Electronic and geometric structures of the blue copper site of azurin investigated by QM/MM hybrid calculations

2009 J. Phys.: Condens. Matter 21 064235

(http://iopscience.iop.org/0953-8984/21/6/064235)

J. Phys.: Condens. Matter **21** (2009) 064235 (8pp) doi:10.1088/0953-8984/21/6/064235

Electronic and geometric structures of the blue copper site of azurin investigated by QM/MM hybrid calculations

Jiyoung Kang[1], Takehiro Ohta[2,3], Yohsuke Hagiwara[1], Keigo Nishikawa[4], Tetsunori Yamamoto[4], Hidemi Nagao[4] and Masaru Tateno[1,2,5]

[1] Graduate School of Pure and Applied Sciences, University of Tsukuba, Tennodai 1-1-1, Tsukuba Science City, Ibaraki 305-8571, Japan
[2] Center for Computational Sciences, University of Tsukuba, Tennodai 1-1-1, Tsukuba Science City, Ibaraki 305-8577, Japan
[3] CREST, Japan Science and Technology Agency, 4-1-8 Honcho, Kawaguchi 332-0012, Japan
[4] Graduate School of Natural Science and Technology, Kanazawa University, Kakuma-machi, Kanazawa, Ishikawa 920-1192, Japan

E-mail: tateno@ccs.tsukuba.ac.jp

Received 19 July 2008, in final form 1 September 2008
Published 20 January 2009
Online at stacks.iop.org/JPhysCM/21/064235

Abstract
The electronic and geometric structures of the copper-binding site in a fully solvated azurin were investigated using quantum mechanics (QM) and molecular mechanics (MM) hybrid calculations. Two types of computational models were applied to evaluate the effects of the environment surrounding the active site. In model I, long-distance electrostatic interactions between QM region atoms and partial point charges of the surrounding protein moieties and solvent water were calculated in a QM Hamiltonian, for which the spin-unrestricted Hartree–Fock (UHF)/density functional theory (DFT) hybrid all-electron calculation with the B3LYP functional was adopted. In model II, the QM Hamiltonian was not allowed to be polarized by those partial point charges. Models I and II provided different descriptions of the copper coordination structure, particularly for the coordinative bonds including a large dipole. In fact, the Cu–O(Gly45) and Cu–S(Cys112) bonds are sensitive to the treatment of long-distance electrostatic interactions in the QM Hamiltonian. This suggests that biological processes occurring in the active site are regulated by the surrounding structures of protein and solvent, and therefore the effects of long-range electrostatic interactions involved in the QM Hamiltonian are crucial for accurate descriptions of electronic structures of the copper active site of metalloenzymes.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Ab initio electronic structure calculations are indispensable to elucidate the functional mechanisms of metalloenzymes, such as electron/proton transfer, catalytic reactions and binding of dioxygen. Density functional theory (DFT) provides useful information on such systems and uses reasonable computational resources [1, 2]. However, for biological macromolecular systems, calculations of the entire system using DFT are too intensive on CPU time and amount of memory used; thus, in those calculations, extracted small modeled structures involving metal active sites have been conventionally exploited so far in combination with polarizable continuum models to represent the surrounding environment of protein and solvent [3].

In order to carry out *ab initio* calculations in which active sites are included in native protein structures, without intensively increasing computational costs, joint methods of quantum mechanics (QM) and molecular mechanics

[5] Author to whom any correspondence should be addressed.

(MM) [4–9] have become a popular strategy to analyze large biological molecules. QM/MM methods allow modeling of long-range electrostatic effects and steric constraints of the surrounding environment. In particular, as for the effects of long-range electrostatic interactions, a number of reports have addressed this issue so far, indicating that such effects are critical to biological macromolecular systems [11–18]. Accordingly, the long-range electrostatic interactions are presumed to also affect the electronic structure of the metal active sites significantly, and thus proper modeling of the environment using a QM/MM method is required for the study of metalloenzymes.

The energy expression for QM/MM calculations is classified into two types: additive and subtractive schemes [10]. In an additive scheme, the QM energy of a QM inner region and the MM energy of an MM outer region are added together, and the interaction energy between the QM and MM regions is added to complete the energy scheme. On the other hand, the subtractive scheme is obtained by calculating the QM energy of the inner region and the MM energy of the entire molecule, from which the inner region energy obtained by MM calculation is subsequently subtracted [5, 10].

One important application of QM/MM methods for metalloenzymes is investigation of active sites in blue copper proteins [19–23], for which the biological function is electron transfer (ET) through the bound Cu ion(s) [24]. X-ray crystallographic analyses have been reported for such enzymes including azurin, plastocyanin and stellacyanin. Common features of one-cysteine coordination and two-histidine coordination to the copper centers, forming an approximate trigonal plane, are observed and their coordination environments completed with weak axial ligand(s) [22, 24]. Azurin includes five-coordinated copper having methionine (Met) and a backbone carbonyl oxygen of glycine (Gly) as weak axial ligands. Stellacyanin and plastocyanin have a four-coordinated copper site coordinated with glutamine (Gln) and Met as a weak axial ligand, respectively. Those weakly coordinative bonds are presumed to contribute critically to the redox potential related to electron transfer, thereby regulating biological functions. Thus, blue copper proteins have been used so far as standard metalloenzymes to examine the accuracy of calculations [1, 19, 20, 23].

Here, we report a QM/MM study of azurin using our newly developed program [26, 27] that interfaces the quantum chemical calculation program GAMESS [28] with the molecular dynamics simulation program AMBER [29]. Exploiting our interface program, we adopted the two distinct QM/MM schemes mentioned, i.e. additive and subtractive schemes, to evaluate long-range effects on the copper active site. Thus, an advantage of our strategy is that the interface program enables us to compare results of QM/MM calculations using the distinct schemes; to the best of our knowledge, this is the first case to analyze such calculations in terms of detailed geometries and electronic structures of the active centers in proteins. In fact, the results presented indicate that the treatment of the electrostatic interactions of the metal site, the surrounding protein moiety and the solvent in the QM Hamiltonian is important for accurate description of the electronic structure and geometry of the blue copper site in azurin. Thus, it is suggested that electronic structures of active sites are actually modulated by the surrounding regions through long-range electrostatic interactions, resulting in contributions to the biological functions of metalloenzymes, such as electron transfer.

2. Methodology

2.1. QM/MM calculations

In our QM/MM method, a macromolecular system of interest is first divided into two regions, i.e. a region including an active site in a protein where the QM calculation is performed and a region outside the QM region that is treated with an MM potential; then, the MM region within a distance r from the center of mass of the QM region is defined as the MM1 region, and the region outside the MM1 region is defined as the MM2 region.

The total Hamiltonian of our QM/MM scheme is expressed as the sum of a QM Hamiltonian, a classical MM-based Hamiltonian, a QM–MM1 hybrid Hamiltonian and a QM–MM2 hybrid Hamiltonian [26, 27]. A characteristic feature of our hybrid scheme is that the partial charges of the MM1 atoms are incorporated into the one-electron integrals, and thereby polarization of the QM region by MM1 atoms can be considered: the QM–MM1 hybrid Hamiltonian evaluates interactions of the nuclei of the QM atoms with partial charges and the Lennard-Jones (LJ) potential representing the van der Waals (vdW) energy between the QM and MM1 atoms. The QM–MM2 hybrid Hamiltonian calculates the Coulomb and vdW interactions between QM and MM2 atoms at the MM level. In our QM/MM scheme, the LJ potential is incorporated into the QM Hamiltonian since DFT, which is used for QM calculation in this study, is known to fail in the estimation of vdW interaction energy [25]. Detailed descriptions of our QM/MM calculation schemes are presented in [26, 27].

2.2. Setup of the azurin structure for QM/MM calculation

The coordinates of a crystal structure of the *Pseudomonas aeruginosa* azurin were obtained from the Protein Data Bank (PDB) (PDB ID: 4AZU). Our procedures to set up the solvated molecular system and calculation schemes to relax the system using molecular dynamics (MD) simulations are described previously [26]; the number of total atoms of the solvated protein system was 25 618 (figures 1(a) and (b)).

For QM/MM calculations, QM atoms were assigned for the active site region (figure 2(a)), as shown in figure 1(c). The crystal water molecule that forms an H bond with the H atom of the imidazole moiety of His117 was also included in the QM region to investigate the effect of the H bond on the electronic structure and the coordination geometry of the active copper site. QM calculations were performed with GAMESS [28] using the UHF/DFT hybrid B3LYP [30] functional in combination with the triple zeta valence (TZVP) basis set for copper [31] and the 6-31G^* basis set [32] for other atoms. In order to evaluate the basis set dependence, we adopted other basis sets, i.e. the 6-311G^*, 6-311G^{**}, 6-31+G^* and 6-311 + +G^* basis sets [32–37], instead of the 6-31G^*

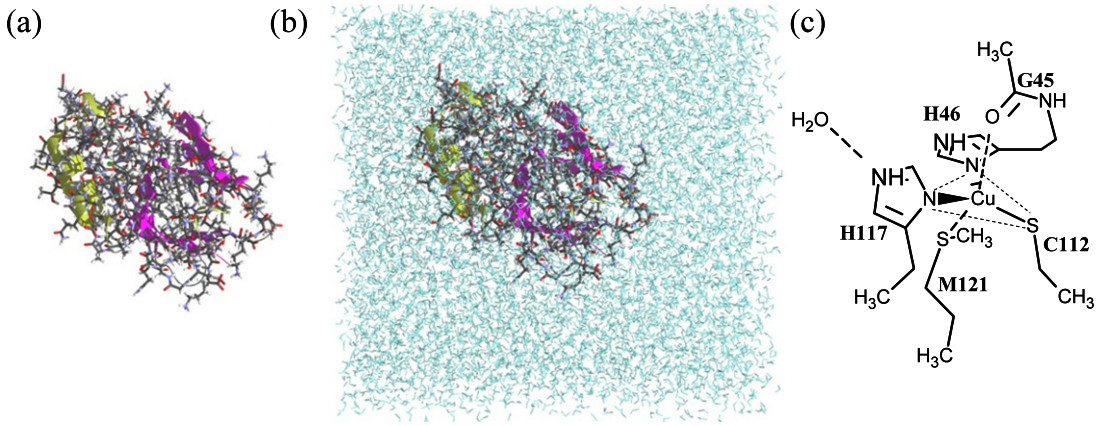

Figure 1. Structural model of azurin built by the addition of H atoms to the crystal structure (PDB: 4AZU) (a), the calculation model of azurin solvated with TIP3P water molecules (b) and the QM region selected by termination using the link atom scheme (c).

Table 1. Optimized geometries of the active site of azurin [26].

	Xtal[a]	MD–MM	Model I	Model II	ONIOM[b]	Xtal[c]	EXAFS[d]
$Cu–S_{Met121}$	3.16	3.35	3.49	3.50	3.53(3.41)	2.87–3.26	3.39
$Cu–S_{Cys112}$	2.27	2.25	2.20	2.24	2.17(2.17)	2.12–2.27	2.12
$Cu–O_{Gly45}$	2.95	2.96	3.01	2.81	2.55(2.49)	2.75–3.16	2.82
$Cu–N_{His117}$	1.98	1.95	2.00	2.10	2.01(2.03)	1.99–2.12	1.94
$Cu–N_{His46}$	2.06	2.06	2.03	1.93	1.99(2.01)	1.99–2.12	1.86
$O_{H_2O}–N_{His117}$	2.93	2.95	3.00	3.04	—	—	—
$O_{H_2O}–H_{His117}$	—	1.95	1.98	2.03	—	—	—
$H–N_{His117}$	—	1.02	1.03	1.02	—	—	

[a] Experimental values in the crystal structure used in this study as the initial structure of the MD simulation.
[b] See [22]; EE(ME) optimized geometries.
[c] Experimental values summarized in [19].
[d] Experimental values listed in [22].

basis set. In this study, only QM region atoms were allowed to relax by fixing all MM atoms during geometry optimization.

The effects of electrostatic interactions from the protein and solvent around the copper active site were analyzed in this study, using the following two QM/MM models. In model I, the partial point charges of the protein and solvent within a distance of 25 Å from the center of the QM region ($r = 25$ Å) were defined as the MM1 region, and thus were allowed to perturb the QM Hamiltonian for the QM inner region. In model II, r was set as 0, i.e. the electrostatic interaction between the inner region and the surrounding environment was not evaluated in the QM calculation. Thus, the QM/MM Hamiltonian of model II is close to that of a subtractive scheme, while the additional van der Waals energy in the QM region is included in our model. In this way, the effect of the additive and subtractive QM/MM energy schemes on the description of the electronic and geometric properties of the metal center would be clarified by comparing the results of models I and II.

3. Results and discussion

3.1. Comparison of x-ray crystallographic analyses and QM/MM studies

Since the detailed geometric features of the active site would be directly concerned with the functional properties, such as the redox potential and the ET rates [1, 19, 20, 23, 24, 38], unraveling the accurate geometry of the copper coordination site is of particular importance in elucidating the structure/function correlation of metalloenzymes. However, to evaluate the geometric properties of the azurin copper site, it should be noted that the available crystallographic data for this site are not sufficiently accurate to allow exact comparisons with the calculated data. In fact, the crystal structure of azurin II obtained from *A. xylosoxidans* was also reported at the higher resolution of 1.13 Å [22], but it was pointed out that the intense x-ray beam reduced the blue copper site during data collection, and thus ambiguities are still present in the experimental data (table 1).

Accordingly, computational investigations are crucial to elucidate the outstanding electronic and geometric features of the weak axial Cu–S(Met121) and Cu–O(Gly45) ligands and the relatively strong Cu–S(Cys112) ligand observed in azurin. Fortunately, we can anticipate that experimental values of the spin populations in the blue copper site obtained using spectroscopic techniques [1, 23] can be used to assess the results of our calculations.

3.2. Molecular dynamics simulation of fully solvated azurin

In order to obtain an equilibrated structure of the protein in the environment of an aqueous solution, we constructed a

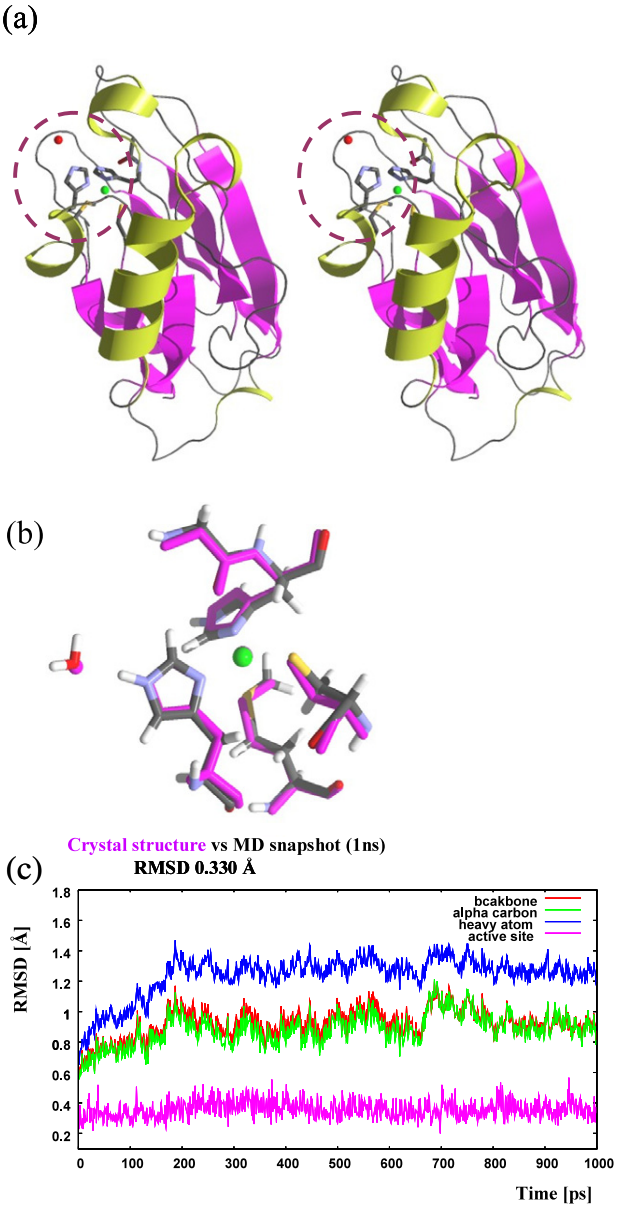

Figure 2. Stereo view of the crystal structure of azurin (PDB: 4AZU); the active site is shown by the circle (broken line) (a), overlay drawing of the blue copper site; the crystal structure versus 1 ns MD snapshot (b) and the RMSD plot for the 1 ns MD simulation (c) [26].

structural model of fully solvated azurin and then performed MD simulation for 1 ns. In the calculation, the structure of the protein was very stable as shown in figure 2(c); in particular, for the amino acid residues consisting of the copper site, the structure converged in the early phase of the simulation (∼300 ps) and the averaged root mean square deviations (RMSDs) after the convergence were as small as ∼0.3 Å. The final snapshot of the 1 ns MD simulation was subjected to energy minimization and thus we obtained a completely solvated structure of the protein optimized at the MM–MD level (figure 2(b)), and also obtained a reasonable copper coordination geometry, as indicated in the column MD–MM (table 1). Accordingly, we used the MD–MM coordinate as the starting geometry for our QM/MM calculations.

3.3. Geometry optimization using distinct QM/MM calculation schemes

In the QM/MM geometry optimization for model I, the RMS force was reduced to 0.13 kcal mol^{-1} Å^{-1} after 140 cycles. As shown in figure 3(a), the RMSD value of the copper site in model I against that of the azurin crystal structure is 0.311 Å, which is smaller than the value for the initial structure for the QM/MM calculation optimized using an MM potential (0.330 Å; figure 2(b)). This indicated that the static QM/MM optimized structure is more consistent with the crystal structure, which can be considered as the averaged structure in the thermal fluctuation. On the other hand, the structure of model II was hardly optimized, showing a slow decrease in the RMS force in the calculation; a larger RMS force of 1.84 kcal mol^{-1} Å^{-1} still remained even after 100 cycles of geometry optimization. Thus, we terminated the geometry optimization of model II at 100 cycles. Nevertheless, the 100 cycle structure of model II did not deviate significantly from the crystal structure, although it showed a slightly larger RMSD of 0.332 Å (figure 3(b)). Thus, the result of model II can be used in the following discussions to delineate the effect of the different QM/MM energy schemes on the geometric and electronic properties of the copper site.

3.4. Comparison of geometric and electronic features

As for the Cu–S(Met121) coordination, comparison of models I and II shows almost the same Cu–S(Met121) distances of ∼3.5 Å, indicating that, in our QM/MM energy scheme, the weak Cu–S(Met121) coordination is somewhat insensitive to the polarization effect. For the Cu–O(Gly45) coordination, Hasnain and coworkers obtained remarkably short Cu–O(Gly45) distances of 2.55 and 2.49 Å in both the presence and absence of the electrostatic energy term, respectively [22]. Thus, the results of our calculations showed much better agreement with the experimental data, compared with the results of the ONIOM calculations. In addition, we found a significant difference in the Cu–O(Gly45) distance between the calculated structures, 3.0 and 2.8 Å, for models I and II, respectively, indicating that this weak axial Cu–O(Gly45) coordination is sensitive to treatment of the electrostatic interaction in the QM Hamiltonian. This is a distinctive feature between the Cu–S(Met121) and Cu–O(Gly45) bonds. It should be noted here that the Cu–O(Gly45) bond is fairly well polarized, whereas the Cu–S(Met121) bond is just slightly polarized, as discussed below (table 2).

Next, to investigate electronic structures related to those bonds, we searched for molecular orbitals (MOs) that most dominantly include electrons involved in S(Met121)/O(Gly45) atoms in models I and II. Here, O(Gly45) is included in a peptide group of the protein backbone, and the electrons of the atom are delocalized onto the peptide group; therefore, we sum up contributions of all electrons of the peptide group involving O(Gly45) atoms. To compare energy levels of MOs between models I and II, the lowest unoccupied molecular orbitals (LUMOs) of β electrons are used as the reference, since LUMOs in the two models correspond to each other.

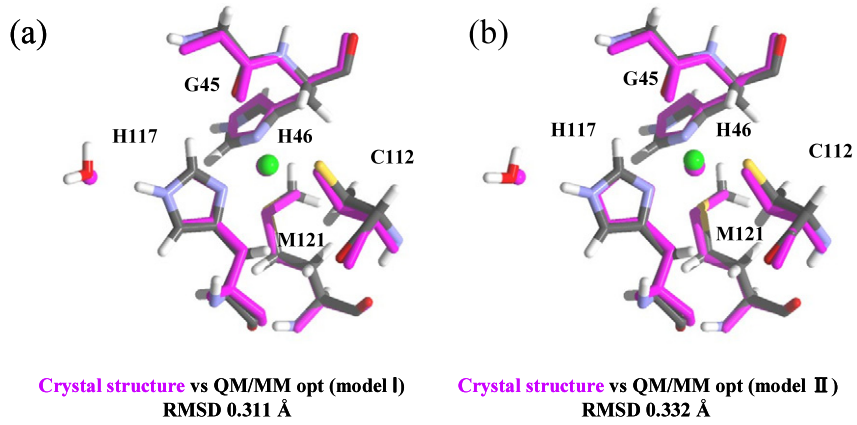

Figure 3. Overlay drawings of the copper site in azurin; the crystal structure versus model I (a) and model II (b). RMSD values are also indicated.

Table 2. Mulliken charge (spin population).

		6-31G^{*}[b]		6-311G^{*}		6-311G^{**}		6-31 + G^{*}		6-311 + +G^{**}	
	Basis set[a]	Charge	Spin	Charge	Spin	Charge	Spin	Charge	Spin	Charge	Spin
Cu	model I	0.489	0.391	0.603	0.388	0.504	0.387	0.516	0.386	0.482	0.387
	model II	0.483	0.341	0.562	0.342	0.466	0.342	0.488	0.341	0.451	0.341
S_{Cys112}	model I	−0.088	0.499	−0.144	0.505	−0.262	0.507	−0.159	0.502	−0.284	0.506
	model II	−0.064	0.568	−0.024	0.570	−0.124	0.571	−0.068	0.567	−0.143	0.571
S_{Met121}	model I	0.057	0.000	0.100	0.000	−0.068	0.000	0.023	0.000	−0.127	0.000
	model II	0.072	0.000	0.149	0.000	−0.015	−0.001	0.065	0.000	−0.072	0.000
O_{Gly45}	model I	−0.523	0.001	−0.426	0.001	−0.452	0.001	−0.519	0.001	−0.462	0.001
	model II	−0.515	0.002	−0.431	0.001	−0.455	0.001	−0.516	0.002	−0.464	0.001
N^{δ}_{His117}	model I	−0.573	0.053	−0.506	0.051	−0.322	0.050	−0.587	0.052	−0.304	0.050
	model II	−0.548	0.032	−0.473	0.032	−0.322	0.032	−0.557	0.032	−0.312	0.031
H^{ε}_{His117}	model I	0.402	0.000	0.408	0.000	0.399	0.000	0.313	0.000	0.167	0.000
	model II	0.393	0.000	0.399	0.000	0.393	0.000	0.307	0.000	0.168	0.000
N^{δ}_{His46}	model I	−0.513	0.052	−0.433	0.050	−0.233	0.050	−0.523	0.050	−0.224	0.050
	model II	−0.528	0.052	−0.452	0.052	−0.228	0.052	−0.538	0.052	−0.217	0.052

[a] Basis set used for atoms except for Cu, for which the TZVP basis set was applied in all calculations.
[b] Reported in [26].

As a result, for the Cu–S(Met121) coordination, the highest occupied molecular orbitals (HOMOs) of α electrons, which are the 128 MOs in both models, include electrons of the S(Met121) atoms most dominantly in both models and they are equivalent to each other. The energy difference of the two HOMOs is as small as 1.8 kcal mol^{-1}, which can be considered to be marginal (figure 4(a)). This is almost the case with β electrons of S(Met121) (figure 4(b)).

On the other hand, for the Cu–O(Gly45) coordination, the 96(124) orbital of the $\alpha(\beta)$ electron in model I corresponds to the 97(122) orbital of the $\alpha(\beta)$ electron in model II, which most dominantly include electrons of the peptide group involving O(Gly45). The energy difference between the two MOs is 8.6(10.0) kcal mol^{-1}; these values are definitely larger than those of S(Met121) (figures 4(c) and (d)). Accordingly, it is likely that the energy level of electrons of O(Gly45) is shifted through the electrostatic interactions in the QM Hamiltonian, suggesting that this polarized coordinative bond is sensitive to the environment surrounding the copper active site.

For the Cu–S(Cys) coordination, this short bond distance is one of the key features of blue copper proteins; therefore, an accurate description of the Cu–S(Cys) coordination is crucial to elucidate the ET mechanism. Its calculated bond lengths are also fundamentally within the range of the crystallographic data; however, by comparison between the computational results of models I and II, different bond distances of 2.20 and 2.24 Å, respectively, were found, indicating that polarization of the electronic structure of the copper site can affect the Cu–S(Cys112) coordination. The ONIOM calculations did not lead to changes in the Cu–S(Cys112) bond length (2.17 Å) in both the absence and presence of long-range electrostatic interactions [22], indicating that the polarization effect on Cu–S(Cys112) coordination did not influence its geometry in the ONIOM calculation. In contrast, in our calculations, the Cu–S(Cys112) distance was 2.24 Å, which is remarkably similar to that of the initial structure of our QM/MM calculations, thus suggesting that the subtractive scheme in model II cannot correctly describe the Cu–S(Cys112) coordination.

The Cu–N(His) bonds may be expected to be less polarizable than other coordinations due to the harder donor of the nitrogen atom of His compared to the softer donors of

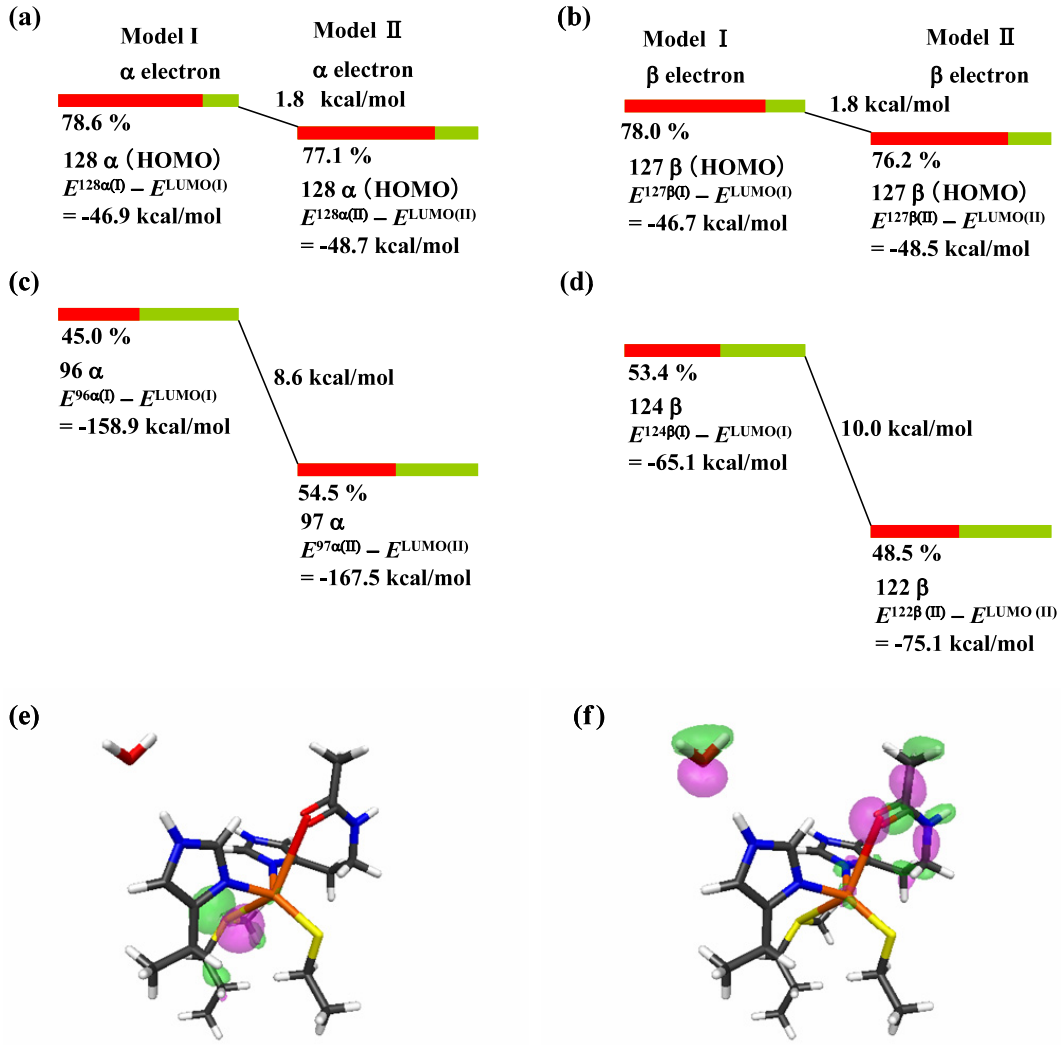

Figure 4. Energy differences of equivalent molecular orbitals (MOs) of models I and II. MOs in which S(Met121) and O(Gly45) atoms are most dominantly included are depicted in (a)/(b) and (c)/(d), respectively. The ratio of the sum of occupancies of the atomic orbitals belonging to the S(Met121) atom ((a)/(b)) or the peptide group involving the O(Gly45) atom ((c)/(d)) in each MO are represented by a red bar. MOs of the 128 α electrons (e) and the 96 α electrons (f) in model I are shown.

the oxygen and sulfur atoms of the other residues. However, inclusion of the explicit electrostatic interaction in the QM Hamiltonian alters the strength of the H bond between H_{His117} and the water molecule, resulting in an O_{H_2O}–H_{His117} distance of 1.98 Å in model I, which is shorter than that in model II (2.03 Å). Consequently, more polarized His117, due to the stronger H bond in model I, yielded stronger Cu–N_{His117} coordination, showing a shorter distance of 2.0 Å than that of 2.1 Å obtained in model II. On the other hand, the counter His (His46) coordinates to the copper more strongly in model II than in model I, as inferred from their bond distances of 1.93 and 2.03 Å, respectively. Thus, an inverse relationship between the strength of the Cu–His117 and Cu–His46 coordinations is apparently observed. Here, our results clearly show that the H bond between His117 and water can be polarizable by the electrostatic effect, which in turn influences the copper coordination environment, again indicating the importance of the additive QM/MM energy scheme for calculation of the copper site in azurin.

3.5. Population analysis

The singly occupied molecular orbital (SOMO) of the copper site of blue copper proteins is known as an antibonding combination of a Cu 3d orbital and a cysteine sulfur 3p orbital [1, 23, 39]. In models I and II, the 128 α- and 127 β-spin orbitals are occupied and the SOMO of models I and II were found in the 126 α-spin orbitals, as shown in figures 5(a), and (b), respectively. In agreement with previous calculations, the Cu 3d orbital and the cysteine sulfur 3p orbital contribute to the SOMO in both models I and II, while some differences in the orbital pattern can be seen at the Gly45 and His117 residues.

Our Mulliken population analyses of models I and II are given in table 2. We first discuss the calculations using the 6-31G^* basis set and then compare the result with the other calculations using various basis sets, since Mulliken charges generally show strong basis set dependence [41, 42].

In the spectroscopic experiments, 45% S 3p orbital character was observed in the SOMO of azurin [1, 43]. The

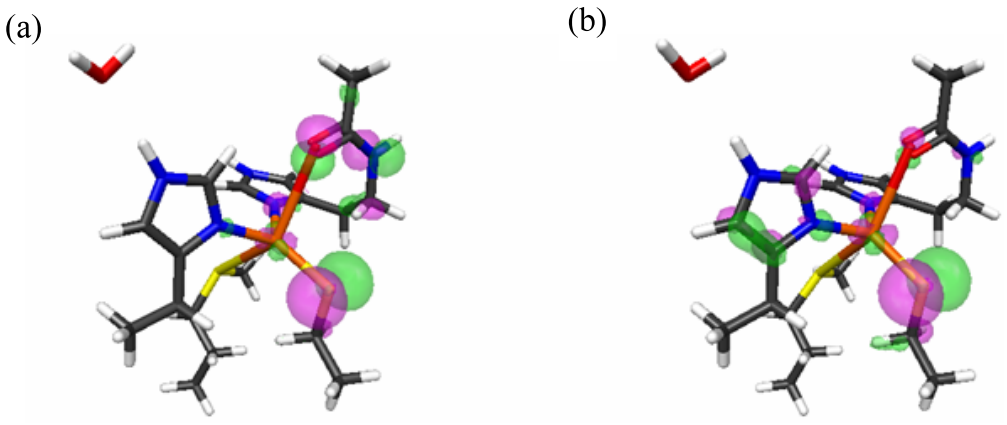

Figure 5. The singly occupied molecular orbitals (SOMOs) in models I (a) and II (b) are shown, both of which are the 126 α-spin orbitals. The figures were generated using the MacMolPlt program [40].

Mulliken charge of the S atom in Cys112 indicates that the electron density of the S atom is significantly delocalized onto the copper d-orbital, resulting in a large spin density on the S atom. In our QM/MM calculation, the spin polarization is significantly improved in model I (49.9%) compared to model II (56.8%), indicating that the explicit inclusion of electrostatic interaction by the QM Hamiltonian is essential for the calculation of the copper site.

The oxygen atom of the backbone of Gly45 is negatively charged in both models I and II; thus, it can be considered that the Cu–O(Gly45) bond involves a relatively large dipole, which interacts with the partial point charges of the protein and solvent water. In fact, models I and II provide different Cu–O(Gly45) bond distances, corresponding to the presence or absence of such long-distance electrostatic interaction, respectively. On the other hand, the S(Met121) atom is almost neutral in both models I and II, which seems consistent with the result that models I and II yielded almost the same Cu–S(Met121) distances of ∼3.5 Å. For His117, in terms of the Mulliken charges, N^{δ} (His117) and H^{ε} (His117) in model I (–0.57 and 0.40) are more negatively and positively charged, respectively, when compared to those in model II (–0.55 and 0.39). It should be noted here that H^{ε} (His117) forms a stronger hydrogen bond with the water molecule involved in the QM region and thereby His117 in model I could be more polarized than that in model II.

Hence, it is suggested that to describe the electronic structure and geometric parameters of metal active sites in metalloenzymes accurately, the QM Hamiltonian of the inner region should be modified to be perturbed by long-range interactions with the surrounding environment. Finally, as for the basis set dependence of Mulliken charges, the values actually showed a strong dependence on the basis sets used in the QM calculations, as reported in the previous studies [41, 42]; however, the characteristic tendency discussed above, which has been found in the calculations using the 6-31G^* basis set, is substantially consistent with that using the other extended basis sets. Thus, our conclusion is independent of the basis sets used in the calculations.

4. Conclusions

In this study, we used our new interface program for QM/MM hybrid calculations that connects QM and MM engines and applied it to a blue copper protein of azurin. The additive and subtractive QM/MM energy schemes provided different descriptions of the copper coordination geometry, particularly when the coordinative bond included a large dipole: models I and II provided different bond distances for the Cu–O(Gly45) and Cu–S(Cys112) bonds. In addition, the H bond between H_{His117} and the water molecule on the surface of the protein was also sensitive to the QM treatment of the electrostatic interaction from the protein and solvent water, which in turn affected the Cu–His coordination significantly.

Thus, our computational results are in good agreement with experimental data, compared with some reports in the literature—they could also be dependent on the starting structure of the QM/MM hybrid calculations. As for our model system, the completely solvated protein structure was well equilibrated in the environment of an aqueous solution using MD simulations, followed by QM/MM calculations. Furthermore, the additive QM/MM energy scheme is crucial for accurate calculation of the spin polarization in the Cu–S(Cys) bond. Thus, to describe the electronic and geometric properties of the metal active sites in metalloenzymes, an additive energy scheme would be preferable to include the effect of polarization, since this may play a key role in the biological functions of proteins.

Acknowledgments

This work is partly supported by grants-in-aid from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) under contract nos. 19019003, 19340108 and 20051003. Computations were performed on the computer facilities under the ‘Interdisciplinary Computational Science Program’ at the Center for Computational Sciences, University of Tsukuba and at the Computer Center for Agriculture, Forestry, and Fisheries Research, MAFF, Japan.

References

[1] Solomon E I, Szilagyi R K, DeBeer George S and Basumallick L 2004 *Chem. Rev.* **104** 419
[2] Siegbahn P E 2006 *J. Biol. Inorg. Chem.* **11** 695
[3] Siegbahn P E and Borowski T 2006 *Acc. Chem. Res.* **39** 729
[4] Morokuma K 2007 *Bull. Chem. Soc. Japan.* **80** 2247
[5] Lin H and Truhlar D G 2007 *Theor. Chem. Acc.* **117** 185
[6] Senn H M and Thiel W 2007 *Curr. Opin. Chem. Biol.* **11** 182
[7] Bruice T C 2006 *Chem. Rev.* **106** 3119
[8] Riccardi D, Schaefer P, Yang Y, Yu H, Ghosh N, Prat-Resina X, Konig P, Li G, Xu D, Guo H, Elstner M and Cui Q 2006 *J. Phys. Chem.* B **110** 6458
[9] Murphy R B, Philipp D M and Friesner R A 2000 *J. Comput. Chem.* **21** 1442
[10] Sherwood P 2000 *Modern Methods and Algorithms of Quantum Chemistry* vol 1, ed J Grotendorst (Jülich: John von Neumann Institute for Computing) pp 257–77
[11] Ryde U 2003 *Curr. Opin. Chem. Biol.* **7** 136
[12] Mulholland A 2005 *Biosilico* **10** 1393
[13] Jayapal P, Sundararajan M, Hillier I H and Burton N A 2008 *Phys. Chem. Chem. Phys.* **10** 4249
[14] Markus K, Söderhjelm P, Heimdal J and Ryde U 2008 *J. Chem. Theor. Comput.* **4** 985
[15] Mladenovic M, Fink R F, Thiel W, Schirmeister T and Engels B 2008 *J. Am. Chem. Soc.* **130** 8696
[16] Solomon E I 2006 *Inorg. Chem.* **45** 8012
[17] Morokuma K 2007 *Bull. Chem. Soc. Japan* **80** 2247
[18] Lin H and Truhlar D G 2007 *Theor. Chem. Acc.* **117** 185
[19] Ryde U, Olsson M H M, Roos B O, De Kerpel J O A and Pierloot K 2000 *J. Biol. Inorg. Chem.* **5** 565
[20] Ryde U and Olsson M H M 2001 *Int. J. Quantum Chem.* **81** 335
[21] Sinnecker S and Neese F 2006 *J. Comput. Chem.* **27** 1463
[22] Paraskevopoulos K, Sundararajan M, Surendran R, Hough M A, Eady R R, Hillier I H and Hasnain S S 2006 *Dalton Trans.* 3067
[23] Solomon E I 2006 *Inorg. Chem.* **45** 8012
[24] Gray H B, Malmström B G and Williams R J P 2000 *J. Biol. Inorg. Chem.* **5** 551
[25] Cybulski S M and Seversen C E 2005 *J. Chem. Phys.* **122** 014117
[26] Ohta T, Hagiwara Y, Kang J, Nishikawa K, Yamamoto T, Nagao H and Tateno M 2009 *J. Comput. Theor. Nanosci.* at press
[27] Hagiwara Y, Ohta T and Tateno M 2009 *J. Phys.: Condens. Matter* at press
[28] Schmidt M W *et al* 1993 *J. Comput. Chem.* **14** 1347
[29] Case D A 2006 *AMBER 9* University of California, San Francisco
[30] Becke A D 1993 *J. Chem. Phys.* **98** 5648
[31] Schaefer A, Huber C and Ahlrichs R 1994 *J. Chem. Phys.* **100** 5829
[32] Rassolov V A, Pople J A, Ratner M A and Windus T L 1998 *J. Chem. Phys.* **109** 1223
[33] Hehre W J, Ditchfield R and Pople J A 1972 *J. Chem. Phys.* **56** 2257
[34] Hariharan P C and Pople J A 1973 *Theor. Chim. Acta* **28** 213
[35] Francl M M, Petro W J, Hehre W J, Binkley J S, Gordon M S, DeFrees D J and Pople J A 1982 *J. Chem. Phys.* **77** 3654
[36] Clark T, Chandrasekhar J, Spitznagel G W and Schleyer P von R 1983 *J. Comput. Chem.* **4** 294
[37] Krishnan R, Binkley J S, Seeger R and Pople J A 1980 *J. Chem. Phys.* **72** 650
[38] Holm R H, Kennepohl P K and Solomon E I 1996 *Chem. Rev.* **96** 2239
[39] Sugiyama A, Sugimori K, Shuku T, Nakamura T, Saito H, Nagao H, Kawabe H and Nishikawa K 2005 *Int. J. Quantum Chem.* **105** 588
[40] Bode B M and Gordon M S 1998 *J. Mol. Graph. Model.* **16** 133
[41] Julg A 1975 *Top. Curr. Chem.* **58** 1
[42] Smith V H Jr 1977 *Phys. Scr.* **15** 147
[43] George S D, Basumallick L, Szilagyi R K, Randall D W, Hill M G, Nersissian A M, Valentine J S, Hedman B, Hodgson K O and Solomon E I 2003 *J. Am. Chem. Soc.* **125** 11314