Review Article

Biological Applications of Hybrid Quantum Mechanics/Molecular Mechanics Calculation

Jiyoung Kang,[1] Yohsuke Hagiwara,[1] and Masaru Tateno[2]

[1] *Graduate School of Pure and Applied Sciences, University of Tsukuba, Tsukuba Science City, Tsukuba 305-857, Japan*
[2] *Graduate School of Life Science, University of Hyogo, 3-2-1 Koto, Kamigori-cho, Ako-gun, Hyogo 678-1297, Japan*

Correspondence should be addressed to Masaru Tateno, tateno@sci.u-hyogo.ac.jp

Since in most cases biological macromolecular systems including solvent water molecules are remarkably large, the computational costs of performing *ab initio* calculations for the entire structures are prohibitive. Accordingly, QM calculations that are jointed with MM calculations are crucial to evaluate the long-range electrostatic interactions, which significantly affect the electronic structures of biological macromolecules. A UNIX-shell-based interface program connecting the quantum mechanics (QMs) and molecular mechanics (MMs) calculation engines, GAMESS and AMBER, was developed in our lab. The system was applied to a metalloenzyme, azurin, and PU.1-DNA complex; thereby, the significance of the environmental effects on the electronic structures of the site of interest was elucidated. Subsequently, hybrid QM/MM molecular dynamics (MD) simulation using the calculation system was employed for investigation of mechanisms of hydrolysis (editing reaction) in leucyl-tRNA synthetase complexed with the misaminoacylated tRNA^{Leu}, and a novel mechanism of the enzymatic reaction was revealed. Thus, our interface program can play a critical role as a powerful tool for state-of-the-art sophisticated hybrid *ab initio* QM/MM MD simulations of large systems, such as biological macromolecules.

1. Introduction

Ab initio quantum mechanical (QM) calculation is an important tool to investigate functional mechanisms of biological macromolecules based on their three-dimensional and electronic structures. The system size which QM calculations can treat is usually up to a few hundred atoms despite of huge sizes of biomacromolecules including solvent water molecules. Accordingly, isolated models of interesting parts of proteins (e.g., catalytic centers) have been studied in QM calculations. However, the ignored regions in the QM calculations, that is, proteins and solvent, surrounding the catalytic centers, have been shown to contribute to the regulation of electronic structures and geometries of the regions of interest [1–12]. This implies that such QM calculations using the isolated models of the catalytic centers have difficulty in the understanding of biological systems.

To overcome these difficulties, quantum mechanics/molecular mechanics (QM/MM) calculations are utilized, in which the system is divided into QM and MM regions. Here,

QM regions correspond to active sites to be investigated, and they are described quantum mechanically. MM regions correspond to the remainder of the system and are described molecular mechanically. The pioneer work of the QM/MM method was accomplished by Warshel and Levitt [12, 13], and since then, there has been much progress on development of QM/MM algorithm and applications to biological systems [14–25]. Here, we describe the procedure of a novel interface program connecting general software packages to perform QM/MM calculations, and its applications for investigating environmental effects of electronic structures of active centers and mechanisms of enzymatic reactions. This could be a solid basis for further applied scientific fields such as biological and medical molecular design.

2. Energy Expression of QM/MM Calculation

In QM/MM method, an entire system is divided into two into QM region, which is described by quantum mechanics, and MM region described by molecular mechanics. Due to

the presence of the interactions between the QM and MM regions, the total energy of the entire system can be formally written as follows:

$$E(\text{ES}) = E(\text{QM}) + E(\text{MM}) + E(\text{QM}, \text{MM}). \qquad (1)$$

The energy term of QM-MM interactions, $E(\text{QM}, \text{MM})$, enables a more realistic description of the system, compared with isolated QM calculations. With respect to the treatment of the electrostatic interaction between the QM and MM regions, QM/MM methods are divided into two groups, subtractive and additive schemes. Subtractive schemes consist of the three steps as follows; (1) an MM calculation on the entire system, (2) a QM calculation on the QM region, and (3) an MM calculation on the QM region. Then, QM/MM energy of the entire system can be formulated as follows:

$$E^{\text{sub}}(\text{ES}) = E_{\text{MM}}(\text{ES}) + E_{\text{QM}}(\text{QM}) - E_{\text{MM}}(\text{QM}). \qquad (2)$$

The subscript indicates the type of calculation (QM or MM calculation), and the region on which the calculation is performed is described in parentheses. An advantage of the subtractive schemes is simplicity. Explicit descriptions of interactions between QM and MM regions are not required. In addition, artifacts which might be caused by using link atom schemes to cap the truncated bonds at the QM-MM boundary (described below) can be avoided. On the other hand, disadvantages of the subtractive schemes are the following: (1) force fields are required for describing the QM region which often includes ligands and intermediate structures of enzymatic reactions; in general, reliable force fields of the molecules are not prepared, and additional QM calculations should be carried out for the parameterization every time a new system is studied; (2) the electrostatic interactions between the QM and MM regions are described at molecular mechanics level; that is, the interactions are calculated by the Coulomb interactions between fixed atomic charges in the QM and MM regions. Such description cannot represent polarization of the QM region induced by environment surrounding the QM region.

In contrast, the additive schemes can take into account the polarization effects. The energy for the additive schemes is given in (3)

$$E^{\text{add}}(\text{ES}) = E_{\text{MM}}(\text{MM}) + E_{\text{QM}}(\text{QM}) + E_{\text{QM/MM}}(\text{QM}, \text{MM}). \qquad (3)$$

A characteristic feature of this scheme is the presence of the energy term with respect to the interactions between QM and MM regions, described as follows:

$$\begin{aligned} E_{\text{QM/MM}}(\text{QM}, \text{MM}) = &\ E_{\text{QM/MM}}^{\text{elec}}(\text{QM}, \text{MM}) \\ &+ E_{\text{QM/MM}}^{\text{bonded}}(\text{QM}, \text{MM}) \\ &+ E_{\text{QM/MM}}^{\text{vdW}}(\text{QM}, \text{MM}). \end{aligned} \qquad (4)$$

To calculate electrostatic interactions, that is, the first term in the left side of (4), one-electron integrals in the QM Hamiltonian incorporating MM partial charges can be used

$$\begin{aligned} \hat{H}_{\text{QM/MM}}^{\text{elec}}(\text{QM}, \text{MM}) = &-\sum_i^N \sum_j^M \frac{q_j}{\left| \mathbf{r}_i - \mathbf{R}_j \right|} \\ &+ \sum_k^L \sum_j^M \frac{q_j Q_k}{\left| \mathbf{R}_k - \mathbf{R}_j \right|}. \end{aligned} \qquad (5)$$

The symbols q_j are the MM partial charges located at $\mathbf{R}_j$. Q_k are the nuclear charges of the QM atoms at $\mathbf{R}_k$, and $\mathbf{r}_i$ represents electron positions. N, M, and L represent number of electrons, MM atoms to be incorporated into the one-electron integrals, and QM atoms, respectively. Using the additive scheme, the electronic structures of the QM region are affected by the charge distribution of the environment. An advanced approach is to consider polarization of the MM region by the QM region (i.e., to allow the partial charges to be changed according to changes in the electronic structure of the QM region). However, polarizable force fields with broader applications have not yet emerged, despite many efforts to account for the polarization effects [26, 27]. A solution of this problem was recently proposed, and the scheme was applied to molecular dynamics (MD) calculation involving Na^+-π interaction, since the polarization effect is important to describe this type of interaction [28].

Note that a problem arises in the case that a covalent bond crosses the boundary between QM and MM regions (e.g., the side chain of an amino acid is assigned to the QM region, while its backbone is assigned to the MM region). Obviously, electronic valence state of the QM subsystem should be closed in the QM Hamiltonian. A traditional approach to reconstruct a saturated valence structure is link atom method [29]. Each bond between such boundary atoms in the QM and MM regions is capped by a hydrogen atom; this link atom is placed along the bond vector between the QM and MM atoms at a distance of ~ 1 Å. This new particle is treated as a QM atom. Although its simplicity is an advantage of this method, introducing the additional atom which does not exist in the real system causes several problems. A serious one is overpolarization problem, where the electronic structure of the QM region is unphysically polarized by the "strong" partial charge of the boundary MM atom in the case of using the additive scheme, since the link atom is spatially very close to the boundary atoms [21, 29]. In the cases of calculations where small, localized basis sets are employed, the unphysical polarization effects are less serious. However, the effects are more severe in the cases where flexible basis sets are used, such as larger basis sets including polarization and diffuse functions and, in particular, plane waves [21].

Several schemes have been devised to avoid the overpolarization issue: (1) not computing the one-electron integrals regarding the link atoms and MM partial charges [13, 30], (2) assigning zero to the partial charges of the MM atoms replaced with the link atoms [31], and (3) introducing distributed charges by using Gaussian functions centered at the positions of the MM atoms close to the QM region

[32, 33]. Generalized hybrid orbitals (GHOs) method is a major method to overcome the disadvantage of the link atom method, which originats from local self-consistent field (LSCF) method [34–40]. In this method, "auxiliary" and "active" orbitals are located on the MM boundary atoms, respectively. By incorporating the active orbital, which is directed towards the QM boundary atoms into SCF calculation, the valency of the QM subsystem is satisfied, while the auxiliary orbitals are fixed during the SCF calculation.

3. Computer Programs for QM/MM Calculations

To conduct QM/MM calculation, several programs are available. For example, for MM calculations, AMBER [41, 42] and CHARMM [43], both of which are widely used program packages, include components for QM calculation in their main packages. A widely used software package for QM calculation, Gaussian [44], has the MM components for QM/MM calculation. An effective approach to carry out QM/MM calculation is to connect existing QM and MM software packages by an interface program. ChemShell [45], QoMMM [46], and PUPIL [47] belong to this category. An advantage of this approach is to enable us to utilize characteristic program modules that each package has (e.g., efficient algorithms for expanding conformational spaces sampled in molecular dynamics simulations implemented in MM packages and algorithms to find optimal reaction paths and stationary points on potential energy surface implemented in QM packages).

We recently developed an interface program to couple AMBER [48] and GAMESS [49] for QM/MM hybrid calculations [11]. Thereby, we can employ various *ab initio* calculations as the QM part in QM/MM calculations, while just semiempirical methodologies are available in AMBER. Figure 1 shows a schematic architecture of our calculation system. Roles of the interface program are (i) to control the sequence of steps for the QM/MM optimization and MD simulation and (ii) to exchange information between those two program packages. At the first stage of a calculation cycle, a single-point calculation is performed in GAMESS regarding a coordinate set of a system. This yields two types of forces, that is, (1) the forces on QM atoms from the QM-QM and QM-MM interactions and (2) the forces on the MM atom derived from the QM-MM interactions. To calculate the second force, the contribution from the QM atoms to the MM atoms is calculated by integrating the interaction between the partial charges of MM atoms and the electron density at each grid point of the quantum region, as shown in (6)

$$F_j^{\mathrm{MM}} = \sum_i^N \mathbf{r}_{ij} \frac{q_j}{\left| r_{ij} \right|^3} dq_i. \tag{6}$$

In the equation, q_m is a partial charge in a MM atom and N is the total number of electron density points. dq_j designates a small volume on the electron density grid, $dq_j = \rho_j dx\, dy\, dz$. This manner of the calculation of the forces on

the MM atoms is more accurate than the manner that such forces are calculated by approximating QM charges using Mulliken and RESP charges. With respect to the coordinate set, AMBER is responsible for calculating (1) forces from van der Waals (vdW) interactions between all atoms of the system and (2) forces on MM atoms from MM-MM interactions. After those calculations, the interface program combines the forces calculated by GAMESS and AMBER. At this step, mapping of the relationship between QM and MM atoms is required, since the manner of atom numbering in one program package is different from that of the other one. In order to reduce the overhead cost regarding this process (in particular, the computational time required would be increased, when treating a solvated biomacromolecules over 100,000 atoms), we employ the combination of UNIX shell and C programming. Finally, using the combined forces, AMBER updates the coordinate set in the case of geometry optimization and MD simulation. The issues on the treatments of electrostatic interactions between QM and MM regions are important in the calculation efficiency and have also been discussed previously [50, 51].

4. Investigation of Environmental Effects on Electronic Structure of QM Region

In this section, we discuss the effects of the environmental structures on the electronic state of the active centers of some biological macromolecules. First, we take a metalloenzyme as an example of such an analysis. Metalloenzymes are involved in various biological functions such as electron transfer, binding of dioxygen, storage of metals, turnover of substrates, and configuration of protein structure [52, 53] and are studied in various research fields, such as biochemistry, inorganic chemistry, biophysics, bioengineering, and medical sciences. Understanding the coordination geometry and electronic structure of metal active sites in the metalloenzymes is of fundamental biophysical importance to gain insight into the structure/function correlation of these biological macromolecules. Azurin is one of the metalloenzymes, whose biological function is electron transfer. The copper ion bound to its active site is coordinated by five amino acid residues, a cystein (Cys), two histidines (His), methionine (Met), and a backbone carbonyl oxygen of glycine (Gly) (Figure 2) [54].

We performed QM/MM calculations of the active site of azurin using both of the subtractive and additive schemes to assess the polarization effects of the environment on the geometry and electronic properties of the active center and compared those two results [55, 56]. The spin-unrestricted Hartree-Fock (UHF)/density functional theory (DFT) hybrid all-electron calculation with the B3LYP functional was adopted as the QM Hamiltonian. The detailed theoretical methodologies are previously described [11]. The comparison of the optimized structures obtained by the two schemes showed that the additive scheme gave the electronic property more consistent with spectroscopic experiment than the subtractive scheme. In the experiments, the contribution of $3p$-orbital character of sulphur atom (S) of Cys112 was 45% in the SOMO (Figure 3) [57]. In our QM/MM calculation,
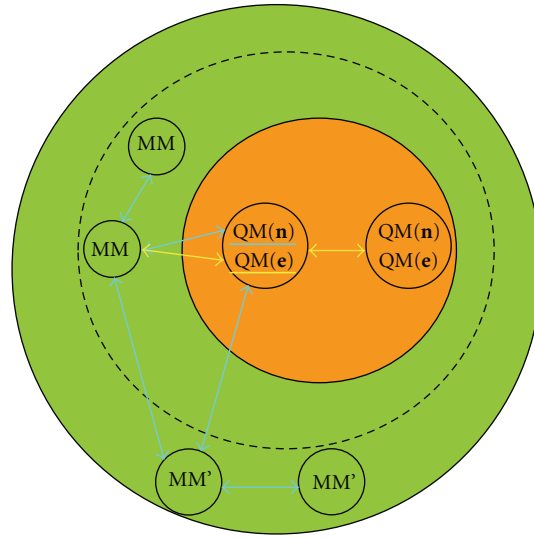
QM engine

$$\hat{H}_{\mathrm{QM}} = -\frac{1}{2}\nabla_i^2 + \sum_{i<j}\frac{1}{r_{ij}} - \sum_{i\alpha}\frac{Z_\alpha}{r_{i\alpha}} + \sum_{\alpha<\beta}\frac{Z_\alpha Z_\beta}{R_{\alpha\beta}} - \sum_{iM}\frac{q_M}{r_{iM}}$$

$$\left\{ -\frac{\hbar}{2m}\nabla^2 + V[\rho] \right\}\psi_n(\mathbf{r}) = \varepsilon n_{\mathrm{QM}}\psi_n(\mathbf{r}),\, \rho(\mathbf{r}) = \sum_n f_n |\psi_n(\mathbf{r})|^2$$

$$F_{\mathrm{QM}}(\mathbf{r}) = -\nabla E_{\mathrm{QM}}$$

MM engine

$$E_{\mathrm{MM}} = \sum_{\mathrm{bonds}} K_r(r - r_{\mathrm{req}})^2 + \sum_{\mathrm{angels}} K_\theta(\theta - \theta_{\mathrm{req}})^2 + \sum_{\mathrm{dihedrals}} \frac{V_n}{2}(1 + \cos[n\phi - \gamma])$$

$$+ \sum_{i<j}^{\mathrm{atoms}} \frac{q_i q_j}{\varepsilon R_{ij}} + \sum_{i<j}^{\mathrm{atoms}} \left(\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} \right) + \sum_{\alpha M} \frac{Z_\alpha q_M}{R_{\alpha M}}$$

$$F_{\mathrm{MM}}(\mathbf{r}) = -\nabla E_{\mathrm{MM}}$$

Merging of the forces

$$F(\mathbf{r}) = F_{\mathrm{QM}} + F_{\mathrm{MM}}$$

Equation of motion

$$M\frac{d^2\mathbf{r}}{dt^2} = F(\mathbf{r})$$

FIGURE 1: The QM/MM scheme used in the present calculation system. QM(n) and QM(e) represent nuclei and electrons in the QM regions, respectively.
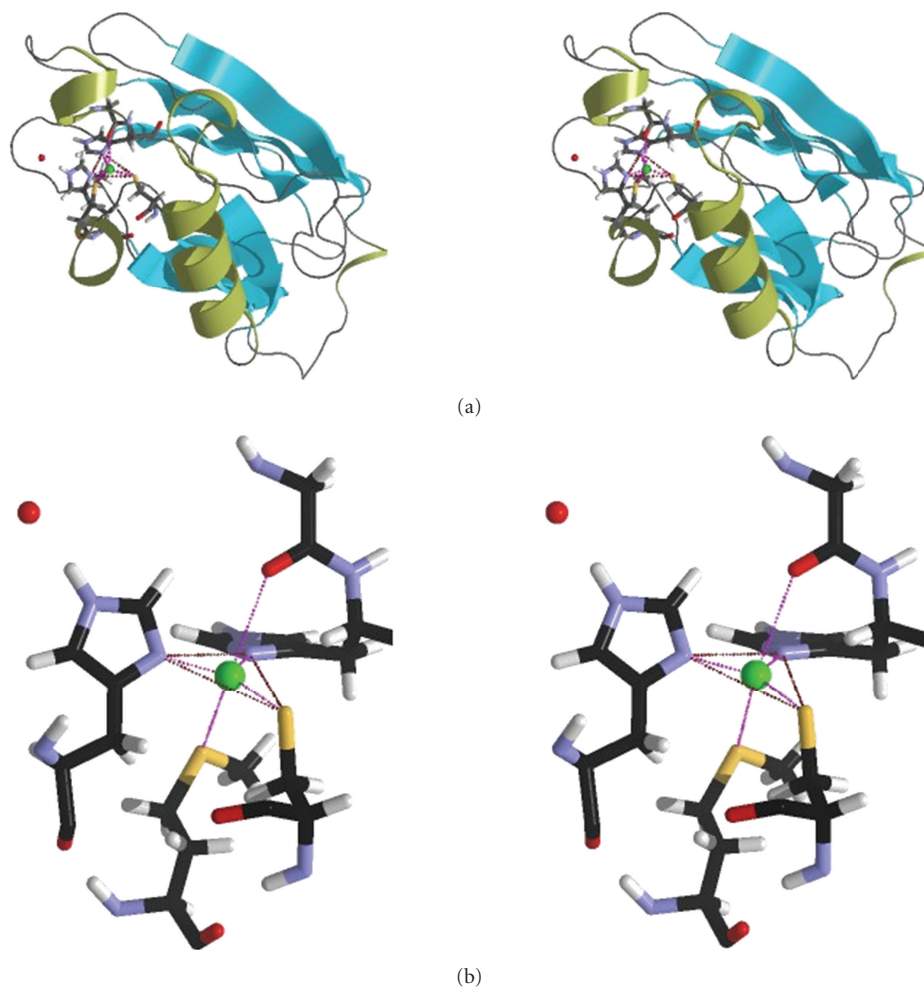
(a)

(b)

FIGURE 2: Stereoview of the three-dimensional structure of azurin (a) and its active site (b).

the spin polarization was significantly improved in model I (50%) compared to model II (57%). This indicates that the explicit inclusion of electrostatic interaction by the QM Hamiltonian is essential for the calculation of the copper site.

Further, we found that some occupied MOs are affected by the environment. For example, with respect to the MOs dominantly occupied by $2p$ orbital of O atom of Gly45 (the backbone carbonyl), their energy levels in models I and II are different, as shown in Figure 4 [56]. In terms of the geometries, different descriptions of the copper coordination geometry were obtained, particularly for the coordinative bonds including a large dipole: Cu–O (Gly45) and Cu–S (Cys112). Thus, it is suggested that modification of the QM Hamiltonian to interact with the long-distance partial point charges of the environment is crucial for an accurate QM/MM description of the geometrical and electronic properties of metal active sites.

Next, to quantify the effects of the surrounding MM atoms on the electronic structures of the QM region, we take a solvated protein-DNA complex as the second example, that is, PU.1 in a complex with the target double-stranded DNA molecules (Figure 5). PU.1 is a member of the Ets family of

proteins, which consists of transcriptional factors that regulate gene expression for biological growth and development. They share a conserved domain composed of around 85 amino acid residues, which binds with the consensus DNA sequence, 5′-GGAA-3′, as the core sequence in the DNA. As the calculation of azurin mentioned above, we employed the two schemes, where the QM region was assigned to the consensus DNA sequence and the amino acid residues which recognize the sequence. The comparison of the molecular orbitals (MOs) obtained in the additive scheme with those obtained in the subtractive scheme revealed that the order of some MOs is shifted and that the surrounding MM atoms have significant impacts on the energy of MOs [11]. This result also emphasizes the crucial importance of considering the long-range electrostatic effects from MM atoms and suggests that biological functions involved in active sites are modulated through long-range electrostatic interactions with the surrounding regions. Actually, it was reported experimentally that Bam HI, which binds the specific DNA core sequence, is supposed to protect the native electronic states of guanine bases from damages (caused by radicals) by regulating the oxidation potential of the DNA [58]. This
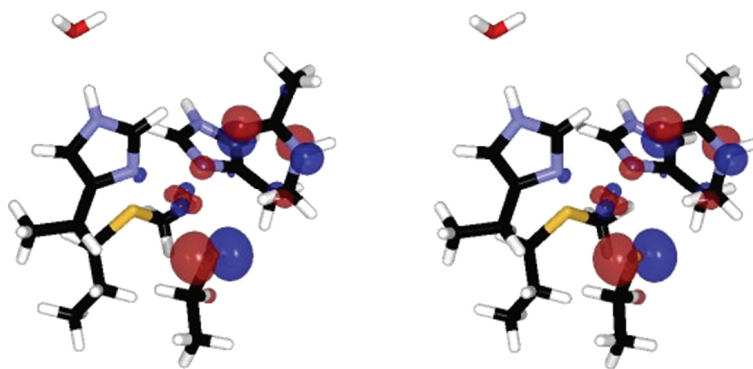
Figure 3: Singly occupied molecular orbital (SOMO) of the catalytic site of azurin in the oxidized state (stereoview).
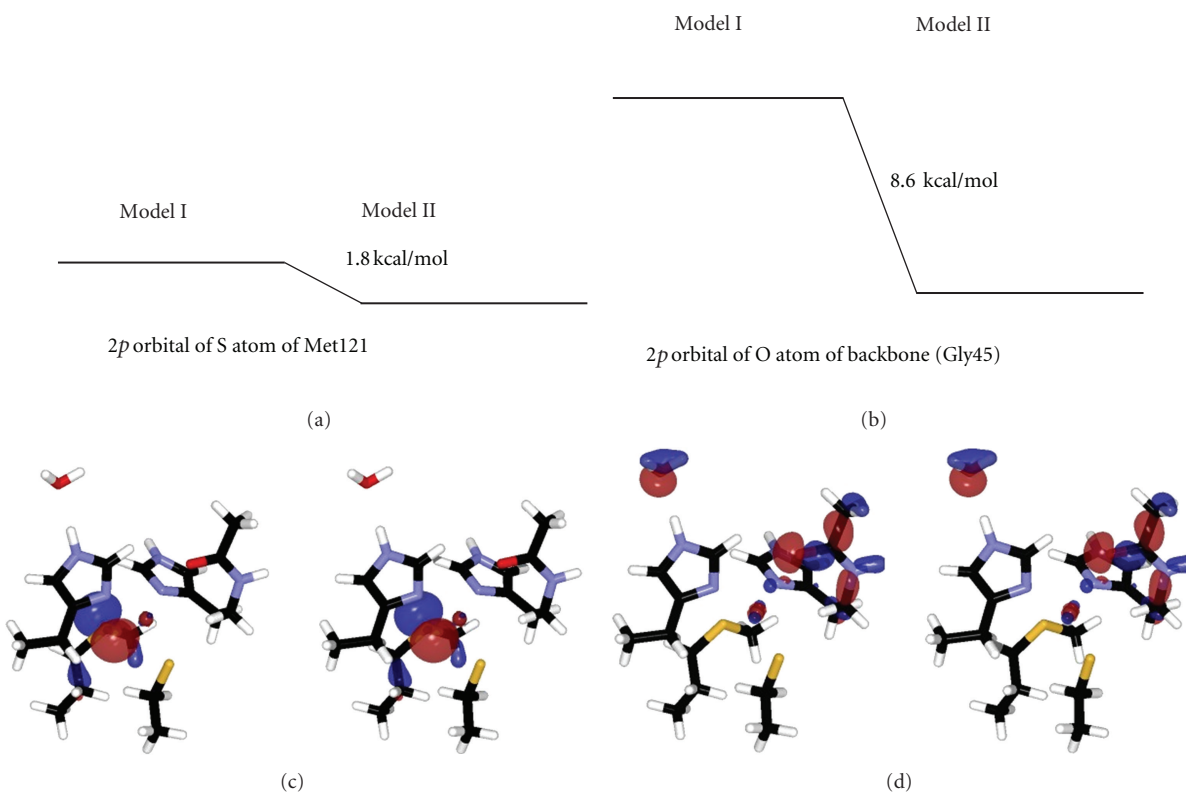
Model I   Model II

8.6 kcal/mol

Model I   Model II

1.8 kcal/mol

2p orbital of S atom of Met121

2p orbital of O atom of backbone (Gly45)

(a)   (b)

(c)   (d)

Figure 4: (a) The energy difference of 2p orbitals of S atom of Met121 between models I and II is small. (b) In contrast, the energy difference of 2p orbitals of O atom (the backbone) of Gly45 between models I and II is larger. (c) The MO which is mostly occupied by 2p orbital of S atom of Met121 is shown (stereoview). (d) The MO which is highly occupied by the peptide group of Gly45 is shown (stereoview).

suggests that protein can modulate the electronic states of DNA molecules, thereby playing a crucial role in the biological functions.

Furthermore, with respect to the PU.1-DNA complex, we investigated relationship between energies of MOs of DNA bases and the DNA recognition by PU.1 using the hybrid QM/MM calculations [59]. In the complex, PU.1 masks some DNA bases from the solvent through hydrogen bonds and van der Waals contacts. Accordingly, the magnitude of the masking by the protein is, in general, corresponding to the recognition mode of DNA and the protein. The analysis using the hybrid QM/MM calculations revealed that the

MO energies of some DNA bases strongly correlate with the magnitude of masking of the bases from the solvent by the protein. Moreover, in the complex, PU.1 causes a variation in the magnitude of the masking among DNA bases through directly recognizing the bases and inducing structural changes of the DNA structure from the canonical one. In this way, the strong correlation of the MO energies and the magnitude of masking DNA bases from the solvent by protein is the first evidence showing the close quantitative relationship between recognition modes of bases and the energy levels of the corresponding MOs. This also means that the electronic state of each base is highly regulated by the DNA recognition of
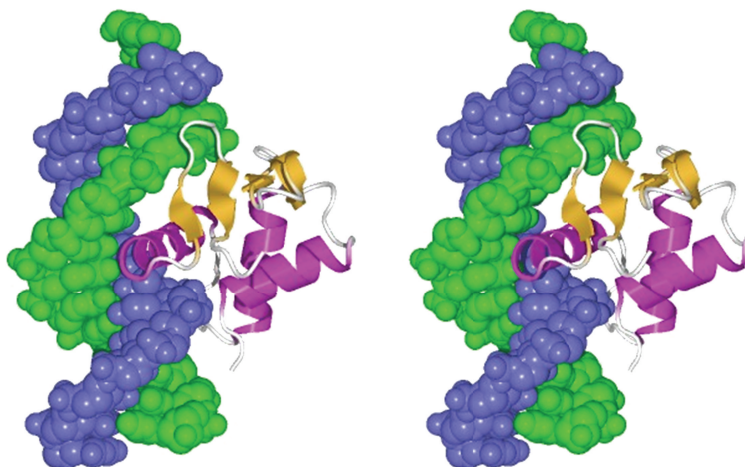
FIGURE 5: Stereoview of the three-dimensional structure of the complex of PU.1 and DNA. DNA recognized by PU.1 is shown by the green and light blue CPK model. PU.1 is represented by ribbon model.

the protein. Other biological macromolecular systems can be expected to also possess modulation mechanisms, suggesting that this provides a novel basis for the understanding of the regulation functions of biological macromolecular systems.

5. Investigations of Enzymatic Reactions: A Novel Catalytic Mechanism by a Hybrid Ribozyme/Protein Catalyst

The following study is an example of applications of hybrid QM/MM MD simulation to enzymatic reactions. Aminoacyl-tRNA synthetases play a critical role in decoding genetic information through catalyzing attachment of their cognate amino acid to $3'$-end of the specific tRNA. The fidelity of translation is assured by their strict discrimination of the cognate amino acids from noncognate ones. However, several systems have difficulties in the strict discrimination of their specific amino acids, producing misaminoacylated tRNAs [60, 61]. To avoid this error, the enzymes have editing function to hydrolyze the mis-products by themselves, but detailed mechanisms are not known due to the absence of experimental structures of aaRSs in complex with tRNA bound to the active site of the editing reaction. The crystal structure of *Thermus thermophilus* leucyl-tRNA synthetase (LeuRS) in complex with tRNA^{Leu} bound to the connective polypeptide (CP) 1 domain on which the active site of the editing reaction is located was determined [62]. LeuRS is one of aaRSs possessing the editing function. Exploiting this crystal structure, we performed structural modeling of the LeuRS•tRNA^{Leu} complex, for which an incognate amino acid, valine, is attached to A76 of tRNA^{Leu} [63], and identified the water molecule corresponding to the nucleophile of the hydrolysis (Figure 6) [64].

Furthermore, we elucidated the mechanism for the nucleophilic water to approach the electrophilic atom, that is, the carbonyl carbon of the misattached valine: the hydrogen atom attached to the O$^{3'}$ atom of the ribose of A76 ($3'$-HO) is required to be translocated for the nucleophilic water

to approach the carbon. In other words, the $3'$-HO plays roles as a "gate" to regulate the access of the nucleophile (Figure 6(b)) [64]. Then, in order to investigate changes in electronic structures on the gate opening, we performed hybrid QM/MM calculations using the above-mentioned interface program.

As a result, in both of opened and closed states, the lowest unoccupied molecular orbital (LUMO) is localized on the carbonyl group of the substrate, which contains an anti-bonding character regarding the C–O$^{2'}$ bond. This indicates that the carbonyl carbon atom can accept electrons from the nucleophile and that the electron donation results in the cleavage of the bond. Furthermore, with respect to the nucleophile, the energy difference between LUMO and the MO which includes a major contribution from the valence *p*-orbitals of the oxygen atom of the nucleophile decreases on gate opening. This indicates that the nucleophilic attack is facilitated on gate opening. Thus, the initial phase of the editing reaction in the LeuRS•tRNA^{Leu} complex was identified in terms of the changes of the electronic structures.

Moreover, to elucidate the editing mechanisms, we performed hybrid QM/MM MD simulations [65]. We examined four possible schemes of the reaction and identified the most feasible pathway, in which the nucleophilic water is activated by the $3'$-OH of tRNA^{Leu} and the O$^{2'}$ atom is capped by a hydrogen atom of the water. This means that the editing is actually a self-cleavage reaction of the tRNA; that is, editing is a ribozymal reaction. However, it should be noted here that the protein also contributes to the reaction, by (i) stabilizing the transition state and (ii) increasing the catalytic power of the $3'$-OH to activate the water through the formation of the hydrogen bond networks. Furthermore, the importance of opening the H-gate was also confirmed in the hybrid QM/MM MD simulations. Considering the contributions of the protein moiety to this ribozymal reaction, we concluded that editing is achieved by a "hybrid ribozyme/protein catalyst," which represents a novel enzyme category.

Recently, a similar theoretical analysis of the editing mechanism was conducted by employing our fully-solvated
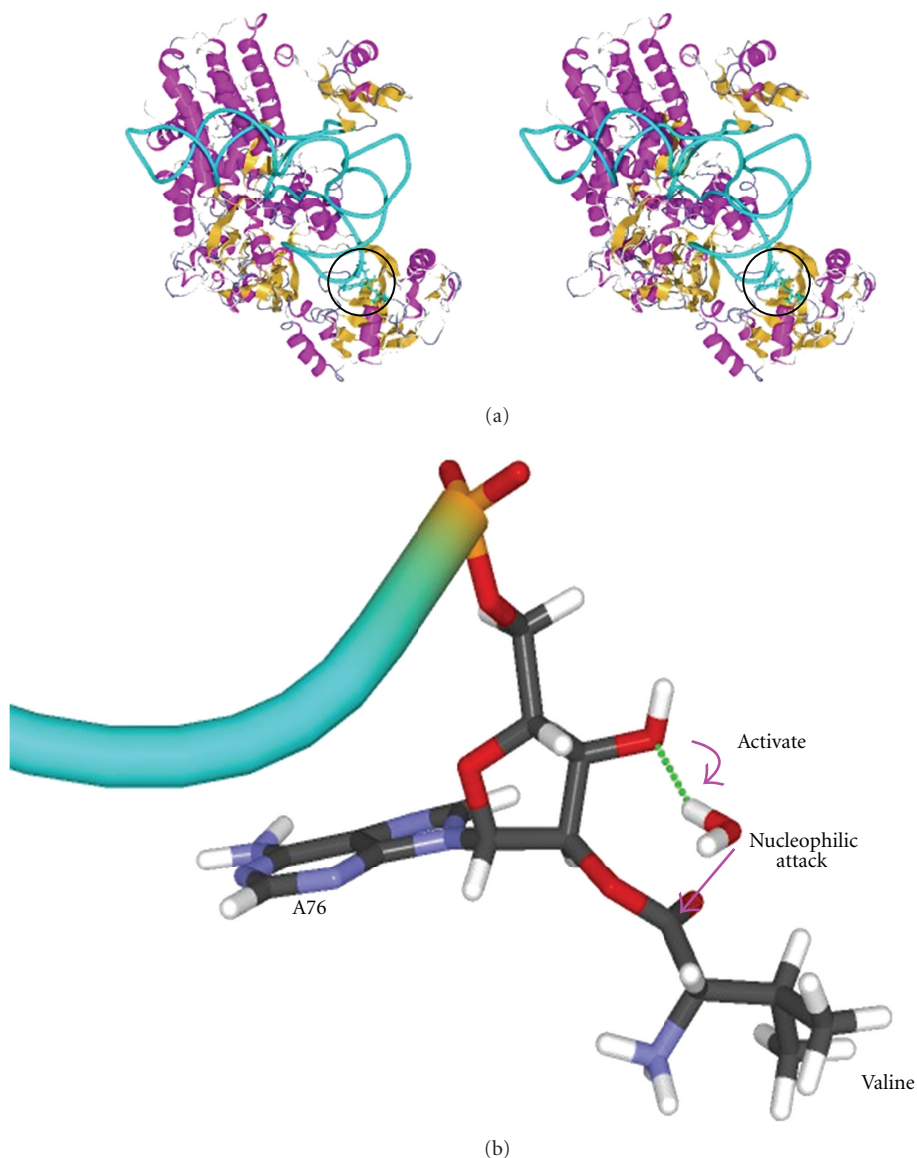
(a)

Activate

Nucleophilic
attack

A76

Valine

(b)

FIGURE 6: (a) The three-dimensional structure of the LeuRS•valyl-tRNA[Leu] complex (stereoview). The active site of editing is shown by a black circle. (b) Close-up view of the editing site in the LeuRS•valyl-tRNA[Leu] complex. The valine-attached A76 and the nucleophilic water molecule are shown. The backbone of tRNA[Leu] is shown using the tube representation colored in cyan. Our calculations showed that $OH^{3'}$ activates the nucleophilic water molecule, which is the trigger of editing reaction as well as the opening of the "H-gate".

structural model [63, 64], and the author reported almost the same reaction mechanisms as described above [66]. The small difference seems to be another alternative initial reaction pathway that was observed in the more recent study. However, this difference found in the more recent study would be due to the inappropriate QM regions lacking the crucial atoms, and so on (also see Simulations 2 and 4 and the discussions in Hagiwara et al., [65]).

The above-mentioned result is consistent with the widely accepted, but unproven, scenario of the evolution of life, which proposes that proteins previously just acted as chaperones and subsequently replaced such functional ribozymes [67]. Further, we could propose a broader scenario by introducing a novel stage of evolution, where tRNA itself

acquired the ability to play the catalytic roles and employed the proteins to enhance the catalysis; this can be called the "hybrid catalyst world," which previously existed on the earth and played a critical role in generating our modern world of life.

6. Conclusions

To connect the QM and MM calculation engines, GAMESS and AMBER, a UNIX shell-based interface program, was developed in our group. Overhead costs of the UNIX shell which arise from treating huge amount of information such as number of atoms of a solvated system of the biological system were avoided using C programs and

modifications of the AMBER and GAMESS source code. The GAMESS and AMBER engines used in our system are highly efficient for parallelization; therefore, our system is suitable for simulations of large molecular systems. This system was applied to a metalloenzyme, azurin, and PU.1-DNA complex, and the significance of the environmental effects on the electronic structures of the site of interest was revealed. Subsequently, the system was employed for investigation of mechanisms of editing reaction in LeuRS, and the productive enzyme·substrate complex was identified as an initial state of the enzymatic reaction. Then, the reaction pathways of the editing reaction occurring in the complex of LeuRS and mis-aminoacylated tRNA$^{\text{Leu}}$ were investigated, and thereby the catalytic mechanism was identified. The revealed mechanism is consistent with the biochemical experimental result. This mechanism led us to finding a novel type of enzyme, that is, a hybrid ribozyme/protein catalyst. Thus, our interface program can play a critical role as a powerful tool for sophisticated hybrid QM/MM MD simulations of large systems, such as biological macromolecules. Such a state-of-the-art theoretical methodology could also be a solid basis for applied scientific fields such as pharmaceutical molecular design.

Acknowledgments

References

[1] E. J. Merino, A. K. Boal, and J. K. Barton, "Biological contexts for DNA charge transport chemistry," *Current Opinion in Chemical Biology*, vol. 12, no. 2, pp. 229–237, 2008.

[2] K. Nakatani, C. Dohno, A. Ogawa, and I. Saito, "Suppression of DNA-mediated charge transport by BamHI binding," *Chemistry and Biology*, vol. 9, no. 3, pp. 361–366, 2002.

[3] T. Kubař, P. Benjamin Woiczikowski, G. Cuniberti, and M. Elstner, "Efficient calculation of charge-transfer matrix elements for hole transfer in DNA," *Journal of Physical Chemistry B*, vol. 112, no. 26, pp. 7937–7947, 2008.

[4] I. Solt, I. Simon, A. G. Császár, and M. Fuxreiter, "Electrostatic versus nonelectrostatic effects in DNA sequence discrimination by divalent ions Mg2+ and Mn2+," *Journal of Physical Chemistry B*, vol. 111, no. 22, pp. 6272–6279, 2007.

[5] J. B. MacNaughton, A. Moewes, J. S. Lee et al., "Dependence of DNA electronic structure on environmental and structural variations," *Journal of Physical Chemistry B*, vol. 110, no. 32, pp. 15742–15748, 2006.

[6] J. Ladik, A. Bende, and F. Bogár, "The electronic structure of the four nucleotide bases in DNA, of their stacks, and of their homopolynucleotides in the absence and presence of water," *Journal of Chemical Physics*, vol. 128, no. 10, 2008.

[7] P. Jayapal, M. Sundararajan, I. H. Hillier, and N. A. Burton, "QM/MM studies of Ni-Fe hydrogenases: the effect of enzyme environment on the structure and energies of the inactive and active states," *Physical Chemistry Chemical Physics*, vol. 10, no. 29, pp. 4249–4257, 2008.

[8] M. Mladenovic, R. F. Fink, W. Thiel, T. Schirmeister, and B. Engels, "On the origin of the stabilization of the zwitterionic resting state of cysteine proteases: a theoretical study," *Journal of the American Chemical Society*, vol. 130, no. 27, pp. 8696–8705, 2008.

[9] E. I. Solomon, "Spectroscopic methods in bioinorganic chemistry: blue to green to red copper sites," *Inorganic Chemistry*, vol. 45, no. 20, pp. 8012–8025, 2006.

[10] T. Vreven and K. Morokuma, "Investigation of the S0 → S1 excitation in bacteriorhodopsin with the ONIOM(MO:MM) hybrid method," *Theoretical Chemistry Accounts*, vol. 109, no. 3, pp. 125–132, 2003.

[11] Y. Hagiwara, T. Ohta, and M. Tateno, "QM/MM hybrid calculation of biological macromolecules using a new interface program connecting QM and MM engines," *Journal of Physics Condensed Matter*, vol. 21, no. 6, 2009.

[12] A. Warshel and M. Levitt, "Theoretical studies of enzymatic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme," *Journal of Molecular Biology*, vol. 103, no. 2, pp. 227–249, 1976.

[13] M. J. Field, P. A. Bash, and M. Karplus, "By hand, to find the optimum values of the parameters by calculating interaction curves for charge/ion systems and comparing them with the MP2/6-311++G** ab initio results," *Journal of Computational Chemistry*, vol. 11, p. 700, 1990.

[14] T. C. Bruice and K. Kahn, "Computational enzymology," *Current Opinion in Chemical Biology*, vol. 4, no. 5, pp. 540–544, 2000.

[15] M. J. Field, "Simulating enzyme reactions: challenges and perspectives," *Journal of Computational Chemistry*, vol. 23, no. 1, pp. 48–58, 2002.

[16] A. J. Mulholland, "Modelling enzyme reaction mechanisms, specificity and catalysis," *Drug Discovery Today*, vol. 10, no. 20, pp. 1393–1402, 2005.

[17] H. Lin and D. G. Truhlar, "QM/MM: what have we learned, where are we, and where do we go from here?" *Theoretical Chemistry Accounts*, vol. 117, no. 2, pp. 185–199, 2007.

[18] H. M. Senn and W. Thiel, "QM/MM studies of enzymes," *Current Opinion in Chemical Biology*, vol. 11, no. 2, pp. 182–187, 2007.

[19] U. Ryde, "Combined quantum and molecular mechanics calculations on metalloproteins," *Current Opinion in Chemical Biology*, vol. 7, no. 1, pp. 136–142, 2003.

[20] J. Gao, S. Ma, D. T. Major, K. Nam, J. Pu, and D. G. Truhlar, "Mechanisms and free energies of enzymatic reactions," *Chemical Reviews*, vol. 106, no. 8, pp. 3188–3209, 2006.

[21] H. M. Senn and W. Thiel, "QM/MM methods for biomolecular systems," *Angewandte Chemie*, vol. 48, no. 7, pp. 1198–1229, 2009.

[22] K. Spiegel and A. Magistrato, "Modeling anticancer drug-DNA interactions via mixed QM/MM molecular dynamics simulations," *Organic and Biomolecular Chemistry*, vol. 4, no. 13, pp. 2507–2517, 2006.

[23] B. Kirchner, F. Wennmohs, S. Ye, and F. Neese, "Theoretical bioinorganic chemistry: the electronic structure makes a difference," *Current Opinion in Chemical Biology*, vol. 11, no. 2, pp. 134–141, 2007.

[24] R. A. Friesner and V. Guallar, "Ab initio quantum chemical and mixed quantum mechanics/molecular mechanics (QM/MM) methods for studying enzymatic catalysis," *Annual Review of Physical Chemistry*, vol. 56, pp. 389–427, 2005.

[25] M. C. Colombo, L. Guidoni, A. Laio et al., "Hybrid QM/MM Car-Parrinello simulations of catalytic and enzymatic reactions," *Chimia*, vol. 56, no. 1-2, pp. 13–19, 2002.

[26] A. Warshel, M. Kato, and A. V. Pisliakov, "Polarizable force fields: history, test cases, and prospects," *Journal of Chemical Theory and Computation*, vol. 3, no. 6, pp. 2034–2045, 2007.

[27] W. L. Jorgensen, "Special issue on polarization," *Journal of Chemical Theory and Computation*, vol. 3, no. 6, p. 1877, 2007.

[28] Y. Hagiwara, H. Matsumura, and M. Tateno, "Functional roles of a structural element involving Na+-π interactions in the catalytic site of T1 lipase revealed by molecular dynamics simulations," *Journal of the American Chemical Society*, vol. 131, no. 46, pp. 16697–16705, 2009.

[29] H. Lin and D. G. Truhlar, "Redistributed charge and dipole schemes for combined quantum mechanical and molecular mechanical calculations," *Journal of Physical Chemistry A*, vol. 109, no. 17, pp. 3991–4004, 2005.

[30] M. J. Field, M. Albe, C. Bret, F. Proust-De Martin, and A. Thomas, "The dynamo library for molecular simulations using hybrid quantum mechanical and molecular mechanical potentials," *Journal of Computational Chemistry*, vol. 21, no. 12, pp. 1088–1100, 2000.

[31] N. Reuter, A. Dejaegere, B. Maigret, and M. Karplus, "Frontier bonds in QM/MM methods: a comparison of different approaches," *Journal of Physical Chemistry A*, vol. 104, no. 8, pp. 1720–1735, 2000.

[32] M. Eichinger, P. Tavan, J. Hutter, and M. Parrinello, "A hybrid method for solutes in complex solvents: density functional theory combined with empirical force fields," *Journal of Chemical Physics*, vol. 110, no. 21, pp. 10452–10467, 1999.

[33] P. Amara and M. J. Field, "Evaluation of an ab initio quantum mechanical/molecular mechanical hybrid-potential link-atom method," *Theoretical Chemistry Accounts*, vol. 109, no. 1, pp. 43–52, 2003.

[34] V. Thery, D. Rinaldi, J. L. Rivail, B. Maigret, and G. G. Ferenczy, "Quantum-mechanical computations on very large molecular systems: the local self-consistent field method," *Journal of Computational Chemistry*, vol. 15, no. 3, p. 269, 1994.

[35] G. Monard, M. Loos, V. Théry, K. Baka, and J. L. Rivail, "Hybrid classical quantum force field for modeling very large molecules," *International Journal of Quantum Chemistry*, vol. 58, no. 2, pp. 153–159, 1996.

[36] X. Assfeld and J. L. Rivail, "Quantum chemical computations on parts of large molecules: the ab initio local self consistent field method," *Chemical Physics Letters*, vol. 263, no. 1-2, pp. 100–106, 1996.

[37] N. Ferré, X. Assfeld, and J. L. Rivail, "Specific force field parameters determination for the hybrid ab initio QM/MM LSCF method," *Journal of Computational Chemistry*, vol. 23, no. 6, pp. 610–624, 2002.

[38] J. Gao, P. Amara, C. Alhambra, and M. J. Field, "A generalized hybrid orbital (GHO) method for the treatment of boundary atoms in combined QM/MM calculations," *Journal of Physical Chemistry A*, vol. 102, no. 24, pp. 4714–4721, 1998.

[39] P. Amara, M. J. Field, C. Alhambra, and J. Gao, "The generalized hybrid orbital method for combined quantum mechanical/molecular mechanical calculations: formulation and tests of the analytical derivatives," *Theoretical Chemistry Accounts*, vol. 104, no. 5, pp. 336–343, 2000.

[40] J. Pu, J. Gao, and D. G. Truhlar, "Generalized hybrid orbital (GHO) method for combining Ab initio Hartree-Fock wave functions with molecular mechanics," *Journal of Physical Chemistry A*, vol. 108, no. 4, pp. 632–650, 2004.

[41] R. C. Walker, I. F. Crowley, and D. A. Case, "The implementation of a fast and accurate QM/MM potential method in Amber," *Journal of Computational Chemistry*, vol. 29, no. 7, pp. 1019–1031, 2008.

[42] G. D. M. Seabra, R. C. Walker, M. Elstner, D. A. Case, and A. E. Roitberg, "Implementation of the SCC-DFTB method for hybrid QM/MM simulations within the Amber molecular dynamics package," *Journal of Physical Chemistry A*, vol. 111, no. 26, pp. 5655–5664, 2007.

[43] H. L. Woodcock, M. Hodošček, A. T. B. Gilbert, P. M. W. Gill, H. F. Schaefer, and B. R. Brooks, "Interfacing Q-Chem and CHARMM to perform QM/MM reaction path calculations," *Journal of Computational Chemistry*, vol. 28, no. 9, pp. 1485–1502, 2007.

[44] M. J. Frisch et al., *Gaussian 03*, Gaussian Inc., Pittsburgh Pa, USA, 2003.

[45] P. Sherwood, A. H. De Vries, M. F. Guest et al., "QUASI: a general purpose implementation of the QM/MM approach and its application to problems in catalysis," *Journal of Molecular Structure*, vol. 632, pp. 1–28, 2003.

[46] J. N. Harvey, "Spin-forbidden CO ligand recombination in myoglobin," *Faraday Discussions*, vol. 127, pp. 165–177, 2004.

[47] J. Torras, G. M. De Seabra, E. Deumens, S. B. Trickey, and A. E. Roitberg, "A versatile AMBER-Gaussian QM/MM interface through PUPIL," *Journal of Computational Chemistry*, vol. 29, no. 10, pp. 1564–1573, 2008.

[48] D. A. Case, T. E. Cheatham III, T. Darden et al., "The Amber biomolecular simulation programs," *Journal of Computational Chemistry*, vol. 26, no. 16, pp. 1668–1688, 2005.

[49] M. W. Schmidt, "General atomic and molecular electronic structure system," *Journal of Computational Chemistry*, vol. 14, pp. 1347–1363, 1993.

[50] A. Laio, J. VandeVondele, and U. Rothlisberger, "A Hamiltonian electrostatic coupling scheme for hybrid Car-Parrinello molecular dynamics simulations," *Journal of Chemical Physics*, vol. 116, no. 16, pp. 6941–6947, 2002.

[51] T. Laino, F. Mohamed, A. Laio, and M. Parrinello, "An efficient real space multigrid QM/MM electrostatic coupling," *Journal of Chemical Theory and Computation*, vol. 1, no. 6, pp. 1176–1184, 2005.

[52] R. H. Holm, P. Kennepohl, and E. I. Solomon, "Structural and functional aspects of metal sites in biology," *Chemical Reviews*, vol. 96, no. 7, pp. 2239–2314, 1996.

[53] E. I. Solomon, R. K. Szilagyi, S. DeBeer George, and L. Basumallick, "Electronic structures of metal sites in proteins and models: contributions to function in blue copper proteins," *Chemical Reviews*, vol. 104, no. 2, pp. 419–458, 2004.

[54] H. Nar, A. Messerschmidt, M. Van De Kamp, G. W. Canters, and R. Huber, "Crystal structure analysis of oxidized Pseudomonas aeruginosa azurin at pH 5.5 and pH 9.0. A pH-induced conformational transition involves a peptide bond flip," *Journal of Molecular Biology*, vol. 221, no. 3, pp. 765–772, 1991.

[55] T. Ohta, Y. Hagiwara, J. Kang et al., "Evaluation of electronic and geometrical properties of the blue copper site in fully solvated azurin by QM/MM hybrid calculations using a new interface program connecting QM and MM engines," *Journal of Computational and Theoretical Nanoscience*, vol. 6, no. 12, pp. 2648–2655, 2009.

[56] J. Kang, T. Ohta, Y. Hagiwara et al., "Electronic and geometric structures of the blue copper site of azurin investigated by QM/MM hybrid calculations," *Journal of Physics Condensed Matter*, vol. 21, no. 6, Article ID 064235, 2009.

[57] S. D. George, L. Basumallick, R. K. Szilagyi et al., "Spectroscopic investigation of stellacyanin mutants: axial ligand interactions at the blue copper site," *Journal of the American Chemical Society*, vol. 125, no. 37, pp. 11314–11328, 2003.

[58] K. Nakatani, C. Dohno, A. Ogawa, and I. Saito, "Suppression of DNA-mediated charge transport by BamHI binding," *Chemistry and Biology*, vol. 9, no. 3, pp. 361–366, 2002.

[59] Y. Hagiwara, H. Kino, and M. Tateno, "Modulation of electronic structures of bases through DNA recognition of protein," *Journal of Physics Condensed Matter*, vol. 22, no. 15, Article ID 152101, 2010.

[60] M. Ibba and D. Söll, "Quality control mechanisms during translation," *Science*, vol. 286, no. 5446, pp. 1893–1897, 1999.

[61] R. Sankaranarayanan and D. Moras, "The fidelity of the translation of the genetic code," *Acta Biochimica Polonica*, vol. 48, no. 2, pp. 323–335, 2001.

[62] M. Tukalo, A. Yaremchuk, R. Fukunaga, S. Yokoyama, and S. Cusack, "The crystal structure of leucyl-tRNA synthetase complexed with tRNA^{Leu} in the post-transfer-editing conformation," *Nature Structural and Molecular Biology*, vol. 12, no. 10, pp. 923–930, 2005.

[63] Y. Hagiwara, O. Nureki, and M. Tateno, "Structural modelling of the complex of leucyl-tRNA synthetase and mis-aminoacylated tRNA^{Leu}," *FEBS Letters*, vol. 583, no. 4, pp. 825–830, 2009.

[64] Y. Hagiwara, O. Nureki, and M. Tateno, "Identification of the nucleophilic factors and the productive complex for the editing reaction by leucyl-tRNA synthetase," *FEBS Letters*, vol. 583, no. 12, pp. 1901–1908, 2009.

[65] Y. Hagiwara, M. J. Field, O. Nureki, and M. Tateno, "Editing Mechanism of Aminoacyl-tRNA synthetases operates by a hybrid ribozyme/protein catalyst," *Journal of the American Chemical Society*, vol. 132, no. 8, pp. 2751–2758, 2010.

[66] M. Boero, "LeuRS Synthetase: A First-Principles Investigation of the Water-Mediated Editing Reaction," *Journal of Physical Chemistry B*, vol. 115, no. 42, pp. 12276–12286, 2011.

[67] L. Ribas de Pouplana and P. Schimmel, "Aminoacyl-tRNA synthetases: potential markers of genetic code development," *Trends in Biochemical Sciences*, vol. 26, no. 10, pp. 591–596, 2001.