JCTC Journal of Chemical Theory and Computation

Protein Free Energy Corrections in ONIOM QM:MM Modeling: A Case Study for Isopenicillin N Synthase (IPNS)

Tsutomu Kawatsu,[†] Marcus Lundberg,[†] and Keiji Morokuma*[,†,‡]

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

Received September 27, 2010

Abstract: The protein environment can have significant effects on the enzyme catalysis even though the reaction occurs locally at the reaction center. In this paper, we describe an efficient scheme that includes a classical molecular dynamics (MD) free-energy perturbation (FEP) correction to the reaction energy diagram, as a complement to the protein effect obtained from static ONIOM (QM:MM) calculations. The method is applied to eight different reaction steps, from the O_2-bound reactant to formation of a high-valent ferryl-oxo intermediate, in the nonheme iron enzyme isopenicillin N synthase (IPNS), for which the QM:MM energy diagram has previously been computed [Lundberg, M. et al. *J. Chem. Theory Comput.* **2009**, *5*, 220−234]. This large span of the reaction coordinate is covered by dividing each reaction step into microsteps using a virtual reaction coordinate, thus only requiring ONIOM information about the stationary points themselves. Protein effects are important for C−H bond activation and heterolytic O−O bond cleavage because both these two steps involve charge transfer, and compared to a static QM:MM energies, the dynamics of the protein environment changes the barrier for O−O bond cleavage by several kcal/mol. The origin of the dynamical contribution is analyzed in two terms, the geometrical effect caused by the change in average protein geometry (compared to the optimized geometry) in the room temperature MD simulation with the solvent, and the statistical (entropic) effect resulting from fluctuations in the interactions between the active site and the protein environment. These two effects give significant contributions in different steps of the reaction.

1. Introduction

When modeling enzymatic reactions, it is common to separate the reactivity of the active site from the effects of the surrounding protein matrix. This approximation seems especially valid for transition metal enzymes, for which the activity of biomimetic complexes[1] indicates that the reactivity mainly depends on the electronic structure of the metal center.[2] However, enzymes with very similar active sites catalyze different reactions, and an explicit description of the protein environment is necessary to fully understand the reaction mechanism, relative reaction rates and substrate selectivity.

QM/MM (quantum mechanics/molecular mechanics) models take advantage of the separation between reaction active site and protein environment by treating these regions at different levels of theory.[3] Our group has developed the ONIOM multiscale method that calculates the total energy of the molecular system by an extrapolation scheme including different QM and MM calculations.[4−9] The QM:MM label

* To whom correspondence should be addressed: E-mail: morokuma@fukui.kyoto-u.ac.jp. Telephone: +81-75-711-7843.

[†] Kyoto University.

[‡] Emory University.

10.1021/ct1005592
Published on Web 12/29/2010

separates ONIOM from standard QM/MM methods that employ additive schemes. The interaction between the QM part and the MM part can be included in the calculations either classically by mechanical embedding (ONIOM-ME) or semiclassically by electronic embedding (ONIOM-EE).[9] We have previously used the ONIOM QM:MM method to describe the protein effects on several metalloenzyme reactions.[10] Recently, we have employed an advanced algorithm, the "fully coupled macro/micro-iterative" optimization scheme,[11] to efficiently locate transition states in complex molecular systems, specifically in mammalian glutathione peroxidase,[12] isopenicillin N synthase (IPNS),[13] and methylmalonyl-CoA mutase.[14,15] The protein environment influences the description of the reactivity, but the effect on the calculated energy barriers varies significantly between different enzymes, significantly lowering the barrier in methylmalonyl-CoA mutase, while having only a modest effect in glutathione peroxidase.

However, describing a large system with optimization techniques requires special care to avoid artificial changes in geometry that can lead to large errors in relative energies.[9] Static methods also cannot describe situations where the environment changes during the chemical reaction, for example, new alignment of side chains or solvent water, thermal fluctuations, or large-scale protein motions. Because of the lack of geometric polarization, the static approach may overestimate electrostatic effects. In the present study, we replace the static interactions between protein and QM region by classical free-energy corrections from dynamical sampling of millions of protein configurations. QM/MM approaches with free-energy perturbation (FEP) have previously been used to describe reactions in both protein and solvent.[16−21] To separate the present approach from others efforts in the area, we use the description QM:[MM-FEP] for the ONIOM QM:MM approach, where the effect of the MM layer is described by free-energy perturbation.[22]

One of the main objectives of the present method is the capability to estimate the dynamic effects on the reaction energy profile of complicated enzymatic reactions, for example, a multistep redox reaction in a transition metal enzyme. The difference in electronic structure and nuclear coordinates between two stationary states can be large, so in FEP each reaction step is divided into several intermediate points, for example, by following selected reaction coordinates (typically bond distances) or the intrinsic reaction coordinate (IRC). However, transition-metal systems have complicated multidimensional reaction coordinates, so to avoid a detailed mapping, we adopt the standard alchemical FEP technique that only requires information about the initial and the final state.[23] Intermediate points are generated by a virtual reaction coordinate that gradually mixes the initial and the final state. The required information about these states is obtained by full QM:MM optimizations of all stationary points, including transition states, using the fully coupled Hessian algorithm.[9]

Transition-metal systems require relatively expensive QM methods, for example, hybrid DFT. We therefore freeze the geometry of the QM part and perform the FEP calculations with fully classical samplings (ONIOM-ME). In this approximation the QM energy (for a given QM geometry) does not depend on the protein structure, so only one calculation of the QM wave function is required for each geometry. This approximation is similar to the QM/MM-FE method,[24,25] and calculations by Rod et al.,[18,26] show that this method differs by less than 3 kcal/mol from their more elaborate QTCP method. In this context, it must be kept in mind that modeling of transition metal reactions is a difficult task and that the inherent error in the QM treatment has been estimated to be 3−5 kcal/mol.[2]

We apply the QM:MM method with free-energy corrections to the nonheme iron enzyme isopenicillin N synthase that catalyzes the formation of isopenicillin N (IPN), a key reaction in penicillin synthesis that is still used in large-scale production.[27] Our previous QM and QM:MM studies[13,28,29] identified 19 intermediates and transition states for the reaction leading from the ACV substrate to the IPN product. For the free-energy treatment, we selected the first half of this reaction, C−H bond activation from the iron-bound dioxygen species, followed by heterolytic O−O bond cleavage to form a ferryl-oxo species (9 stationary points). Of the two alternative mechanisms for Fe(IV)−oxo formation previously investigated, only the "ligand donor" mechanism is chosen here, partly because it shows larger protein effects.

In the following sections, we first describe the computational details of the QM:[MM-FEP] method, followed by a presentation of the IPNS free energy diagram and a discussion of how it differs from a potential energy diagram calculated with the standard QM:MM optimization method.

2. Methods of Computation

2.1. Free Energy in the ONIOM QM:MM Scheme. The relative energy for a standard ONIOM QM:MM calculation is obtained as follows:[6]

$$\Delta E_{\mathrm{ONIOM}} = \Delta E_{\mathrm{QM}}^{\mathrm{model}} + \Delta E_{\mathrm{MM}}^{\mathrm{real}} - \Delta E_{\mathrm{MM}}^{\mathrm{model}} \qquad (1)$$

where real includes all atoms in the system and model includes the selected reaction center, with hydrogen link atoms for truncated covalent bonds. In mechanical embedding (ME), the interaction between model and real system is described at the low level of theory of the real system. The QM:MM-ME energy thus includes a QM-level description of the relative energy, and an MM-level description of the protein effect on the relative energy. An alternative to mechanical embedding is electronic embedding (EE), where the QM-MM interactions are evaluated semiclassically by including the MM point charges in the QM calculation of the model system. For a free-energy calculation, the EE approach becomes very expensive as it requires a new QM calculation for each position of the surrounding atoms. We take advantage of the speed of the ME approximation and perform fully classical free-energy calculations.

In our QM:[MM-FEP] approximation, the free energy difference between two states is calculated as

$$\Delta F_{\mathrm{ONIOM}} = \Delta E_{\mathrm{ONIOM}} + \Delta f_{\mathrm{QM}}^{\mathrm{model}} + \Delta F_{\mathrm{FEP}}^{\mathrm{int/NB}} - \Delta E_{\mathrm{MM}}^{\mathrm{int/NB}} \qquad (2)$$

where Δf_{QM}^{model} is the free energy correction of the model system obtained from a Hessian calculation using the harmonic oscillator approximation. In the IPNS QM:MM potential diagram this term has been determined from calculations on an active-site QM-only model.[28] $\Delta F_{FEP}^{int/NB}$ and $\Delta E_{MM}^{int/NB}$ are the classical nonbonded interactions between the QM model system and the protein environment, evaluated using the FEP method and the static ONIOM-ME approach, respectively. By the use of the third and fourth terms, we replace the nonbonded interactions in standard ONIOM by the free energy of the model system and additionally solvent effect.

Note that the free-energy contributions in QM:[MM-FEP] does not explicitly include the protein−protein interactions. The total free energy of the protein fluctuates on a scale much larger than the energy differences between two stationary points, and it is challenging to converge the total energy. These interactions are still taken into account in the calculations, because the low-energy protein configurations dominate the calculation of nonbonded interactions between the model system center and the protein environment.

Free-energy calculations are performed with the real system at the low (MM) level with the coordinates of the model system frozen during the simulations. We have neglected the difference between terms including the hydrogen link atoms in Δf_{QM}^{model} compared to the free energy of the original covalent bonds that were truncated to form the model system. This effect is difficult to estimate correctly in the FEP approach,[30] but the effect is not likely to change significantly during the reaction and is therefore neglected. We have also neglected the cross term between the free energy of the QM model system and the protein environment, that is, how the protein environment affects the QM vibrations and vice versa.

2.2. Free-Energy Perturbation Method. In the alchemical FEP technique, an arbitrary number of intermediate points can be created by increasingly mixing two stationary points using a dual-topology method.[31] Here, we use λ_i, a virtual reaction coordinate that changes from 0 to 1 with increasing i. We run the molecular dynamics (MD) simulation between X (initial) and Y (final) states along the path of virtual intermediate states. The MD Hamiltonian is

$$H_{MD}^{i}(X,Y;R) = H_{env}(R) + (1-\lambda_i)H_X(R) + \lambda_i H_Y(R) \quad (3)$$

where $H_{env}(R)$ is the Hamiltonian excluding the QM model system, as well as the interactions between the model system and protein environment. $H_X(R)$ and $H_Y(R)$ are Hamiltonians of the model system in the two stationary states, X and Y, including their interaction terms with the protein environment. $R \equiv R(X,Y;\lambda_i)$ are the atomic coordinates of the protein environment, obtained by the molecular dynamics using the Hamiltonian $H_{MD}^{i}(X,Y;R)$. As the set of $(X,Y;\lambda_i)$ defines the Hamiltonian of a mixed chemical state, only information about initial and final states are required.

For the calculations of free energy, we define H_X^{NB} and H_Y^{NB} that exclude the bonded interaction terms (for link atoms) between the model system (redox center) and the protein environment from H_X and H_Y. The change in free energy $\Delta F_{FEP}^{int/NB}$ is the sum of free energy differences along the virtual reaction path

$$\Delta F_{FEP}^{int/NB}(X,Y) = \sum_{i=1}^{n} \Delta F_{i,i+1}(X,Y) \quad (4)$$

where the free energy difference for each virtual step is calculated as

$$\Delta F_{i,i+1}(X,Y) = -k_B T \ln\left\langle \exp\left\{ \frac{H_X^{NB}[R(X,Y;\lambda_i)] - H_Y^{NB}[R(X,Y;\lambda_i)]}{k_B T}(\lambda_{i+1}-\lambda_i) \right\} \right\rangle_i \quad (5)$$

Here, k_B is the Boltzmann constant, and T is the temperature. Where we used the energy change in the reaction coordinate from λ_i to λ_{i+1}, $H_{MD}^{i+1} - H_{MD}^{i} = -(H_X - H_Y)(\lambda_{i+1} - \lambda_i) \approx (H_X^{NB} - H_Y^{NB})(\lambda_{i+1} - \lambda_i)$.

From the FEP calculations we obtain the part of the free energy caused by the dynamics of the protein environment (so-called dynamical contribution, thereafter) $\Delta F_{FEP}^{int/NB}(X,Y) - \Delta E_{MM}^{int/NB}(X,Y)$ for each pair of stationary points, X and Y. For convenience we hereafter use the notations $\Delta F_{XY}^{FEB} \equiv \Delta F_{FEP}^{int/NB}(X,Y)$ and $\Delta E_{XY}^{opt} \equiv \Delta E_{MM}^{int/NB}(X,Y)$.

2.3. Geometrical and Statistical Effects. To better understand the origin of the QM:[MM-FEP] dynamical effects on the reaction energy diagram, we separate the contributions to the free energy into two parts: the *geometrical effect* and the *statistical effect*. The geometrical effect is the result of a change in average protein geometry in the room temperature MD simulation compared to the optimized structure. The statistical effect comes from fluctuations around the average geometry, because favorable protein geometries, that is, those that represent low-energy pathways, give larger contributions in the calculation of the free energy for a reaction step. ΔF_{XY}^{FEB} is the difference of the free energy interaction between state X and Y and includes all dynamical effects, while ΔE_{XY}^{opt} is the difference in protein interaction between optimized geometries, and excludes all dynamical effects. Here, we define a value that includes the geometrical effect, but, not the statistical effect. For this purpose we investigated the QM-MM interaction difference between $\lambda = 0$ (state X) and 1 (state Y) in MD simulations with the state-X geometry and charges. The interaction energy changes as

$$\Delta H_{XY}(X) \equiv H_X^{NB}(X) - H_Y^{NB}(X) \quad (6)$$

where $H_X^{NB}(X) \equiv H_X^{NB}(R(X,Y,0)) = H_X^{NB}(R(Y,X,1))$ MM Hamiltonians of the nonbonded interaction between the QM part and protein environment in the MD simulation were defined in section 2.2. $H_Y^{NB}(X) \equiv H_Y^{NB}(R(X,Y,0)) = H_Y^{NB}(R(Y,X,1))$ is written in same indices, but the interaction is between the state-Y QM part and the state-X protein environments. Because state-X protein geometry is not the optimized static geometry, $\Delta H_{XY}(X)$ is different from ΔE_{XY}^{opt}. The difference is the geometrical effect. $\Delta H_{XY}(X)$ includes the geometrical effect for each geometry, but not the statistical effect. The average of ΔH_{XY} then includes the average geometrical effect for the reaction barrier between states X and Y. We define the average as

$$\langle \Delta H \rangle_{XY} \equiv \langle [\Delta H_{XY}(X) - \Delta H_{YX}(Y)]/2 \rangle \quad (7)$$

We use the values $\langle \Delta H \rangle_{XY} - \Delta E_{XY}^{\mathrm{opt}}$ to estimate the geometrical contribution and the remaining free-energy contribution, $\Delta F_{XY}^{\mathrm{FEP}} - \langle \Delta H \rangle_{XY}$ is the statistical effect. We must notice that because $\langle \Delta H \rangle_{XY}$ includes only the situation of $\lambda = 0$ and 1, the comparison between $\langle \Delta H \rangle_{XY}$ and $\Delta F_{XY}^{\mathrm{FEP}}$ is inconsistent. However, when states X and Y are enough close, $\langle \Delta H \rangle_{XY}$ is reasonable approximation of the average potential energy shift from state X to Y that excludes the statistical effect.

For discussing the details, we partition the geometrical effect into contributions from each single residue i, $\Delta\Delta H_{XY}^{i}(X) \equiv \Delta H_{XY}^{i}(X) - \Delta E_{XY}^{\mathrm{opt},i}$ for an average geometry of state-X MD simulation, where $\Delta H_{XY}^{i}(X)$ and $\Delta E_{XY}^{\mathrm{opt},i}$ are the contributions of residue i to $\Delta H_{XY}(X)$ and $\Delta E_{XY}^{\mathrm{opt}}$, respectively.

The fluctuation of $\Delta H_{XY}(X)$ can correlate to the statistical effect. To investigate the fluctuation, we define the standard deviation s_{XY} as

$$s_{XY} \equiv \sqrt{\langle [\Delta H_{XY}(X)]^2 \rangle - \langle \Delta H_{XY}(X) \rangle^2} \quad (8)$$

We also define s_{YX} for the backward reaction from state Y to state X in the state-Y protein environment. We calculate the standard deviation (s_{XY}^{i}) of the single residue contribution to the interaction change ($\Delta H_{XY}^{i}(X)$) as

$$s_{XY}^{i} \equiv \sqrt{\langle [\Delta H_{XY}^{i}(X)]^2 \rangle - \langle \Delta H_{XY}^{i}(X) \rangle^2} \quad (9)$$

2.4. Computational Model. The protein setup and the ONIOM system are described in detail in reference.[13] Here we use the small 65-atom model system including Fe, a water ligand, selected parts of the three amino acids His214, Asp216, His270, and the reactive part of the substrate, see Figure 1. The small size of the model system increases the protein effects and makes it easier to evaluate the difference of the static and dynamic approaches. QM calculations were performed with the density functional B3LYP. The 6-31G(d) basis set was used for the geometry optimizations and Hessian calculations, while 6-311+G(d,p) was used for energy evaluations.

The classical nonbonded interactions between model and real system depend on the van der Waals parameters and the assigned point charges. Atoms outside the model system were assigned parameters from the Amber94 force field[34] to be able to compare with the previous QM:MM calculations.[13] Atoms in the model system are assigned point charges from RESP[35] calculations of the model system for each stationary point, using the Gaussian[36] standard geometry for the ESP calculations and the Antechamber module of Amber[37] for the first step of the RESP fitting, see Supporting Information. Charges for the part of the substrate that is not included in the model system were assigned from a calculation in the reactant state and were not changed during the reaction.

2.5. Simulation Details. We started the free-energy calculations based on the optimized QM:MM geometries of the stationary states as reported in ref 13. The protein was placed in an approximately 80 × 68 × 57 Å^3 water box including ~7700 TIP3P water molecules and 11 sodium ions

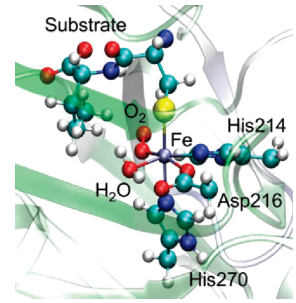

Figure 1. QM:MM model system of isopenicillin N synthase (QM atoms in ball and stick). All protein illustrations in the present paper are prepared using the VMD program.[32,33]

with periodic boundary condition. Simulations were run using the NAMD molecular dynamics program.[38,39] Five thousand steps of energy minimization and 1.5 ns equilibration were applied at each stationary point before the start of the FEP calculation. All simulations used a 1 fs time step. The temperature was controlled at 298 K using a Langevin thermostat every 5 fs.

We split each reaction step into 24 virtual steps (25 states) using the virtual λ coordinate ($\lambda_i = 0$, 0.001, 0.01, 0.05, 0.1, ..., 0.9, 0.95, 0.99, 0.999, 1.0). We use an unevenly spaced reaction coordinate to reduce the impact of steric clashes when the final state is turned on. There are nine stationary points along the reaction path, and the total number of intermediate states is 193. Each virtual step includes 100 ps equilibrium followed by 500 ps sampling simulations. Molecular dynamics is performed with the Hamiltonian for state i, $H_{\mathrm{MD}}^{i}(X,Y;R)$, and the difference in nonbonded interaction energies between two topologies (i and $i + 1$) is evaluated each 10 fs.

Average values of the exponential and statistical errors were estimated using the bootstrap method.[40,41] Our program uses a subroutine for the inverse error function calculation in Ooura's Mathematical Software package.[42] For each reaction step, we ran simulations both in forward ($\lambda = 0$ to 1) and backward ($\lambda = 1$ to 0) directions. The final result is the average of these two values, and the error bars are determined by the difference of the two simulations. The method presents statistically correct estimation of an average value in biased sampling.

3. Results and Analysis

3.1. QM:MM-ME and -EE Potential Energy Profiles. The use of a fully classical mechanical embedding (QM:MM-ME) potential is critical in our method as it allows for longer sampling times and better convergence of the FEP calculations. To check how the classical approximation of the protein−core interactions matches the semiclassical electronic embedding approach, we calculated the static energy diagram using both methods, see Figure 2. All comparisons are made with the same protein geometry for each stationary point, and the energies in Figure 2, therefore, represent situations without any geometric relaxation of either core or protein. This assumption overestimates the electrostatic effects and the difference between the methods, but still serves a purpose for a general comparison. The present

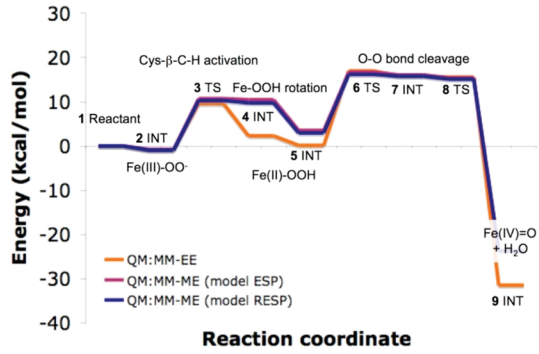

Figure 2. Comparison between ONIOM-EE and ONIOM-ME with charges (ESP and RESP) of the model system updated at every stationary point.

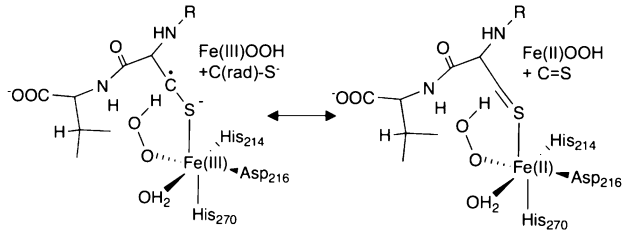

Figure 3. Resonance structures for stationary point **4**. The difference between the left and the right structure is an electron transfer from substrate to iron.

ONIOM-ME energy diagram is different from the energy diagram in ref 13. In the present study, the model charges are updated at each stationary point, but in the previous study the MM parameters were constant to provide a continuous energy surface for optimization of a large number of stationary points, including transition states. To facilitate the discussion in this section and later, each stationary point is described by a number in bold, representing the order it appeared along the reaction energy diagram. For details see ref 28.

Counting all stationary points, the mean absolute deviation between ME and EE approximations is 2.2 kcal. Major deviations between the two methods appear for stationary points **4** (iron-bound peroxide) and **9** (ferryl-oxo + water). In the first case, **4**, two alternative electronic structures can be drawn, see Figure 3, with the difference being a charge transfer between substrate and iron. This charge transfer leads to significant electrostatic repulsion from the surrounding protein, and while the ME calculation indicates a complete charge transfer (no spin population on the substrate), in the EE calculation there remains some unpaired spin population on the substrate carbon (−0.14). The difference in electron density between ME and EE can be ascribed as a polarization effect of the surrounding protein. In the second case, **9**, heterolytic O−O bond cleavage leads to release of water from the active site, and new direct interactions between QM and MM atoms. Fortunately these effects are not essential for the modeling of a reaction mechanism because product formation is exothermal and the energy of reaction does not affect the barriers of the proceeding steps of the reaction.

If we only compare the barrier heights of the three transition states, from state **2** to **3**, **5** to **6**, and **7** to **8**, the mean absolute deviation is only 0.67 kcal/mol. Although the classical approximation shows differences compared to the semiclassical approximation for certain steps, the potential energy diagrams for the present reaction are mostly reasonable. We therefore employ the classical charges for our FEP calculations.

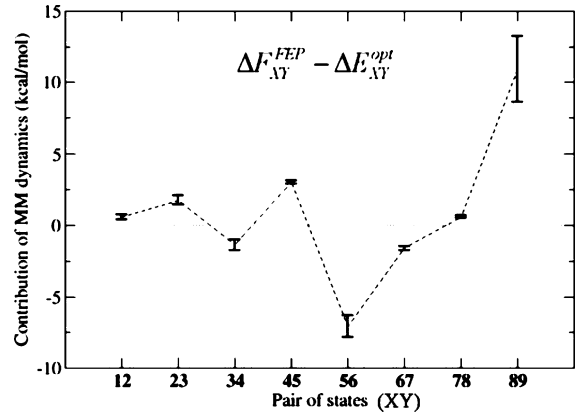

Figure 4. Contribution of MM dynamics, relative to a static QM:MM description, on the relative energies of neighboring stationary points *X* and *Y*.

3.2. Dynamical Contribution to the Free-Energy Diagram. The dynamical contributions to the free-energy diagram $\Delta F_{XY}^{\mathrm{FEP}} - \Delta E_{XY}^{\mathrm{opt}}$ are shown in Figure 4. In this figure each reaction step is written as a pair of initial (*X*) and final state (*Y*). Effects exceed 3 kcal/mol for pairs $XY =$ **45**, **56**, and **89** and must, therefore, be considered significant. The transition from state **4** to state **5** is a rotation of the peroxide formed after C−H bond activation, a step, where there is a large change in the active site geometry. The transition from **5** to **6** corresponds to an electron transfer from iron to an antibonding π^* O−O orbital, a step with large electrostatic effects from the protein environment, see discussion in ref 13. The transition from **8** to **9** is the release of H_2O from the reaction center after O−O bond cleavage, and the QM water molecule makes several new hydrogen bonds with MM residues and water molecules. The error bars for the FEP calculations are relatively small, with the exception of pair **89** where the released water molecule sometimes has strong steric interactions with the MM atoms as the virtual coordinate λ changes from 0 to 1.

The calculated free-energy diagram is shown in Figure 5. The blue line is the static QM:MM result (with updated RESP charges) and the red line is the QM:[MM-FEP] result. The QM:[MM-FEP] results represent a frozen core and a flexible protein, and to make a relevant comparison, the ONIOM-ME results are obtained with the same core geometry and with the protein reoptimized after the RESP charges had been updated.

Compared to the static ONIOM-ME description, the free-energy description leads to an increase in the C−H activation barrier (state **2** to **3**) and a decrease in the barrier for O−O bond activation by significantly stabilizing state **6** relatively to **5**. For evaluation of the reaction mechanism, the most significant difference is the predicted rate-limiting step. The ONIOM-ME (RESP) diagram suggests that O−O bond activation is rate-limiting. The free-energy approach gives a higher barrier for C−H bond activation step (**2** to **3**)

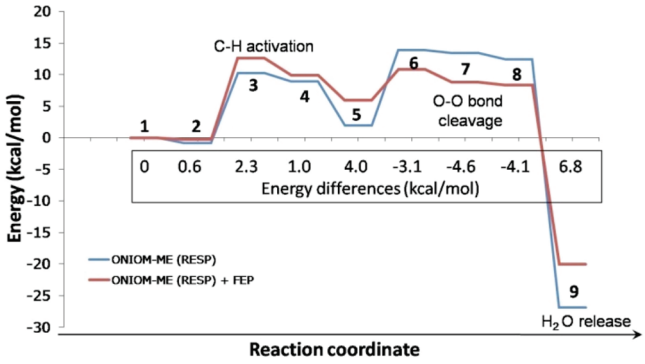

Figure 5. Free-energy diagram for the formation of an Fe(IV)-oxo intermediate after dioxygen binding. The results of the FEP approach are compared to results from a static approach where the protein has been optimized (ONIOM-ME). The relative energy differences between the two approaches are shown in the center box.

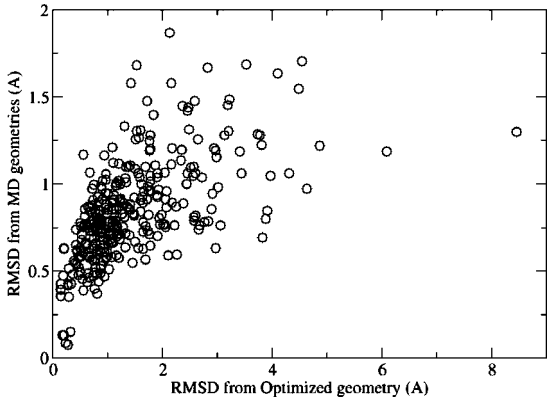

Figure 6. rmsd of the geometry of heavy atoms on each residue. The horizontal axis indicates rmsd between the optimized geometry and the MD averaged geometry. The vertical axis indicates rmsd between the MD averaged geometry and geometries in 10 MD snapshots, from $XY =$ **34** forward simulation. Each point represents a protein residue.

compared to O−O bond cleavage (**5** to **6**), which is consistent with kinetic isotope experiments that show that C−H activation is at least partly rate-limiting.[43] However, the two barriers are relatively close in both cases and the uncertainties in both QM and MM treatment makes it difficult to use the relative barriers as a reliable benchmark. The computed free-energy of C−H activation barrier is 12.9 (±0.3) kcal/mol, where the error bars reflect the statistical error of the FEP calculation. The value is significantly lower than the experimentally estimated reaction barrier of 16.8 kcal/mol,[44] but that barrier is based on a DFT calculation with the B3LYP functional. This method may underestimate barriers of simple hydrogen atom transfer reactions.[45,46]

We discuss the geometrical and statistical effects in the next two subsections.

3.3. Statistical vs Geometrical Effects in the Free Energy Diagram. The geometrical effect is caused by the change in average protein geometry compared to the optimized structure. The statistical effect comes from fluctuations around the average geometry. In Figure 6, each protein residue is represented by a circle. The horizontal axis indicates the rmsd value between the MD average geometry

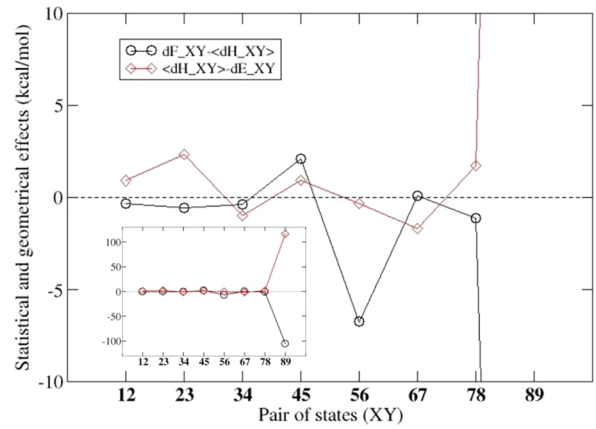

Figure 7. Comparison between the statistical contribution, $\Delta F_{XY}^{\mathrm{FEP}} - \langle \Delta H \rangle_{XY}$ (black circle) and the geometrical contribution $\langle \Delta H \rangle_{XY} - \Delta E_{XY}^{\mathrm{opt}}$ (red diamond) of the protein dynamics on the free energy profile.

and the optimized geometry, that is, how much a residue moves on average in MD simulations. The vertical axis indicates the rmsd in geometry of 10 MD snapshots from the average MD geometry. This value correlates to the flexibility of the residues, that is, how much their positions fluctuate from their average MD positions. Many of these flexible residues are on the protein surface, which is reasonable because the MD system includes explicit water molecules that allow surface residues to move during the simulation. However, we also note that many residues inside the protein also show significant flexibility. We found residues for which both RMSDs are large, and these geometry changes may represent the geometrical and statistical effects. For a more detailed analysis of geometry changes during MD simulations, see the Supporting Information.

Here, we try to determine which effect controls the protein effect on each barrier of the entire reaction free energy profile. We calculated three type of nonbonded interaction energy, $\Delta F_{XY}^{\mathrm{FEP}}$, $\Delta E_{XY}^{\mathrm{opt}}$, and $\langle \Delta H \rangle_{XY}$ that include both geometrical and statistical effects, no effect and the geometrical effect, respectively. When $\Delta F_{XY}^{\mathrm{FEP}}$ and $\Delta E_{XY}^{\mathrm{opt}}$ have similar values, there is no net dynamical contribution to the free energy profile. $\langle \Delta H \rangle_{XY} - \Delta E_{XY}^{\mathrm{opt}}$ and $\Delta F_{XY}^{\mathrm{FEP}} - \langle \Delta H \rangle_{XY}$ are the geometrical and statistical contributions, respectively. The results for $\Delta F_{XY}^{FEP} - \langle \Delta H \rangle_{XY}$ and $\langle \Delta H \rangle_{XY} - \Delta E_{XY}^{\mathrm{opt}}$ for all reaction steps are shown in Figure 7.

We found three cases: (a) the geometrical effect $|\langle \Delta H \rangle_{XY} - \Delta E_{XY}^{\mathrm{opt}}|$ is larger than the statistical effect ($XY =$ **12**, **23**, **34**, **67**, and **78**); (b) the statistical effect $|\Delta F_{XY}^{\mathrm{FEP}} - \langle \Delta H \rangle_{XY}|$ is larger ($XY =$ **45** and **56**); and (c) both values are very large ($XY =$ **89**). In the first case, the geometrical effect dominates the dynamical contribution of the free energy profile. In the free energy diagram (Figure 5), this effect increases the transition state barrier for C−H bond activation from **2** to **3** and affects the energy of steps **34** and**67**. For $XY =$ **12** and **78**, geometrical and statistical effects partly cancel and the total dynamical contribution is small. In second case, the statistical effect dominates the dynamical contribution. This effect decreases the exothermicity of $XY =$ **45** (peroxide rotation) and decreases the reaction barrier of O−O bond activation ($XY =$ **56**). In the third case ($XY =$

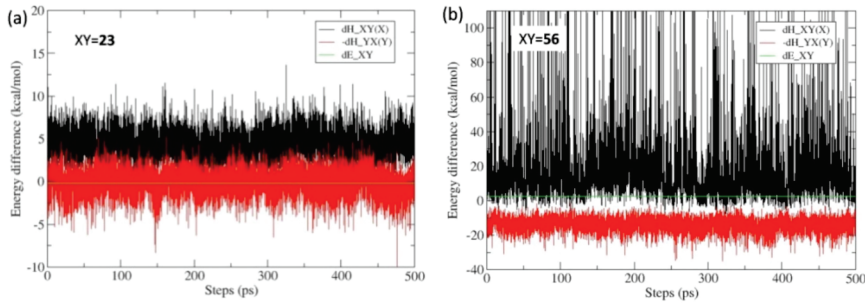

Figure 8. Fluctuation of the interaction energy change $\Delta H_{XY}(X)$ (black line) and the negative reverse $-\Delta H_{YX}(Y)$ (red line) in FEP calculations. (a) $XY = \mathbf{23}$ and (b) $XY = \mathbf{56}$. Green line is the interaction energy difference $\Delta E_{XY}^{\mathrm{opt}}$ between state X and Y in the optimized geometries.

89), both effects are large, but partly cancel, and as the geometrical effect is larger there is a net positive contribution to the total dynamical effect, see Figure 4. However, the case $XY = \mathbf{89}$ may contain an artificial effect of broken criterion of the perturbation calculation (see section 4).

To clarify how these energetic results are coupled to the molecular dynamics simulation, we show the fluctuation of ΔH_{XY} in two typical cases, one where the geometric effect dominates ($XY = \mathbf{23}$) and one where the statistical effect dominates ($XY = \mathbf{56}$), see Figure 8. Black and red lines indicate the forward direction $\Delta H_{XY}(X)$ and the negative reverse direction $-\Delta H_{YX}(Y)$. The green straight line is $\Delta E_{XY}^{\mathrm{opt}}$. In the case of $XY = \mathbf{23}$, Figure 8a, the fluctuations of both black and red lines are small. That is, most geometries have similar opportunity for the reaction from state X to Y or vice versa. All these geometries contribute similarly to the free energy profile and the statistical effect is small. However, the green line, representing the static calculation, is shifted relative to the center of the black and red lines that represent the dynamical calculation. Looking back at the energy contributions in Figure 7, it is clear that the geometrical effect, rather than the statistical effect, dominates the protein contribution to the free energy profile.

In the case of $XY = \mathbf{56}$, Figure 8b, the fluctuation of the black line is large. The possibility that the state X switches to state Y at the geometry R is proportional to $\exp(-\Delta H_{XY}(R)/k_{\mathrm{B}}T)$. Therefore, geometries with lower ΔH_{XY} have larger opportunity to react from state X to Y. When the fluctuation of ΔH_{XY} is large, the reaction is dominated by the few geometries with very low values of ΔH_{XY} rather than the geometric average, and the statistical effect becomes large. Other geometries do not contribute to the reaction from state X to Y.

When the black and red lines cross, these conformational pairs in the same energy can switch states XY and its protein geometries without extra energy. Such condition often appears at the top of the free energy potential barrier. $XY = \mathbf{56}$ has less crossing possibility than $XY = \mathbf{23}$ and intermediate states of X and Y must contribute to the reaction instead of pure states of X and Y. We notice that these two simulations are independent and time axis can shift.

As a measure of the stability of the calculated QM-MM interactions, we use the standard deviation s_{XY} for the interaction energy change ΔH_{XY}, see section 2.3. s_{XY} is defined for the forward reaction from state X to state Y in

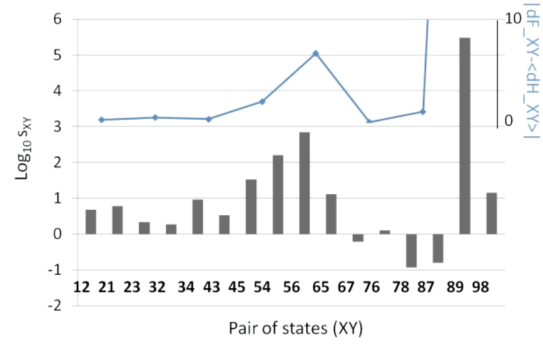

Figure 9. Fluctuation of the interaction change $\Delta H_{XY}(X)$ from state X to state Y for all steps, comparing with pairs of states of stationary points for the IPNS enzymatic reaction, see Figure 5. The vertical axis is the log scale of the standard deviations s_{XY} of the interaction shift $\Delta H_{XY}(X)$. s_{XY} is in kcal/mol. The blue line is the absolute values of the statistical effects to compare with standard deviations.

the state-X protein environment in eq 7. The protein dynamics presents various barrier heights of the reaction steps in time. And when the barrier is low, the reaction can occur easily. During the molecular dynamics simulation, the reaction is more likely to take place through the geometry that has lower energy difference ΔH_{XY} between two stationary points, and the free energy of the reaction ΔF_{XY} becomes lower than in the average geometry. The situation would be the same the reverse reaction with s_{YX} and $\Delta F_{YX} \equiv -\Delta F_{XY}$. When s_{XY} and s_{YX} are similar, the net effect is zero. However, when one of them is larger than the other, fluctuations give a net effect on ΔF_{XY} (if s_{XY} is larger than s_{YX}, ΔF_{XY} decreases). Figure 9 shows s_{XY} and s_{YX} for the eight reaction steps. For example, $s_{\mathbf{54}}$ is much larger than $s_{\mathbf{45}}$for the opposite direction, and $s_{\mathbf{56}}$ is much larger than $s_{\mathbf{65}}$. $\Delta F_{XY}^{\mathrm{FEP}} - \langle \Delta H \rangle_{XY}$ values of $XY = \mathbf{45}$ and **56** are largely positive and negative, respectively, see Figure 7. These dynamical contributions mainly originate from the statistical or entropic effect of the thermal fluctuations. We put the graph of $|\Delta F_{XY}^{\mathrm{FEP}} - \langle \Delta H \rangle_{XY}|$ to compare with the log of standard deviations. Good correlation is shown in Figure 9. For sign of $\Delta F_{XY}^{\mathrm{FEP}} - \langle \Delta H \rangle_{XY}$, see Figure 7.

3.4. Residue Contributions of the Geometrical and Statistical Effects. Figure 10a−d show computed $\Delta\Delta H_{XY}^{i}(X)$ for four pairs with significant dynamical effects, see Figure 4, $XY = \mathbf{23}$, **67** (weak statistical effect) and $XY = \mathbf{45}$ and **56** (strong statistical effect), respectively.

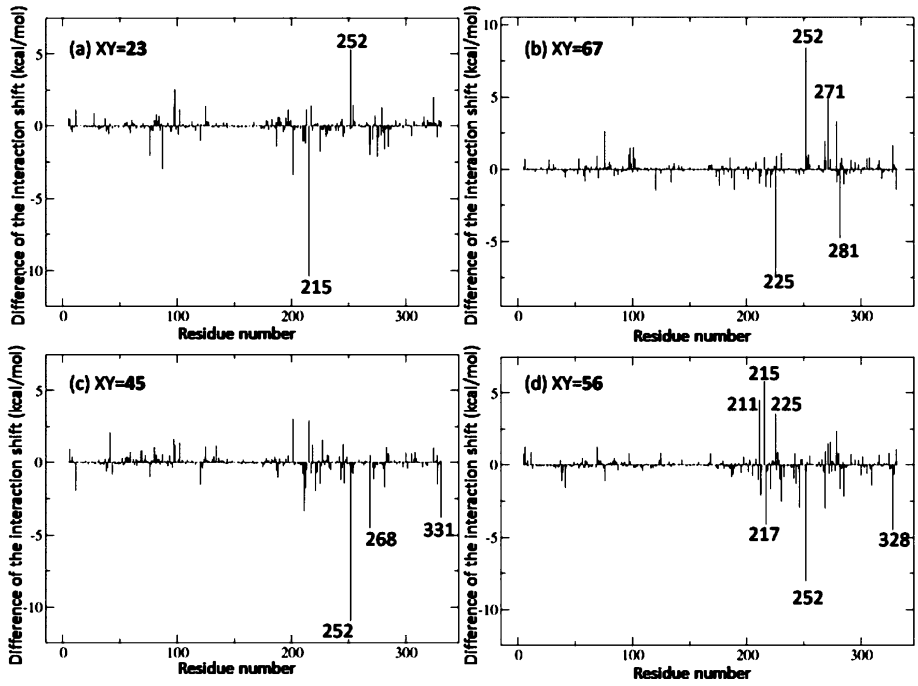

Figure 10. Differences of the interaction energy change $\Delta\Delta H^{i}_{XY}(X)$ between the average MD stationary point geometry and ONIOM QM:MM optimized geometry for individual residue *i*, namely, contributions of residue *i* to the geometrical effect of protein dynamics on the free energy profile. (a, b) Cases of the weak statistical effect ($XY =$ **23** and **67**). (c, d) Cases of the strong statistical effect ($XY =$ **45** and **56**).

In each case, several residues contribute significantly, resulting in partial cancellation of the geometrical effect. A common significant residue, Asn252, has a hydrogen bond with the residue Asp216; Asn252 can rotate almost 180 degrees between the stationary points in MD simulation. This residue is located close to the ligand water molecule that is involved in the reaction. The energetic effect of Asn252 is large in both steps **56** and **67**, but of opposite sign, so it is possible that it switches back and forth as the reaction proceeds. As expected, most other significant residues are also in the vicinity of the QM part. Glu215, Val217, and Arg271 are located next to ligand residues His214, Asp216, and His270, respectively. Asn225 and Ser281 have H-bonds with the ACV substrate. Pro268 has H-bonds with ligand His270. Phe211 is close to the ligand oxygen molecule. Thr331 contacts to MM part of ACV (as shown in the Supporting Information) and Asn328 is a neighbor of Phe211 and His214. We also analyzed whether these geometrical effects are from electrostatic or VDW interactions. In above residues, only Asn252 in $XY =$ **56** and Gln225, Asn252, and Ser281 in $XY =$ **67** have strong geometrical effects because of their electrostatic interaction. Although there are many residues that have similar or even larger size of the electrostatic interaction changes, it is not enough to be strongest without the contribution from VDW. These Asn, Gln, and Ser residues have uncharged polar side chains. Residues with large electrostatic repulsion might not make enough VDW contact with the QM part. The geometrical effects of the other significant residues in Figure 10 are mainly caused by VDW interactions.

Figure 11a−d show S^{i}_{XY} for $XY =$ **23**, **67** (small statistical effects) and **45**, **56** (large statistical effects), respectively. Some residues located near the QM part appear in both the geometrical and statistical effects, but many others only contribute significantly to one of the effects. The statistical effects are mainly caused by VDW interactions, rather than electrostatic interactions because of the stronger distance dependence of the former. Bulky residues, like Phe211 and Pro268, in the vicinity of the QM part have large effects, while residues that contributed to the geometric effect by electrostatic interactions do not appear in the analysis of the statistical effect. Pairs $XY =$ **23** and **67** have relatively small geometric changes of the QM part. In $XY =$ **23**, C−H bond activation, the oxygen molecule moves slightly and the major movement is the hydrogen atom on the substrate ACV that moves closer to the oxygen. Among the significant residues in Figure 11a, Tyr189 and Ser281 both have H-bonds with QM part of ACV, while Phe211 is close to the oxygen molecule. Other significant residues, Phe41 and Pro268 are near the iron ligand His270, and Asn328 is next to the iron ligand His214. In the reaction corresponding to pair $XY =$ **67** (Figure 11b), the O−O distance of the peroxide formed after C−H activation increases in anticipation of O−O bond cleavage. Significant residues Phe221 and Asn252 are both close to the ligand water molecule that donates a proton during O−O bond cleavage, and Pro268 adjoin these two residues.

Both reactions with larger statistical effects (pairs $XY =$ **45** and **56**) also have large geometrical changes in the active site. In $XY =$ **45**, there is a rotation of the peroxide ligand away from the substrate and toward the ligand water that donate its proton. The significant residue Leu231 is located close to the oxygen molecule, while Phe41 is adjacent to the ligand water molecule and Leu231. Tyr189 and Ser281 both have H-bonds with the QM part of ACV, and Ile187 is next to these two residues. Arg271 is next to the ligand

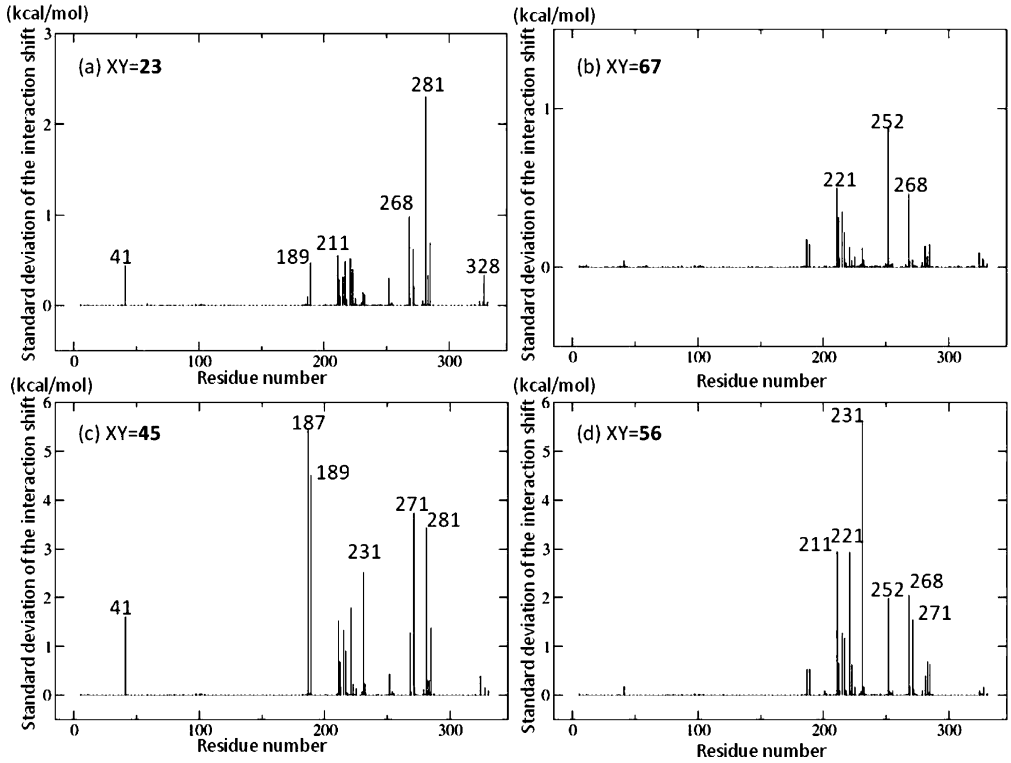

Figure 11. Standard deviation s^i_{XY} of nonbonded interaction change in MD simulation for individual residues, a representation of the statistical effect of the protein dynamics on the free energy profile. (a, b) Cases of weak statistical effect ($XY =$ **23** and **67**). (c, d) Cases of strong statistical effect ($XY =$ **45** and **56**). Note that there are different vertical scales for different figures.

His270 and adjoins the ligand His214 and another significant residue, Leu231. In $XY =$ **56**, an oxygen atom of the ligand OOH moves about 1 Å (another oxygen and hydrogen atoms also move ∼0.5 Å), and the ligand water rotates in the process of the electron transfer from iron to the oxygen molecule. Phe211, Leu231, and Arg271 are close to the ligand oxygen and Thr221 and Asn252 locate next of the ligand water. Pro268 is a neighbor of significant residue Asn252 as written above. Interactions between these residues and QM part significantly change in the reaction from state-**5** to state-**6**, and they contribute to the statistical effect of the free energy profile, see Figure 11d.

Comparing the values of the standard deviation in cases of weak statistical effect ($XY =$ **23** and **67**) within cases of strong statistical effect ($XY =$ **45** and **56**) in Figure 11, the size of the standard deviation for significant residues are much different (see vertical axes). On the other hand, significant values of $\Delta\Delta H^i_{XY}(X)$ are similar size in Figure 10a to 10d. These comparisons suggest that the geometrical effect is relatively stable for different reaction steps and the statistical effect can become large, because these values connect to the statistical and geometrical effects, respectively. When the dynamical contribution is large, the statistical effect dominates it in the reaction step. When the statistical effect is weak, sizes of the geometrical and statistical effects can be comparable.

4. Discussion

The goal in the present paper is to describe a method that can give a broad overview of dynamical contributions for complex multistep reactions. To achieve this, the most important approximations are the neglect of dynamical contributions on the QM region and a classical description of the interaction between QM and MM regions. With these two approximations, we avoid recalculating the QM wave function for each snapshot of the protein geometries.

The QM:[MM-FEP] approach belongs to a family of ONIOM approaches to model interactions between model and real system. ONIOM-ME describes these interactions classically and does not include polarization of the model system or the surrounding. In the semiclassical ONIOM-EE description, the model system is polarized by the charges of the surrounding, but polarization of the environment is not included. The geometry optimization procedure leads only to small changes in the protein structure, and does not really describe geometric polarization. The FEP approach includes a geometric polarization of the actual environment at a finite temperature but uses the classical representation of the electrostatic interactions.

Comparing the static and dynamical results of the potential diagram, the largest change is in the transition state barrier for O−O bond cleavage ($XY =$ **56**), see Figure 5. In this investigation, the dynamical contributions to the energetics have been separated into two types of effects, the geometric and the statistical effects. For the present reaction step, the statistical effect dominates the dynamical contribution of the protein on the free energy profile. The result suggests that the barrier at the QM:MM level includes an artificially high electrostatic repulsion, caused by lack of polarization in the frozen protein environment. The structure fluctuation of the protein environment screens this interaction and decreases the barrier height. We therefore suggest that when redox

center presents significant change of either geometry or electrostatic potential, the dynamical contributions to the free energy reaction diagram should be considered.

In FEP methods, all coordinates other than the selected reaction coordinate are integrated. Any type of the reaction coordinate can be selected, either a real geometry space coordinate or a virtual space coordinate. When choosing a real space coordinate, various computational developments have been reported using different integration methods for leaving coordinates.[47,48] In the most cases, one or a few reaction coordinates like bond, angle or dihedral torsion have been chosen. The approach is appropriate for discussing the realistic dynamics of the system. In the present investigation, we have employed the alchemical FEP method that uses a nonphysical coordinate. The reaction coordinate describes the appearance of atoms in the final state and disappearance of atoms in the initial state. One advantage is that this method can be used for any change of a system, not only chemical reactions. Here, we use the method out of convenience because it is possible to describe the reaction path between two intermediates without a detailed mapping of the potential energy surface. However, in cases where the appearing and disappearing atoms leads to significant changes in the Hamiltonian, the FEP calculation is sometimes hard to converge, and the alchemical FEP method has therefore been preferably applied to small fragments. The alchemical FEP method uses direct linear interpolation of the interaction potential energy. The potential interpolation for the free energy calculation is a traditional approach.[49] There is another approach to build a virtual reaction coordinate in the linear interpolation of the other physical values like atomic coordinates and charges.[50] This alternative approach still avoids the cost of determining the intermediate state and includes less approximation of the interaction energy. At same time, when the rotations of molecules are involved in the reaction, the direction of linear interpolation of the atomic coordinates must be properly chosen.

Our computed value of the dynamical contribution is not a well-determined component of the free energy, because it includes effects of the solvent, temperature difference (compared with 0 K model) and dynamics of the environment. The term is not for comparing with the nature or experiments, but for connecting QM:MM calculations to these. The geometrical effect includes both solvent and temperature effects. On the other hand, the statistical effect is a value connecting to the entropic energy. Computing the entropy in FEP criterion is challenging because the sampling of the correlation function of interactions is required instead of the sampling of interactions in the free energy calculation. That requires N-square order of the free energy sampling. The value of the statistical effect can be considered as an estimation of the entropic energy.

In the present study, we have selected a large fragment that includes the enzymatic reaction center as the alchemical part. The chemical reaction mainly occurs at the center of the fragment and only the metal ligands describe large geometric changes. We find that for reactions where the change in electronic structure is dominant, compared to changes in geometric structure, the alchemical FEP approach is well behaved. However, convergence is slow when strong VDW contacts appear during the FEP step. The error bars are largest for the release of water (step **89)** where the oxygen atom moves out about 1 Å from its original position in state **8** and occasionally makes strong VDW contacts with Leu231 and MM water molecules in the MD simulation. The same oxygen atom moves significantly (0.8−1.3 Å) also in other five reaction steps. But the difference is that in the final step, the water is released from the active site out into the environment. Water molecules are very flexibility in orientation inside the protein and such strong VDW contacts can be avoided. The perturbation criterion is broken at this moment and an artificial energy may be included in this part of FEP calculation. To avoid such problems, it might be necessary to check the change in the effective volume of the redox center when this method is applied.

Another problem appears when the statistical effect has a large contribution to the free energy profile. In that case the FEP calculation picks up only a small number of samples from the configurations that contribute the most. Therefore, such rare events get a very high weight in the FEP calculation. As an example $XY = \mathbf{56}$ has large fluctuation of ΔH_{XY} and the error bar of the computed dynamical contribution is also relatively large. For accurate calculations of reactions dominated by the statistical effect, longer sampling might be needed to include enough number of events, or it may be required to apply another approach.[51]

As we froze the redox center in the MM calculation, we neglect two free-energy terms caused by the dynamics of the redox center. One is that the average geometry of the redox center can change because of the protein dynamics. This effect can be included by applying free energy gradient methods.[52,53] The other term is the statistical effect of the redox center that influences the dynamics of neighboring residues and solvent. QM/MM sampling or an alternative method is required in order to include this term, which would result in a significant increase in computational cost. These terms depend on the computed system and model and are not clearly negligible. That is a challenging problem in the future.

5. Conclusion

The dynamical contribution of the protein environment to the reaction energy diagram has been included at the ONIOM QM:MM level, using an MM FEP method. We applied the method to eight reaction steps of the nonheme iron enzyme isopenicillin N synthase. Redox active metal enzymes require a relatively expensive QM treatment of the active site, and the method is therefore designed to avoid recalculating the QM density at any point of the dynamics simulation.

The dynamical contribution has been separated into geometric and statistical effects of the protein environment. The geometrical effect comes from a change in the average protein geometry and influences to the overall potential of the reaction diagram. The statistical effect is caused by the fluctuation of the interaction between MM and QM part during the molecular dynamics simulation. With respect to the IPNS enzymatic reaction mechanism, the inclusion of the dynamical contribution, mainly coming from the statisti-

cal effect, decreases the barrier for O−O bond cleavage by several kcal/mol. These results show that the dynamical fluctuations of the protein environment can be a factor when modeling enzymatic reactions.

Acknowledgment. One of the authors (M.L.) acknowledges a Fukui Institute for Fundamental Chemistry Fellowship. The 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 use of computational resources at the Research Center of Computer Science (RCCS) at the Institute for Molecular Science (IMS) is acknowledged.

Supporting Information Available: Details of the modified RESP charge fitting procedure. Analysis of geometry changes during MD simulations. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Que, L., Jr.; Tolman, W. B. *Nature* **2008**, *455*, 333–340.

(2) Siegbahn, P. E. M. *J. Biol. Inorg. Chem.* **2006**, *11*, 695–701.

(3) Senn, H. M.; Thiel, W. *Top. Curr. Chem.* **2007**, *268*, 173–290.

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

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

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

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

(8) Dapprich, S.; Komáromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J. *THEOCHEM* **1999**, *461−462*, 1–21.

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

(10) Morokuma, K. *Proc. Jpn. Acad. Sci., Ser. B* **2009**, *85*, 167–182.

(11) Vreven, T.; Frisch, M. J.; Kudin, K. N.; Schlegel, H. B.; Morokuma, K. *Mol. Phys.* **2006**, *104*, 701–704.

(12) Prabhakar, R.; Vreven, T.; Frisch, M. J.; Morokuma, K.; Musaev, D. G. *J. Phys. Chem. B* **2006**, *110*, 13608–13613.

(13) Lundberg, M.; Kawatsu, T.; Vreven, T.; Frisch, M. J.; Morokuma, K. *J. Chem. Theory. Comput.* **2009**, *5*, 220–234.

(14) Kwiecien, R. A.; Khavrutskii, I. V.; Musaev, D. G.; Morokuma, K.; Banerjee, R.; Paneth, P. *J. Am. Chem. Soc.* **2006**, *128*, 1287–1292.

(15) Li, X.; Chung, L. W.; Paneth, P.; Morokuma, K. *J. Am. Chem. Soc.* **2009**, *131*, 5115–5125.

(16) Hu, H.; Yang, W. *Annu. Rev. Phys. Chem.* **2008**, *59*, 573–601.

(17) Yang, W.; Bitetti-Putzer, R.; Karplus, M. *J. Chem. Phys.* **2004**, *120*, 9450–9453.

(18) Rod, T. H.; Ryde, U. *J Chem. Theory Comput.* **2005**, *1*, 1240–1251.

(19) Kästner, J.; Senn, H. M.; Thiel, S.; Otte, N.; Thiel, W. *J Chem. Theory Comput.* **2006**, *2*, 452–461.

(20) Zhang, X.; Bruice, C. T. *J. Am. Chem. Soc.* **2007**, *129*, 1001–1007.

(21) Kamiya, M.; Saito, S.; Ohmine, I. *J. Phys. Chem. B* **2007**, *111*, 2948–2956.

(22) Kollman, P. *Chem. Rev.* **1993**, *93*, 2395–2417.

(23) Gao, J.; Kuczera, K.; Bruce, T.; Karplus, M. *Science* **1989**, *244*, 1069–1072.

(24) Zhang, Y.; Liu, H.; Yang, W. *J. Chem. Phys.* **2000**, *112*, 3483–3492.

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

(26) Rod, T. H.; Ryde, U. *Phys. Rev. Lett.* **2005**, *94*, 138302.

(27) Andersson, I.; Terwisscha van Scheltinga, A. C.; Valegård, K. *Cell. Mol. Life Sci.* **2001**, *58*, 1897–1906.

(28) Lundberg, M.; Siegbahn, P. E. M.; Morokuma, K. *Biochemistry* **2008**, *47*, 1031–1042.

(29) Lundberg, M.; Morokuma, K. *J. Phys. Chem. B* **2007**, *111*, 9380–9389.

(30) Pearlman, D. A.; Kollman, P. A. *J. Chem. Phys.* **1991**, *94*, 4532–4545.

(31) Pearlman, D. A. *J. Phys. Chem.* **1994**, *98*, 1487–1493.

(32) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33–38.

(33) http://www.ks.uiuc.edu/Research/vmd/ (accessed Nov 13, 2010).

(34) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.

(35) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269–10280.

(36) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; 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.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian*, development version;Gaussian, Inc.: Wallingford CT, 2008.

(37) Case, D. A.; Darden, T. A.; Cheatham, III, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Pearlman, D. A.; Crowley, M.; Walker, R. C.; Zhang, W.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Wong, K. F.; Paesani,

F.; Wu, X.; Brozell, S.; Tsui, V.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Mathews, D. H.; Schafmeister, C.; Ross, W. S. Kollman, P. A. *AMBER 9*; University of California: San Francisco, CA, 2006.

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

(39) Namd scalable molecular dynamics. http://www.ks.uiuc.edu/Research/namd/ (accessed Nov 13, 2010).

(40) Efron, B. *Annu. Stat.* **1979**, *7*, 1–26.

(41) Konishi, S. *Ann. Stat.* **1991**, *19*, 2209–2225.

(42) Ooura, T. Ooura's mathematical software packages. http://www.kurims.kyoto-u.ac.jp/~ooura/ (accessed Nov 13, 2010).

(43) Baldwin, J. E.; Abraham, E. *Nat. Prod. Rep.* **1988**, *5*, 129–145.

(44) Kriauciunas, A.; Frolik, C. A.; Hassell, T. C.; Skatrud, P. L.; Johnson, M. G.; Holbrok, N. I.; Chen, V. J. *J. Biol. Chem.* **1991**, *266*, 11779–11788.

(45) Johnson, B. G.; Gonzales, C. A.; Gill, P. M. W.; Pople, J. A. *Chem. Phys. Lett.* **1994**, *221*, 100–108.

(46) Patchkovskii, S.; Ziegler, T. *J. Chem. Phys.* **2002**, *116*, 7806–7813.

(47) Kumar, S.; Bouzida, D.; Robert, H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011–1021.

(48) Isralewitz, B.; Izrailev, S.; Schulten, K. *Biophys. J.* **1997**, *73*, 2972–2979.

(49) Mruzik, M. R. *Chem. Phys. Lett.* **1977**, *48*, 171–175.

(50) Zeng, X.; Hu, H.; Hu, X.; Yang, W. *J. Chem. Phys.* **2009**, *130*, 164111(1−8).

(51) Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. *Chem. Phys. Lett.* **1994**, *222*, 529–539.

(52) Okuyama-Yoshida, N.; Nagaoka, M.; Yamabe, T. *Int. J. Quantum Chem.* **1998**, *70*, 95–103.

(53) Hu, H.; Lu, Z.; Parks, J. M.; Burger, S. K.; Yang, W. *J. Chem. Phys.* **2008**, *128*, 034105(1−18).

CT1005592