–1854.

(48) Zheng, G. S.; Witek, H. A.; Bobadova-Parvanova, P.; Irle, S.; Musaev, D. G.; Prabhakar, R.; Morokuma, K.; Lundberg, M.; Elstner, M.; Köhler, C.; Frauenheim, T. *J. Chem. Theory Comput.* **2007**, *3*, 1349–1367.

(49) Iordanov, T. D. *THEOCHEM* **2008**, *850*, 152–159.

(50) Elstner, M.; Jalkanen, K. J.; Knapp-Mohammady, M.; Frauenheim, T.; Suhai, S. *Chem. Phys.* **2001**, *263*, 203–219.

(51) Elstner, M. *Theor. Chem. Acc.* **2006**, *116*, 316–325.

(52) Otte, N.; Scholten, M.; Thiel, W. *J. Phys. Chem. A* **2007**, *111*, 5751–5755.

(53) Nakatani, N.; Hasegawa, J. Y.; Nakatsuji, H. *J. Am. Chem. Soc.* **2007**, *129*, 8756–8765.

(54) Wanko, M.; Hoffmann, M.; Frauenheim, T.; Elstner, M. *J. Phys. Chem. B* **2008**, *112*, 11462–11467.

(55) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian Development Version*; Gaussian, Inc.: Wallingford, CT, 2009.

(56) Siegbahn, P. E. M. *J. Comput. Chem.* **2001**, *22*, 1634–1645.

(57) Murphy, R. B.; Philipp, D. M.; Friesner, R. A. *J. Comput. Chem.* **2000**, *21*, 1442–1457.

(58) Zhou, H. Y.; Tajkhorshid, E.; Frauenheim, T.; Suhai, S.; Elstner, M. *Chem. Phys.* **2002**, *277*, 91–103.

(59) Elstner, M.; Cui, Q.; Munih, P.; Kaxiras, E.; Frauenheim, T.; Karplus, M. *J. Comput. Chem.* **2003**, *24*, 565–581.

(60) Xu, D. G.; Guo, H.; Cui, Q. *J. Phys. Chem. A* **2007**, *111*, 5630–5636.

(61) Xu, D.; Guo, H.; Cui, G. *J. Am. Chem. Soc.* **2007**, *129*, 10814–10822.

(62) Diaz, N.; Suarez, D.; Merz, K. M. *J. Am. Chem. Soc.* **2001**, *123*, 9867–9879.

(63) Paschke, J.; Kirsch, N.; Korth, H. G.; de Groot, H.; Sustmann, R. *J. Am. Chem. Soc.* **2001**, *123*, 11099–11100.

(64) Wang, X.; Li, S. H.; Jiang, Y. S. *Inorg. Chem.* **2004**, *43*, 6479–6489.

(65) Sicking, W.; Korth, H. G.; Jansen, G.; de Groot, H.; Sustmann, R. *Chem.—Eur. J.* **2007**, *13*, 4230–4245.

(66) Oliveira, K. M. T.; Trsic, M. *THEOCHEM* **2001**, *539*, 107–117.

(67) Ishida, T.; Kato, S. *J. Am. Chem. Soc.* **2003**, *125*, 12035–12048.

(68) Scheidig, A. J.; Hynes, T. R.; Pelletier, L. A.; Wells, J. A.; Kossiakoff, A. A. *Protein Sci.* **1997**, *6*, 1806–1824.

(69) Vriend, G. *J. Mol. Graphics* **1990**, *8*, 52–56.

(70) Li, H.; Robertson, A. D.; Jensen, J. H. *Proteins: Struct., Funct., Bioinf.* **2005**, *61*, 704–721.

(71) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781–1802.

(72) Lundberg, M.; Kawatsu, T.; Vreven, T.; Frisch, M. J.; Morokuma, K. *J. Chem. Theory Comput.* **2009**, *5*, 222–234.

(73) Hratchian, H. P.; Parandekar, P. V.; Raghavachari, K.; Frisch, M. J.; Vreven, T. *J. Chem. Phys.* **2008**, *128*, 034107.

(74) Parandekar, P. V.; Hratchian, H. P.; Raghavachari, K. *J. Chem. Phys.* **2008**, *129*, 145101.

CT100029P