

Article

# Linear Response Functions of Densities and Spin Densities for Systematic Modeling of the QM/MM Approach for Mono- and Poly-Nuclear Transition Metal Systems

Colin K. Kitakawa <sup>1</sup> , Tomohiro Maruyama <sup>1</sup>, Jinta Oonari <sup>1</sup>, Yuki Mitsuta <sup>1</sup>, Takashi Kawakami <sup>1</sup>, Mitsutaka Okumura <sup>1</sup>, Kizashi Yamaguchi <sup>2</sup> and Shusuke Yamanaka <sup>1,3,\*</sup>

<sup>1</sup> Graduate School of Science, Osaka University, Osaka 565-0871, Japan; kitakawac15@chem.sci.osaka-u.ac.jp (C.K.K.); maruyamat16@chem.sci.osaka-u.ac.jp (T.M.); oonarij16@chem.sci.osaka-u.ac.jp (J.O.); mitsutay13@chem.sci.osaka-u.ac.jp (Y.M.); kawakami@chem.sci.osaka-u.ac.jp (T.K.); ok@chem.sci.osaka-u.ac.jp (M.O.)

<sup>2</sup> Institute for Nanodesign, Osaka University, Osaka 560-8531, Japan; yama@chem.sci.osaka-u.ac.jp

<sup>3</sup> Quantum information, Quantum Biology division, Institute for Open and Transdisciplinary Research Initiatives, Osaka University, Osaka 560-8531, Japan

\* Correspondence: syama@chem.sci.osaka-u.ac.jp; Tel.: +81-6-6850-5406

Academic Editors: Yasutaka Kitagawa, Ryohei Kishi and Masayoshi Nakano

Received: 30 December 2018; Accepted: 15 February 2019; Published: 25 February 2019



**Abstract:** We applied our analysis, based on a linear response function of density and spin density, to two typical transition metal complex systems—the reaction centers of P450, and oxygen evolving center in Photosystem II, both of which contain open-shell transition metal ions. We discuss the relationship between LRF of electron density and spin density and the types of units and interactions of the systems. The computational results are discussed in relation to quantum mechanics (QM) cluster and quantum mechanics/molecular mechanics (QM/MM) modeling that are employed to compute the reaction centers of enzymes.

**Keywords:** linear response function; QM/MM; transition metal complexes; enzymes

## 1. Introduction

The quantum mechanics/molecular mechanics (QM/MM) approach, which was first proposed in the 1970s [1,2], is now widely applied to biochemical reactions and molecular materials [3–6]. The essential idea of this approach is that the entire system can be divided into the QM region, that is described by quantum mechanics and the MM region, that is described by using classical point charges and force fields. The region of primary interest is the QM region, where QM events, such as chemical reactions and electronic excitation processes occur. The MM region is, however, also important, because the surrounding environment, such as proteins, lipids, and solvents often strongly affect the electronic structure of the QM region. For this reason, in most contemporary QM/MM methods, the electronic structure of the QM region is calculated in the presence of the MM point charges. This “electric embedding” type of QM/MM treatment usually leads to an improved modeling of the target system as desired, but sometimes it leads to over-polarization of the QM region near the QM/MM boundaries [4], resulting in less accuracy. To begin with, the validity of the QM description critically depends on the appropriate choices of important residues, lipids, and water molecules that are included in the QM region [5]. Then it is the most important step of the QM/MM calculation to determine how we model the target system. Some of the QM/MM researchers carefully examined the convergence of their QM/MM computational results with the size of the QM region [6–15]. These types

of careful approaches for the QM/MM modeling is not new: Karplus and his coworkers analyzed the contributions of residues to the reaction pathway of Triosephosphate isomerase in their study in 1991 [6]. In this type of approach, the researchers examined several models, in which the residues included in the QM regions differed for each model [6]. This charge (atom) deletion analysis (CDA) method was originally used to find residues that are important for the catalytic activity of the reaction centers [6], but it is also used to examine the accuracy of the energies of QM/MM models [7–9]. As an extended version of CDA, Ryde and his coworkers performed a very comprehensive analysis of the effects of QM and QM/MM modeling on the proton transfer in [Ni, Fe] hydrogenase [8]. An alternative version is the charge shift analysis (CSA) proposed by Kulik and her coworkers [9]. Instead of focusing on the energy differences among the models, they estimated the difference of the charge of non-core residues between the full enzyme model and the model in which the active site residues have been removed, as follows:

$$\Delta q_{Res} = q_{Res}^{holo} - q_{Res}^{apo} \quad (1)$$

Here, we use the notation *holo*, and *apo*, respectively for the former and the latter models according to their papers [9,10]. The validity of the CSA model is based on their following observations [9]: The residues that exhibited a charge redistribution when the core active site substrates were removed could be used to construct an appropriate QM region, which is consistent with the sufficiently accurate QM region that reaches at the asymptotically converged QM region. They further proposed the Fukui shift analysis (FSA) [10]: In FSA, they estimated the condensed Fukui function (CFF) [16,17] of the core active site, and then they added surrounding QM residues back, one at a time and examined the CFF values of the core site. If the change of the CFF values of the core site due to the addition of the surrounding residue is large, the residue is determined to be included in the QM region. The heart of the FSA approaches is that the CFF values are used to judge whether the peripheral residues affect the electronic structure of the reaction center.

We also proposed an approach to determine the QM/MM boundary [18–21]. Our approach shares the same goal, i.e., to determine the QM/MM boundary, but it is a simpler method than the CSA or the FSA, and is based on the linear response function (LRF) [22–24] defined by,

$$\frac{\delta\rho(\mathbf{r})}{\delta v(\mathbf{r}')}. \quad (2)$$

In our method [18–21],  $\delta v(\mathbf{r}')$  is a virtual perturbation, which is applied to the point  $\mathbf{r}'$  (a point in the peripheral regions surrounding the active site), and  $\delta\rho(\mathbf{r})$ , which is the density change at the QM region, due to the perturbation. If  $\delta\rho(\mathbf{r})/\delta v(\mathbf{r}')$  is large enough, the replacement of the peripheral regions by MM point charges is inappropriate. Since  $\delta\rho(\mathbf{r})$  includes the effects of electron inflow and outflow from the peripheral regions to the active site, Equation (2) incorporates the effects covered by the FSA described above. At first glance, the relationship between the CSA and our LRF analysis seems to be inside-out. In fact, from the viewpoint of the response theory, a deletion of the active site residues is equivalent to a perturbation that has occurred at the active site, and Equation (1),  $\Delta q_{Res} = q_{Res}^{holo} - q_{Res}^{apo}$ , is the density change at peripheral residues, which can be expressed by the following equation:

$$\Delta q_{Res} \cong \frac{\delta\rho(\text{peripheral residues})}{\delta v(\text{active site residues})} \cong \frac{\delta\rho(\mathbf{r}')}{\delta v(\mathbf{r})}. \quad (3)$$

However, since the linear response function is symmetric with respect to  $\mathbf{r}$  and  $\mathbf{r}'$ , i.e.,

$$\frac{\delta\rho(\mathbf{r}')}{\delta v(\mathbf{r})} = \frac{\delta\rho(\mathbf{r})}{\delta v(\mathbf{r}')}, \quad (4)$$

the CSA can be considered as a special version of the LRF analysis. We previously applied our LRF analysis to the fundamental covalent bonding systems,  $\pi$  conjugated systems, and polypeptide systems [18–20]. Our approach relies on the fact that if we consider  $\delta\rho(\mathbf{r})/\delta v(\mathbf{r}')$  as a propagation of

the effect from the perturbation  $\delta v$  at  $r'$  to the response,  $\delta\rho$ , at  $r$ , the decay of propagation will vary depending on the types of the mediated chemical bonds, such as a covalent-bond,  $\pi$  conjugation, hyper-conjugation, hydrogen-bonding, d- $\pi$  interactions, and through-space interactions (the examples will be shown in the results section). The reason we use the “density” changes is that density is the fundamental property in the context of the density functional theory (DFT), which determines all other electronic properties of molecular systems [25–27].

In this study, we extended our LRF analysis to spin-polarized systems. As is well known, the transition metal complexes, containing open-shell transition metal ions, play important roles in the catalytic activity of enzymes, and electric and magnetic properties of the molecular materials [28–30]. For such cases, the most stable solutions for the Kohn-Sham DFT equation [26] became spin-symmetry broken solutions, where an additional property, such as spin-density [31], on-top pair density [32], and radical density [33], is needed to identify the electronic structures, because the radical character [34] of transition metal complexes is deeply related to the functionality of the active center of such enzymes and molecular materials. These various properties describe the radical character and are related to each other via known relations [35]. The spin-density is another fundamental variable of spin-unrestricted KS-DFT, and we employed the spin-density for the LRF analysis of transition metal complexes described below. We applied our LRF analysis of both the density and the spin-density to the reaction center of P450, which is a mono-nuclear Fe(III) porphyrin complex, and the reaction center of photosystem II,  $Mn_4O_5Ca$  complex. The results are discussed in relation to the QM cluster and QM/MM modeling that has been implemented so far, in particular from the viewpoint of chemical-bonding and intermolecular interactions.

## 2. Theoretical Background

### 2.1. Linear Response Function of Density and Spin Density

As described in the Introduction section, we have investigated the linear response function of the density, because it determines all other electronic properties uniquely in DFT [25].

In the context of QM/MM modeling, the perturbations  $\delta v(r')$ , correspond to the errors, due to the replacement of the QM electronic structures of peripheral and environmental parts ( $r'$ ) of large molecular systems by MM point charges. The density changes,  $\delta\rho(r)$ , correspond to the errors in the QM region ( $r$ ). Then we assume that the following relation holds,

$$\delta\rho(r) = \int dr' \frac{\delta\rho(r)}{\delta v(r')} \delta v(r'). \quad (5)$$

We would like to emphasize again that  $\delta v(r')$  depends on the approximation of the QM/MM modeling for the boundary region, such as the link atom method or the localized orbital method. We do not focus our attention on  $\delta v(r')$ , although many QM/MM developers intend to improve  $\delta v(r')$  (reference [5] and references therein). Instead, our purpose is to inspect  $\delta\rho(r)/\delta v(r')$ , which is related to how the errors due to the QM/MM model propagate through the target molecule. If we employed a specific approximation (for instance, the link-H atom method), then we can assume that the error due to the approximation,  $\delta v(r')$ , takes a specific value. Then we can regulate the error of the density at the QM event site,  $\delta\rho(r)$ , by putting the QM/MM boundary ( $r'$ ) away from the QM event site ( $r$ ), on the basis of the  $\delta\rho(r)/\delta v(r')$  value: As the value  $\delta\rho(r)/\delta v(r')$  decreases, the error of the corresponding QM/MM setting decreases. Another noteworthy point is that, although we assume that the calculation of the whole system can be performed, as shown in the examples in this study,  $\delta\rho(r)/\delta v(r')$  can also be useful for estimating the effects on the reaction center  $r$  on extending the system at a peripheral point  $r'$ . Although such calculations are not done below, such type of studies are interesting in relation to the substituent effects, which are often studied in the field of conceptual DFT [23,24]. We also note that density changes are the main cause of errors in dynamic behavior of the systems when we perform the molecular dynamics, based on the QM calculations, because the dynamics is determined by forces

acting on the nuclei. According to Feynman's electrostatic theorem [36], the forces depend only on the positions of nuclei and electron density. In addition, the validity of the QM/MM modeling relies on the near-sightedness of electronic matter (NEM) as proposed by Kohn and Prodan [27], which states that the perturbations,  $\delta v$ , at any points that are far from a specific point, do not change electronic density,  $\delta\rho$ , at the point significantly. Thus, our LRF analysis for the QM/MM modeling can be considered a test of whether, and how, NEM holds for molecular systems.

In our analysis, we estimated the LRF by applying the simple perturbation theory to the Kohn-Sham DFT solution as,

$$\frac{\delta\rho(\mathbf{r})}{\delta v(\mathbf{r}')} = \sum_{\sigma}^{\alpha,\beta} \frac{\delta\rho^{\sigma}(\mathbf{r})}{\delta v(\mathbf{r}')} = \sum_{\sigma}^{\alpha,\beta} \sum_i^{N_{occ}^{\sigma}} \sum_j^{N_{unocc}^{\sigma}} \frac{\psi_{i\sigma}(\mathbf{r})\psi_{j\sigma}(\mathbf{r})\psi_{i\sigma}(\mathbf{r}')\psi_{j\sigma}(\mathbf{r}')}{\varepsilon_{i\sigma} - \varepsilon_{j\sigma}}. \quad (6)$$

Here,  $\psi_{i\sigma}(\mathbf{r})$  and  $\psi_{j\sigma}(\mathbf{r})$  are occupied and unoccupied Kohn-Sham orbitals with spin  $\sigma$ . The first, second, and third summations run over the spin variables occupied orbitals, and all unoccupied orbitals, respectively. Readers should note that the perturbation,  $\delta v(\mathbf{r}')$  is not specified and is virtual. In other words, we do not need any actual perturbations, such as point charges or electric fields to estimate  $\delta\rho(\mathbf{r})/\delta v(\mathbf{r}')$ . All we need are Kohn-Sham orbitals and their orbital energies. The use of the LRF is not common, but has been investigated in the field of conceptual DFT. In [22], the authors related the LRF to softness, ( $S$ ), local softness ( $s(\mathbf{r})$ ), and the kernel of local softness ( $s(\mathbf{r}, \mathbf{r}')$ ) as,  $\delta\rho(\mathbf{r})/\delta v(\mathbf{r}') = -s(\mathbf{r}, \mathbf{r}') + s(\mathbf{r})s(\mathbf{r}')/S$  and discussed the hard and soft acid and bases (HSAB) principle on the basis of DFT. Geerlings and his coworkers showed that the LRF can be a good descriptor for the inspection of substitution effects, aromaticity, and in the chemical reactivity of molecules [23,24]. We have applied the LRF in order to determine guidelines for the appropriate QM/MM modeling, not to discuss the specific "chemistry" of the target molecules. We intend to make full use of the fact that the LRF values,  $\delta\rho(\mathbf{r})/\delta v(\mathbf{r}')$ , strongly depends on the numbers and types of chemical bonds lied between  $\mathbf{r}$  and  $\mathbf{r}'$ . For the current purpose, it would be useful to employ the condensed versions of the LRF of density, i.e., the density change at  $\mathbf{r}$ , due to the perturbation that is applied to the  $J$ -th atom,  $\delta\rho(\mathbf{r})/\delta v(J)$  and the density change at the  $I$ -th atom for the perturbation to the  $J$ -th atom,  $\delta\rho(I)/\delta v(J)$ . We describe the computational details of these condensed versions of the LRF of density in the Materials and Methods section below.

The extension of the LRF to the spin density version is straightforward as follows. If we note that the spin density,  $\rho_z(\mathbf{r})$ , is the difference between  $\alpha$  spin density,  $\rho^{\alpha}(\mathbf{r})$ , and  $\beta$  spin density,  $\rho^{\beta}(\mathbf{r})$ , and that the primary perturbation to the spin density is the virtual magnetic field,  $v_m$ , we have

$$\frac{\delta\rho_z(\mathbf{r})}{\delta v_m(\mathbf{r}')} = \frac{\delta\rho^{\alpha}(\mathbf{r})}{\delta v_m(\mathbf{r}')} - \frac{\delta\rho^{\beta}(\mathbf{r})}{\delta v_m(\mathbf{r}')} \quad (7)$$

where the LRF of the  $\sigma$  spin density is given by

$$\frac{\delta\rho^{\sigma}(\mathbf{r})}{\delta v_m(\mathbf{r}')} = \sum_i^{N_{occ}^{\sigma}} \sum_j^{N_{unocc}^{\sigma}} \frac{\psi_{i\sigma}(\mathbf{r})\psi_{j\sigma}(\mathbf{r})\psi_{i\sigma}(\mathbf{r}')\psi_{j\sigma}(\mathbf{r}')}{\varepsilon_{i\sigma} - \varepsilon_{j\sigma}}. \quad (8)$$

This formulation is not new. For instance, Fias et al. formulated the LRF of spin density in terms of  $\delta\rho^{\sigma}(\mathbf{r})/\delta v^{\sigma'}(\mathbf{r}')$  for all the combination,  $(\sigma, \sigma') = (\alpha, \alpha), (\alpha, \beta), (\beta, \alpha), (\beta, \beta)$  and inspected the LRF of spin density of atoms from the viewpoint of shell-structure and electron-electron repulsion in atoms [37]. Nevertheless, the chemistry, based on the LRF of spin density, has not been investigated thoroughly so far, even in the field of the conceptual DFT. One of the reasons is that many researchers consider the spin-density in Kohn-Sham (KS) DFT as an artifact of the theory: In fact, the KS-DFT solutions, exhibiting a singlet bi-radical do not have spin symmetry, although the spin symmetry condition must be satisfied for the non-relativistic many-electrons systems. Still, we would like to emphasize that the correct description of spin density is essential for understanding electronic

structures of transition metal complexes, because it is closely related to the effective bond-order and the radical character of the solutions. In the next section, we will describe the relation between the spin density and the effective bond-order, that can be derived from the KS DFT solutions, and an interpretation of spin density as an indicator of electron-electron correlation effects, that are important for computing transition metal complexes containing open-shell transition metal ions.

## 2.2. Spin Density as an Indicator of Correlation Effects in Kohn-Sham DFT

Closed shell molecules can be described by pairs of electrons in the KS orbitals that share the same spatial distribution. However, the closed-shell picture often breaks down in the case of transition metal complexes, containing open-shell transition metal ions, yielding more stable broken-symmetry (BS) solutions of spin-unrestricted KS-DFT (UKS-DFT) [38,39]. Typically, the BS solutions of UKS-DFT are expressed with the bonding and anti-bonding closed-shell orbitals as,

$$\psi_i^\alpha(\mathbf{r}) = \cos\theta_i \phi_i^{\text{bond}}(\mathbf{r}) + \sin\theta_i \phi_i^{\text{antibond}}(\mathbf{r}), \quad (9a)$$

$$\psi_i^\beta(\mathbf{r}) = \cos\theta_i \phi_i^{\text{bond}}(\mathbf{r}) - \sin\theta_i \phi_i^{\text{antibond}}(\mathbf{r}). \quad (9b)$$

Note that the antibonding orbitals are mixed with the bonding orbitals in BS solutions implying that the BS solutions suffer the instability of chemical bonds of the open-shell species [34]. In fact, the first-order density matrix of the BS solutions given by Equations (9a) and (9b) can be diagonalized in the form,

$$\begin{aligned} \rho(\mathbf{r}, \mathbf{r}') &= \sum_i \left\{ (1 + \cos 2\theta_i) \phi_i^{\text{bond}}(\mathbf{r}) \phi_i^{\text{bond}}(\mathbf{r}') + (1 - \cos 2\theta_i) \phi_i^{\text{antibond}}(\mathbf{r}) \phi_i^{\text{antibond}}(\mathbf{r}') \right\} \\ &= \sum_i \left( n_i^{\text{bond}} \phi_i^{\text{bond}}(\mathbf{r}) \phi_i^{\text{bond}}(\mathbf{r}') + n_i^{\text{antibond}} \phi_i^{\text{antibond}}(\mathbf{r}) \phi_i^{\text{antibond}}(\mathbf{r}') \right) \end{aligned} \quad (10)$$

Thus the effective bond-order ( $b$ ) can be expressed as

$$b \equiv \sum_i b_i = \sum_i \frac{n_i^{\text{bond}} - n_i^{\text{antibond}}}{2} = \sum_i \cos 2\theta_i = \sum_i T_i \quad (11)$$

Here,  $T_i$  is defined as the overlap integral between  $\psi_i^\alpha(\mathbf{r})$  and  $\psi_i^\beta(\mathbf{r})$ ,

$$T_i \equiv \langle \psi_i^\alpha | \psi_i^\beta \rangle. \quad (12)$$

We should note that  $T_i$  takes a value from 1 (the closed-shell solutions) to 0 (the completely localized singlet biradical), and so do  $b_i$ . Using Equations (9a) and (9b), together with Equation (12), the spin density can be expressed by

$$\rho_z(\mathbf{r}) = 2 \sum_i \sqrt{1 - T_i^2} \phi_i^{\text{bond}}(\mathbf{r}) \phi_i^{\text{antibond}}(\mathbf{r}) \quad (13)$$

If we consider a two-electron two-site model, bonding and antibonding orbitals are given by

$$\phi^{\text{bond}}(\mathbf{r}) = \frac{1}{\sqrt{2}} (\chi_1(\mathbf{r}) + \chi_2(\mathbf{r})) \quad (14a)$$

$$\phi^{\text{antibond}}(\mathbf{r}) = \frac{1}{\sqrt{2}} (\chi_1(\mathbf{r}) - \chi_2(\mathbf{r})) \quad (14b)$$

and the spin density reduces to

$$\rho_z(\mathbf{r}) = \sqrt{1 - b^2} \chi_1(\mathbf{r}) \chi_1(\mathbf{r}) - \sqrt{1 - b^2} \chi_2(\mathbf{r}) \chi_2(\mathbf{r}). \quad (15)$$

The first and the second terms on the right side are the spin densities on the 1st, and the 2nd sites, respectively. Equation (15) implies that the spin density includes the information concerning the effective bond-order for radical systems. A simple spin-projection method to recover the spin-symmetry of the BS KS-DFT solutions has also been derived [34,38,39]. In this case, the effective bond-order of the spin-projected solution is rewritten in the form,  $B = \sum_i 2T_i / (1 + T_i^2)$ . As an index for the radical character of the spin-projected solutions, the di-radical character is defined as  $y \equiv 1 - B$  [34,35]. Recently, the di-radical character was used to design molecular materials exhibiting hyper-polarizability [40]. However, such properties based on the spin-projection procedure are beyond the scope of this paper. Instead, we note the implication of spin density of the BS-KS-DFT solutions, as pointed out by Perdew et al.: spin density is an alternative variable of the on-top pair density, which is an indicator of the correct description of correlation effects between  $\alpha$  and  $\beta$  spins for magnetic systems [32], via the following relation,

$$\rho_2^{\alpha\beta}(\mathbf{r}; \mathbf{r}') \Big|_{\mathbf{r}=\mathbf{r}'} = \frac{1}{2}(\rho(\mathbf{r})\rho(\mathbf{r}) - \rho_z(\mathbf{r})\rho_z(\mathbf{r})) \quad (16)$$

This property indicates the degree of correlation between  $\alpha$  and  $\beta$  electrons. Whether the spin-projection is applied to the BS KS-DFT solution or not, it is needed to describe the spin-polarized state of transition metal complexes appropriately. This is the reason why we focus on spin density. Of course, for closed shell systems, the BS KS-DFT solutions reduce to the spin restricted KS-DFT solutions. Thus the spin density, di-radical character, and the on-top pair density reduce to zero. Of course,  $\delta\rho_z(\mathbf{r})/\delta v_m(\mathbf{r}')$  also vanishes. In the next section, we will present the computational results of the reaction centers of metallo-protein that contain open-shell transition metal ions. We will focus on the use of the LRF of density and spin density to determine the QM/MM modeling to obtain the appropriate KS-DFT solutions.

### 3. Results and Discussion

We will present the computational results for two types of transition-metal complexes. For all computations of LRFs, we used a modified version of GAMESS [41]. The B3LYP functional is employed for all DFT calculations [42]. Basis sets for each transition metal complexes and other computational details are described in Materials and Methods section.

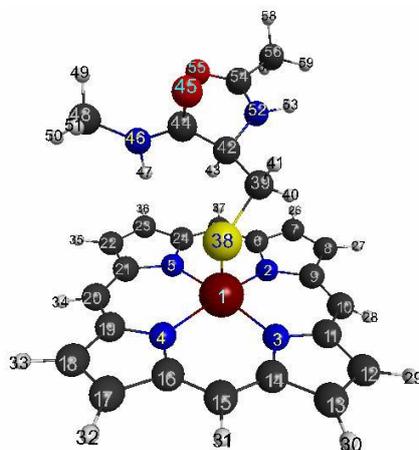
#### 3.1. Heme System~Reaction Center of P450

##### 3.1.1. The Target System

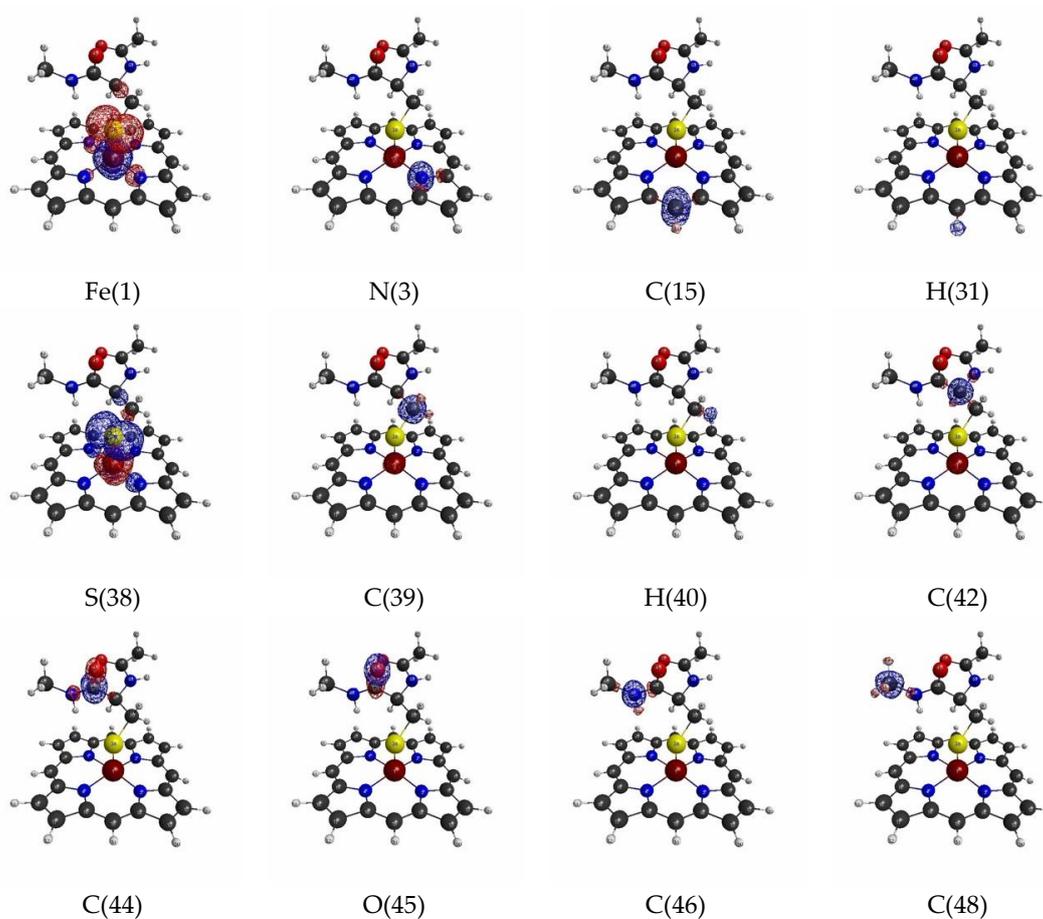
Metalloporphyrins play various roles, such as transportation of  $O_2$ , catalase, monooxidase, and electron transfer in living bodies [43]. Thus, as a typical mono nuclear transition metal complex, we chose metalloporphyrin, that is extracted from the active center of P450 (pdb entry:5vws) [44], of which the geometry is shown in Figure 1. The same, but a larger picture is shown in Figure S1. The numbers, indicated on the atoms in the figures, are used in the following visualizations of LRFs and the discussion. To confirm the numbering of the atoms, please refer to Figure S1.

The isosurfaces of the LRFs for the threshold setting with  $\delta\rho(\mathbf{r})/\delta v(\mathbf{I}) = \pm 0.1$  to the perturbations that are applied to the atomic sites are shown in Figure 2. In the context of QM/MM modeling, roughly speaking, the region where the responses appear in this figure should be included in the QM region for this threshold. As we can see from Figure 2, almost all responses are localized near the perturbation site. However, when the perturbation is applied to the Fe ion (Fe(1)), the density changes are obviously spread over the S ion of the thiol group, and vice versa. This implies that the Fe(III) and S(-I) are strongly coupled to each other. In fact, it is known that the part of residues ligated from the above and/or below sides regulate the functionality of the Heme core, and the Cys of P450 is a typical example [43]. Interestingly, although  $\delta\rho(\mathbf{r})/\delta v(\text{Fe}(1))$  is also delocalized over the porphyrin ring (Por), the coupling between Fe(III) and S ions seems to be stronger than that between Fe(III) and

the porphyrin ring from the viewpoint of LRF of electron density. Also, we note that the perturbation to a site in a specific amide plane yields density changes over the atoms in the same plane (for instance, C(42), C(44), O(45)). In other words, the LRF with this threshold detects the firm units, where the sites are coupled with strong covalent and/or ionic bonds.



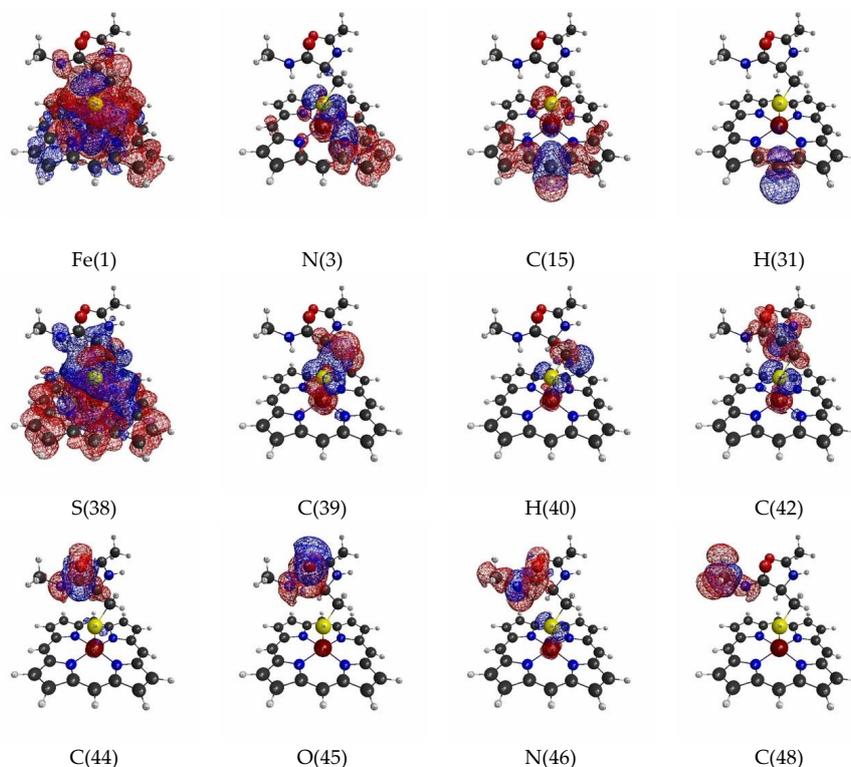
**Figure 1.** Geometry of the active site of P450 (pdb entry:5vws). The numbering of the atoms is used for the following figures to show linear response functions.



**Figure 2.** Isosurfaces of linear response functions of density for the perturbations on atomic sites of P450 (pdb entry:5vws). The threshold of  $\delta\rho(r)/\delta v(I)$  is  $\pm 0.01$  (+: blue, -: red). The atomic symbols and those numbers in parentheses below the figures indicate the atoms (I) to which the virtual perturbations are applied.

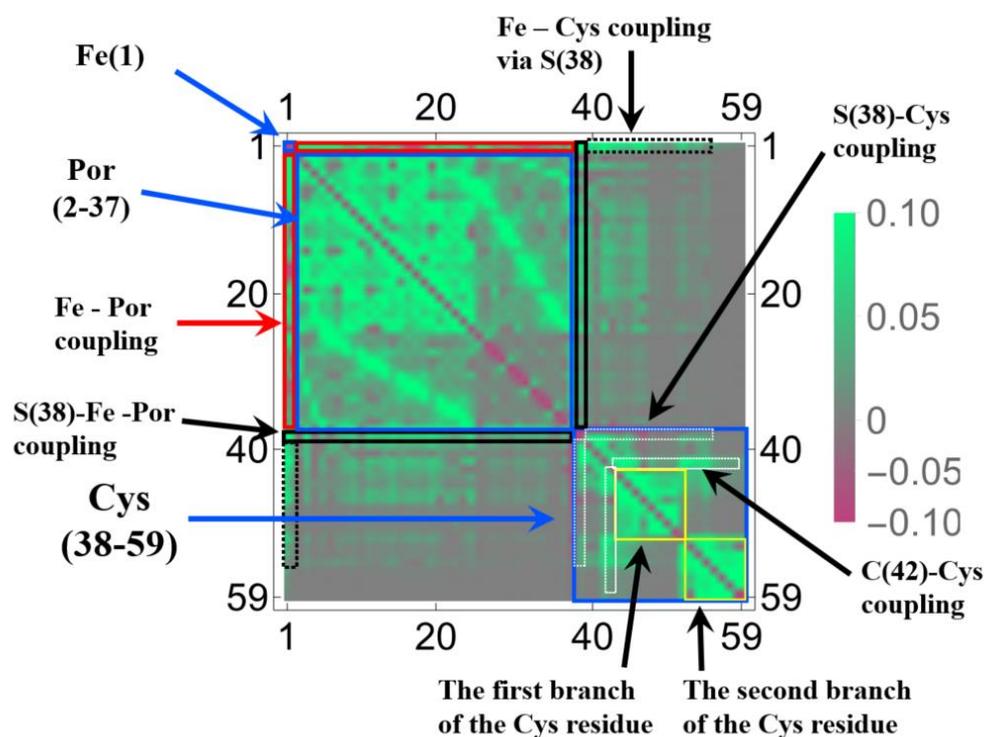
### 3.1.2. Linear Response Functions of Electron Density of P450

Next, Figure 3 shows the isosurfaces of the LRFs with the threshold  $\delta\rho(\mathbf{r})/\delta v(I) = \pm 0.01$ . As suggested in the Theoretical Background section, as the value of the threshold decreases, a large QM region is required to satisfy the threshold of the errors on density. In this “tighter” threshold, the responses naturally become more delocalized than those for  $\delta\rho(\mathbf{r})/\delta v(I) = \pm 0.1$ . The result for when the perturbation is applied to Fe(III) shown in Figure 2 ( $\delta\rho(\mathbf{r})/\delta v(\text{Fe}(1))$ ) is remarkable. Because the Fe(III) and S(-I) ions are coupled with strong ionic bonds, the responses are spread over the heavy atoms in the porphyrin ring, together with the residues of amino acids up to the N(46) atom. Another point of interest is that the propagation of the effects are spread over through the two  $sp^3$  junctions (C(39) and C(42) atoms). This must be via the hyperconjugation path consisting of S(38), C(39), C(42), and C(44) atoms. The interaction between the plus charge at H(47) in the NH and the minus charge at S(38) must also facilitate this delocalization. In fact, the  $\delta\rho(\mathbf{r})/\delta v(\text{N}(46))$  is also spread to the S(-I) site. As pointed out above, because the Fe(III) and S(-I) are strongly coupled with each other, the extensive delocalization of the density fluctuations is also observed in  $\delta\rho(\mathbf{r})/\delta v(\text{S}(38)) = \pm 0.01$ . Compared to the case of the loose threshold ( $\delta\rho(\mathbf{r})/\delta v(I) = \pm 0.1$ ), there are other differences. First, the considerably strong coupling between Fe(III) and the heavy atoms in the porphyrin ring is observed for  $\delta\rho(\mathbf{r})/\delta v(\text{C}(15))$ . Second, the delocalization effects over hyperconjugation through the  $sp^3$  junctions at the C carbons are also observed ( $\delta\rho(\mathbf{r})/\delta v(\text{C}(39))$ ,  $\delta\rho(\mathbf{r})/\delta v(\text{C}(40))$ ,  $\delta\rho(\mathbf{r})/\delta v(\text{C}(42))$ ). In summary, the LRF with this threshold detects the hyperconjugation and weak bonds in the whole molecule, by which the electronic structure of the whole molecule must be regulated. Thus, it is strongly recommended that the coupled units presented in our computational results of  $\delta\rho(\mathbf{r})/\delta v(I) = \pm 0.01$  are not be divided or partially replaced by capped H atoms. For instance, from  $\delta\rho(\mathbf{r})/\delta v(\text{S}(38)) = \pm 0.01$  shown in Figure 3, the replacement of the terminal C(48)H<sub>3</sub> by H is acceptable, but the replacement of the C(42)N(52)H(43)C(44) by CH<sub>3</sub> should be avoided.

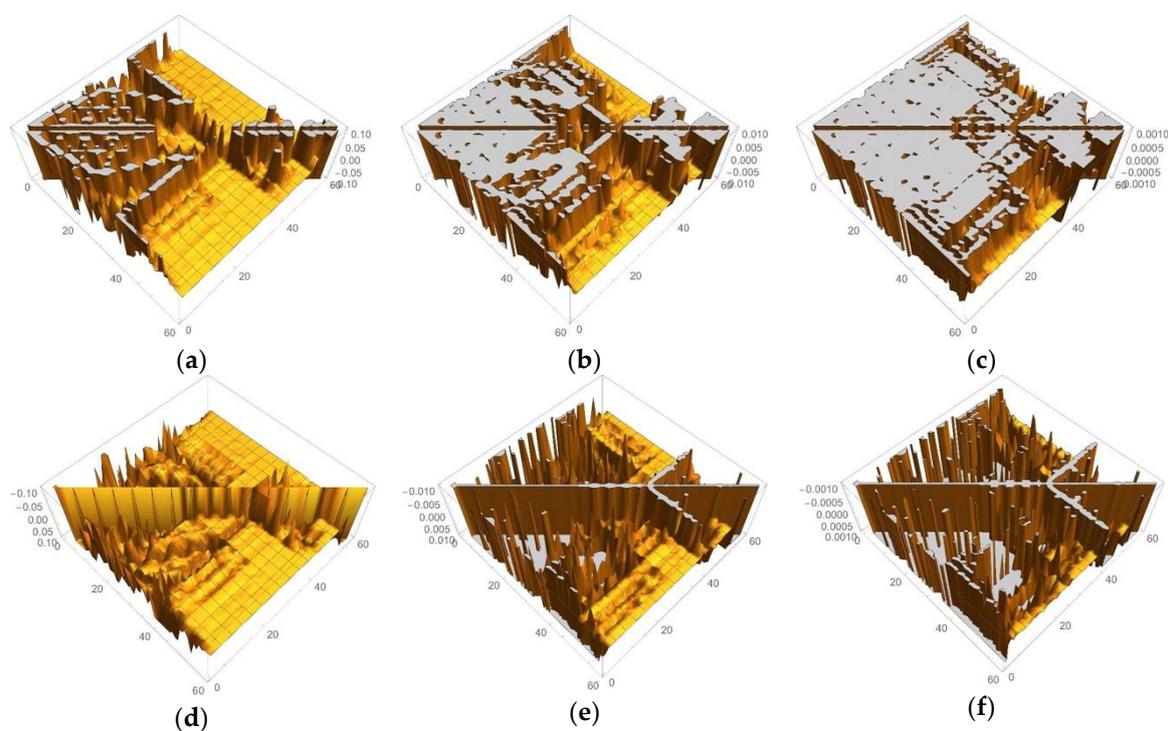


**Figure 3.** Isosurfaces of linear response functions of density for the perturbations on atomic sites of P450 (pdb entry:5vws). The threshold of  $\delta\rho(\mathbf{r})/\delta v(I)$  is  $\pm 0.01$  (+: blue, -: red). The atomic symbols and those numbers in parentheses below the figures indicate the atoms (I) to which the virtual perturbations are applied.

In Figures 1 and 2, we selected the important isosurfaces of LRFs to reduce the numbers of the figures. For more comprehensive data (but with almost equivalent figures removed), please see Figures S2 and S3. The tighter the threshold, the more delocalized isosurfaces are observed. We also plotted  $\delta\rho(\mathbf{r})/\delta v(I) = \pm 0.001$  in Figure S4. For convenience, the more compact representation of Figures S2–S4 is shown in Figure 4 as a heat-map representation of the matrix,  $\{\delta\rho(I)/\delta v(J)\}_{I,J=1\sim 59}$ . It is obvious from this figure that there are three units, the Fe ion, the porphyrin ring (Por), and the His residue as the diagonal blocks, which are indicated by blue squares. The strong coupling between the Fe ion and the Por ring, which is described above, and the coupling between the Por ring and the His residue are indicated in black rectangles. A noteworthy point is that the perturbation to Fe(1) is spread not only to Por, which is indicated by the red square, but even into His via S(38), which is indicated by the dotted black rectangles. These dotted black rectangles correspond to isosurfaces of  $\delta\rho(\mathbf{r})/\delta v(\text{Fe}(1)) = \pm 0.01$  and  $\delta\rho(\mathbf{r})/\delta v(\text{S}(38)) = \pm 0.01$  shown in Figure 2. One of the reasons for such a delocalized character of  $\delta\rho(\mathbf{r})/\delta v(\text{Fe}(1))$  can be explained as follows. In the diagonal block that corresponds to the Cys residue, we can see sub-structures indicated by yellow squares, which correspond to two branches of the Cys residue (C(44)-H(50)) and (C(52)-H(59)). There are two major types of off-diagonal blocks in this diagonal block of the Cys residue: One is between S(38) and these two branches, and another is between the root of the two branches, i.e., C(42), and the two branches, both of which are shown as dotted white rectangles. Obviously, these paths mediate the propagation of the effects, due to the perturbation from the Fe(1)-S(38) unit to the two branches of the Cys residue. In order to discuss the relationship between the structures of the LRF and the threshold for the LRF values more clearly, we plotted the sections for  $\delta\rho(I)/\delta v(J) = 0.1$ ,  $\delta\rho(I)/\delta v(J) = 0.01$ , and  $\delta\rho(I)/\delta v(J) = 0.001$  in Figure 5a–c, respectively. For completeness, we also plotted  $\delta\rho(I)/\delta v(J) = -0.1$ ,  $\delta\rho(I)/\delta v(J) = -0.01$ , and  $\delta\rho(I)/\delta v(J) = -0.001$  in Figure 5d–f, respectively. Because the minus value of  $\delta\rho(I)/\delta v(J)$  indicates an increase in electron density,  $\delta\rho(I)$ , due to the attractive perturbation,  $\delta v(J)$ , the section for  $\delta\rho(I)/\delta v(J) = -0.1$  is localized at the diagonal part as shown in Figure 5d. Thus, the delocalization of the LRFs is more clearly seen from Figure 5a–c than from Figure 5d–f. We can see from Figure 5a that if we employ the criterion,  $|\delta\rho(I)/\delta v(J)| < 0.1$ , this system can be divided into several parts. However, because the Fe(III) and S(-I) ions are strongly coupled, the Fe(III) and Por cannot be separated from the Cys residue as shown in Figure 2 ( $\delta\rho(\mathbf{r})/\delta v(\text{Fe}(1)) = \pm 0.1$  and  $\delta\rho(\mathbf{r})/\delta v(\text{S}(38)) = \pm 0.1$ ). This characteristic of the P450 active site is more explicitly shown in Figure 5b, which implies that the terminal  $\text{CH}_3$  beyond C(48) and the terminal  $\text{CH}_3\text{CO}$  beyond N(52) can be replaced by capped H atoms. When the criterion  $|\delta\rho(I)/\delta v(J)| < 0.001$  must be imposed, the whole system should be included in one QM region, which can be seen from the spread of the section shown in Figure 5c. Roughly speaking, however, this criterion means that even when the QM/MM modeling error  $\delta v(J)$  is 1 Hartree, the error on the active site  $\delta\rho(I)$  is less than 0.001, which is too severe criterion. The approximation to exchange-correlation functional of DFT might lead to larger errors than the QM/MM modeling with the criterion ( $|\delta\rho(I)/\delta v(J)| < 0.001$ ). In fact, almost all QM/MM researchers include up to the second-next residues from the reaction centers (Ref. [5] and references therein), being consistent with the criterion,  $|\delta\rho(I)/\delta v(J)| < 0.01$ .



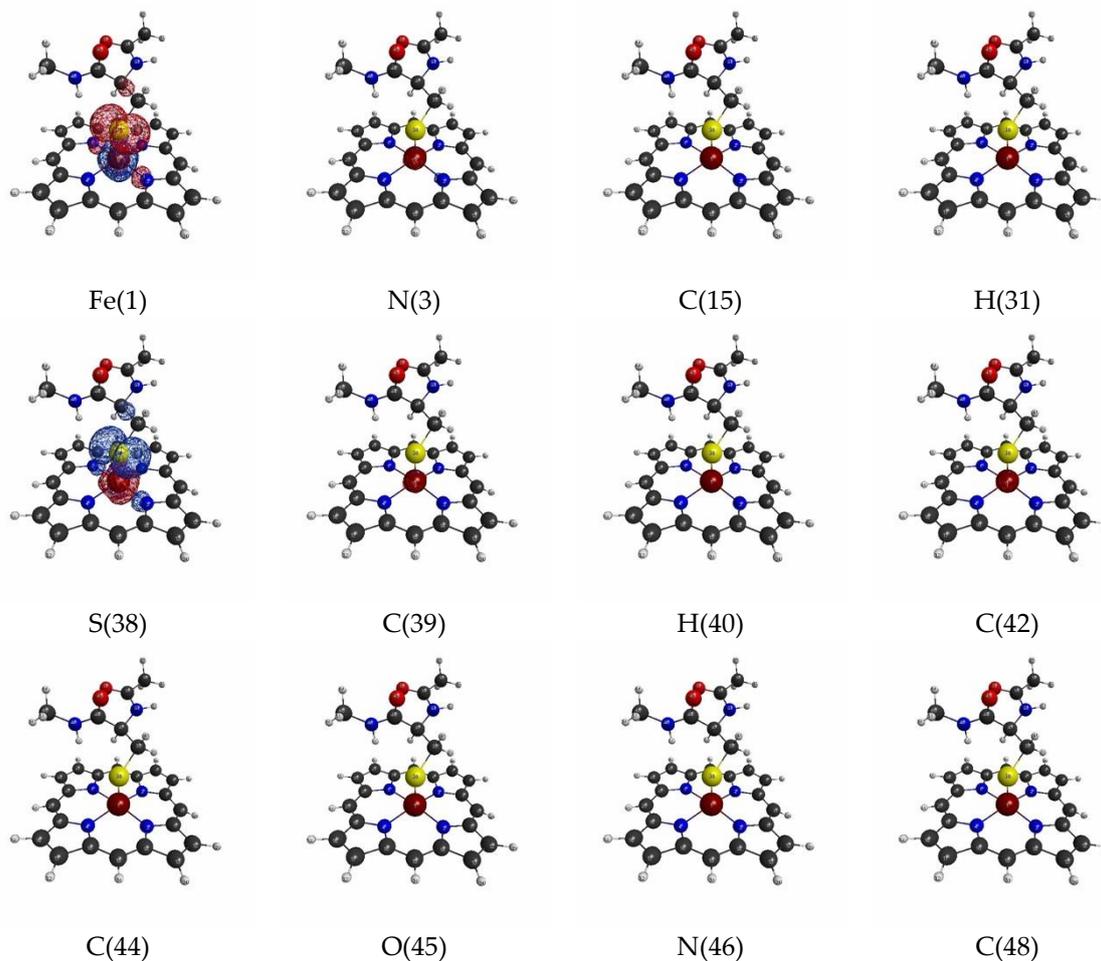
**Figure 4.** The heat map for the matrix representation of condensed version of linear response functions of electron density,  $\{\delta\rho(I)/\delta v(J)\}_{I,J=1\sim 59}$ , for P450. The number of the axes corresponds to the numbering of the atoms indicated in Figure 1. The functional units and the couplings among units are also indicated (see text for the details).



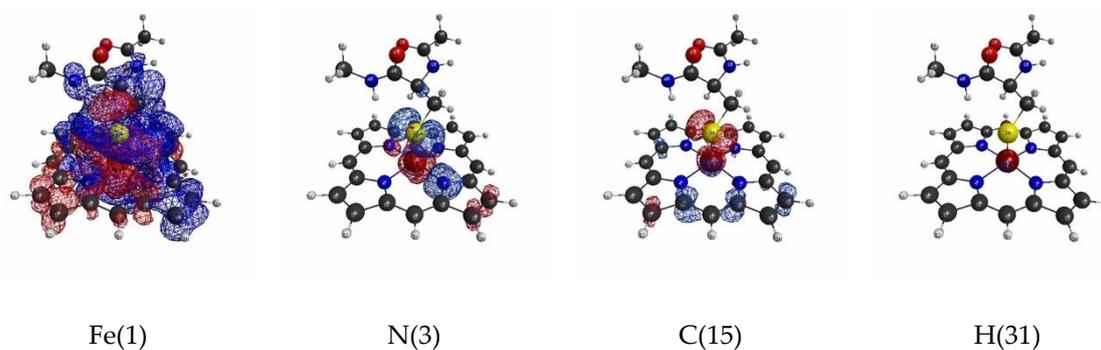
**Figure 5.** Matrix representation of condensed version of linear response functions of electron density for P450 with the six cutting thresholds. (a)  $\delta\rho(I)/\delta v(J) < 0.1$ . (b)  $\delta\rho(I)/\delta v(J) < 0.01$ . (c)  $\delta\rho(I)/\delta v(J) < 0.001$ , (d)  $\delta\rho(I)/\delta v(J) > -0.1$ . (e)  $\delta\rho(I)/\delta v(J) > -0.01$ . (f)  $\delta\rho(I)/\delta v(J) > -0.001$ .

### 3.1.3. Linear Response Functions of Spin Density of P450

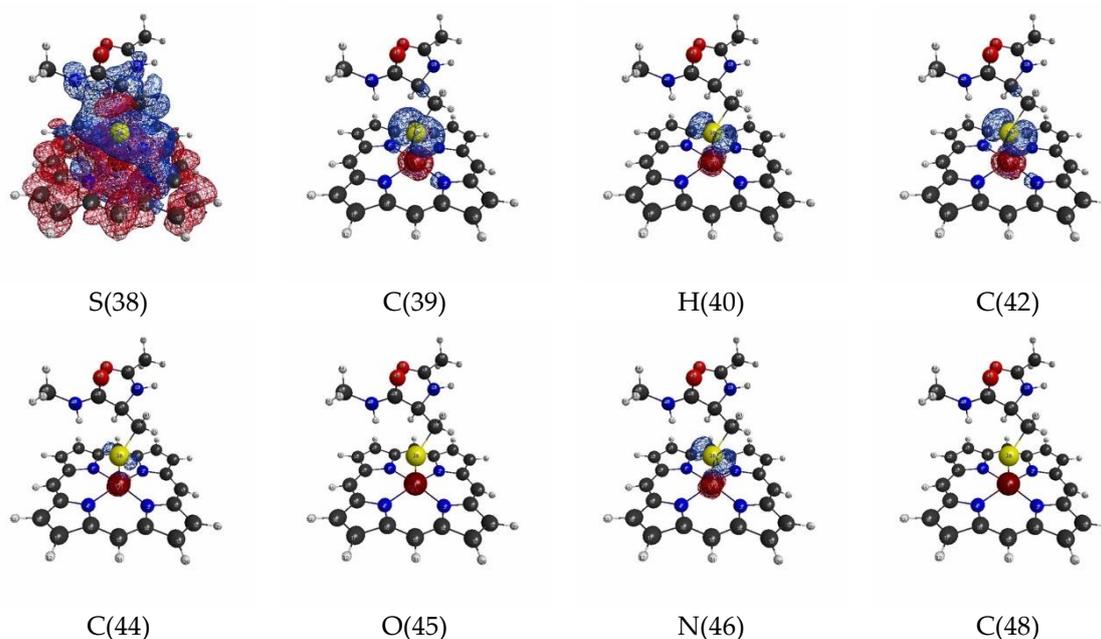
Next, we plotted isosurfaces of LRFs of spin density with  $\delta\rho_z(\mathbf{r})/\delta v_m = \pm 0.1$  in Figure 6. As we can see, except  $\delta\rho_z(\mathbf{r})/\delta v_m(\text{Fe}(1))$  and  $\delta\rho_z(\mathbf{r})/\delta v_m(\text{S}(38))$ , the responses are not remarkable. Still,  $\delta\rho_z(\mathbf{r})/\delta v_m(\text{Fe}(1))$  and  $\delta\rho_z(\mathbf{r})/\delta v_m(\text{S}(38))$  are delocalized into the porphyrin ring and C(42) of the Cys residue. The sensitivity of the sites, Fe(1) and S(38), becomes more obvious for the threshold  $\delta\rho_z(\mathbf{r})/\delta v_m(I) = \pm 0.01$  as shown in Figure 7.



**Figure 6.** Isosurfaces of linear response functions of spin density for the perturbations on atomic sites of P450. The threshold of  $\delta\rho_z(\mathbf{r})/\delta v_m(I)$  is  $\pm 0.1$  (+: blue, -: red). The atomic symbols and those numbers in parentheses below the figures indicate the atoms (I) to which the virtual perturbations are applied.



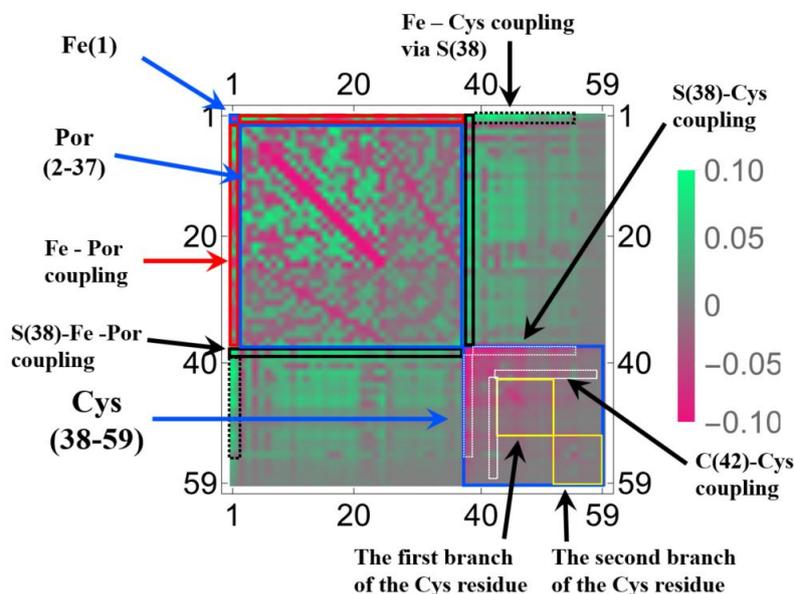
**Figure 7.** Cont.



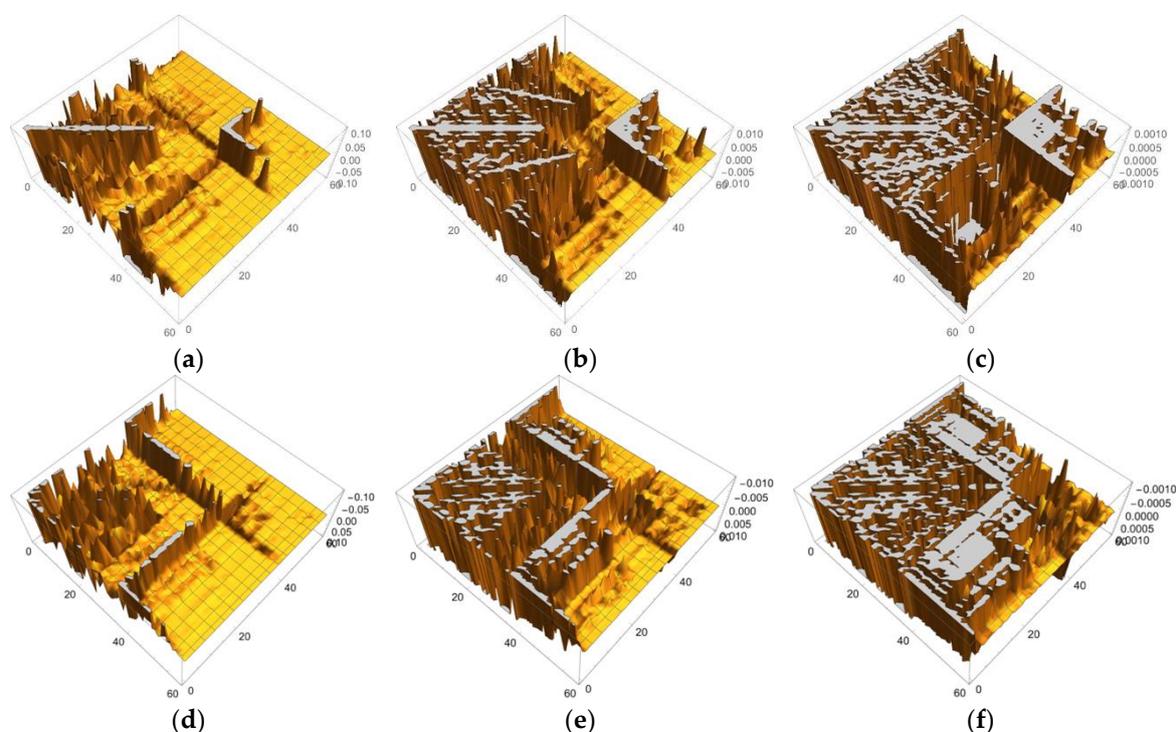
**Figure 7.** Isosurfaces of linear response functions of spin density for the perturbations on atomic sites of P450). The threshold of  $\delta\rho_z(\mathbf{r})/\delta v_m(I)$  is  $\pm 0.01$  (+: blue, -: red). The atomic symbols and those numbers in parentheses below the figures indicate the atoms (I) to which the virtual perturbations are applied.

For reference, Figures S5–S7 plot the comprehensive data of the isosurfaces for  $\delta\rho_z(\mathbf{r})/\delta v_m(I) = \pm 0.1$ ,  $\delta\rho_z(\mathbf{r})/\delta v_m(I) = \pm 0.01$ , and  $\delta\rho_z(\mathbf{r})/\delta v_m(I) = \pm 0.001$ , respectively. The results for  $\delta\rho_z(\mathbf{r})/\delta v_m(I) = \pm 0.001$  shown in Figure S7 are similar to but more compact than those of  $\delta\rho(\mathbf{r})/\delta v(I)$  shown in Figure S4. It should be noted that  $\delta\rho_z(\mathbf{r})/\delta v_m(\text{H}(47))$  and  $\delta\rho_z(\mathbf{r})/\delta v_m(\text{N}(46))$  shown in Figure 7 and  $\delta\rho_z(\mathbf{r})/\delta v_m(\text{N}(52))$  shown in Figure S6 are spread over Fe and S. On the other hand, the perturbations that are applied to terminal  $\text{CH}_3$  beyond C(48) and the terminal  $\text{CH}_3\text{CO}$  beyond N(52) do not significantly affect Fe or S: This is similar to the linear response function (LRF) of electron density. On the basis of these results, the same conclusion can be drawn from the viewpoint of the use of both LRFs of density and spin density for QM/MM modeling of this system: the  $\text{CH}_3$  beyond C(48) and the terminal  $\text{CH}_3\text{CO}$  beyond N(52) can be approximated as MM charges. However, the overview of matrix representation of  $\{\delta\rho_z(I)/\delta v_m(J)\}_{I,J=1\sim 59}$ , which can be seen from Figure 8, is considerably different from the corresponding figure of the LRF of density shown in Figure 4. The first remarkable point is that the disordered checker pattern can be seen in the sub-block of the porphyrin, which indicates the alternative spin-polarization effects on the porphyrin ring. This can be also confirmed from sections shown in Figure 9a–f. For instance, the sections of Figure 9b,e compensate each other in the region of the porphyrin ring. This is the result of the alternative spin polarization effects. It is essentially parallel to the spin-polarization effects on the  $\pi$ -conjugation network, which has been explored in the spin-alignment rule for molecular magnetism by Yamaguchi and his coworkers [45]. The second point is that there is an anti-parallel spin-polarization effect between S(38) and Fe(1) + Por. For instance, we can observe that the large parts of the off-diagonal blocks between S(38) and Fe(1) + Por are colored in green ( $\delta\rho_z(I)/\delta v_m(\text{S}(38)) = -0.1$ ). We also notice that the dull red color is spread over most of the Cys residue as shown in Figure 8. This is because the orthogonality among  $\text{sp}^3$  orbitals causes the weakly high-spin interactions, but of which the magnitudes are quite weak, resulting in the “dull” red color. In contrast, the off-diagonal parts between S(42) and the Cys residue, which are indicated by dotted white rectangles, are considerably clear red color. The result can be more clearly seen in the L-shaped section in the region of S(38)~N(46)H(47) in Figure 9a. This section is increasingly delocalized towards two terminal carbons of the two branches of the Cys residues, as the thresholds increases to  $\delta\rho_z(I)/\delta v_m(J) < 0.01$  (Figure 9b) and further to  $\delta\rho_z(I)/\delta v_m(J) < 0.001$ . This is

consistent with the spin density distribution of the whole system as will be shown in the following discussion: In fact, the spin distribution on S(42) is approximately 0.6, and such up-spin density is slightly delocalized over the Cys residue, indicating that there is originally weak high-spin coupling between S(42) and the Cys residue.



**Figure 8.** The heat map for the matrix representation of condensed version of linear response functions of the spin density  $\{\delta\rho_z(I)/\delta v_m(J)\}_{I,J=1\sim 59}$  for P450. The number of the axes corresponds to the numbering of the atoms indicated in Figure 1. The functional units and the couplings among units are also indicated (see text for the details).



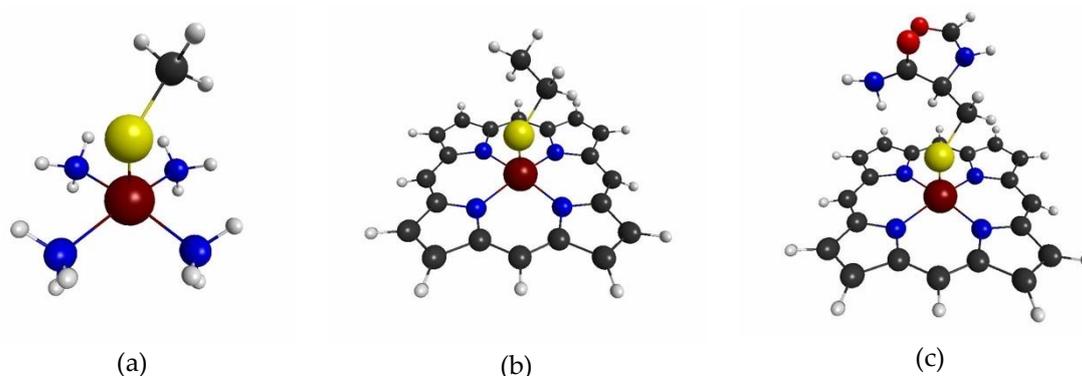
**Figure 9.** Matrix representation of condensed version of linear response functions of spin density for P450 with the six cutting thresholds. (a)  $\delta\rho_z(I)/\delta v_m(J) < 0.1$ . (b)  $\delta\rho_z(I)/\delta v_m(J) < 0.01$ . (c)  $\delta\rho_z(I)/\delta v_m(J) < 0.001$ , (d)  $\delta\rho_z(I)/\delta v_m(J) > -0.1$ . (e)  $\delta\rho_z(I)/\delta v_m(J) > -0.01$ . (f)  $\delta\rho_z(I)/\delta v_m(J) > -0.001$ .

### 3.1.4. QM Cluster and QM/MM Calculations for Several Models of P450

Next, we will show the actual QM/MM calculation results for several models shown in Figure 10. Note that we assumed that the reaction center (RC) of P450 is the Fe(1)S(38) unit. Model 1 shown in Figure 10a is the simplest model, where the porphyrin ring is approximated as four NH<sub>3</sub> molecules that ligate to the Fe ion and the Cys residue is approximated as CH<sub>3</sub>S<sup>−</sup> ion. In model 2, we included the porphyrin ring completely, but approximate the Cys residue as C<sub>2</sub>H<sub>5</sub>S<sup>−</sup> ion. Model 3 suits the criteria

$$|\delta\rho(I)/\delta v(RC)| < 0.01, \quad |\delta\rho_z(I)/\delta v_m(RC)| < 0.01 \quad (17)$$

the most. Here “RC” denotes the reaction center, where the accurate description of the density and spin density is required. Then, Equation (17) determines the peripheral atoms that should be included in the QM region. In all of the models, we replaced the original parts in the full model by capped H atoms. We computed both the QM cluster model and the QM/MM model, with the link atom method for models 1–3 [3–5]. The details of the latter calculations are described in the Materials and Methods section.



**Figure 10.** Geometries of the three models for the active site of P450: (a) model 1, (b) model 2, (c) model 3. We assumed that the full model is the model shown in Figure 1.

The calculated densities and spin densities of the active site are compared among these models and the full model shown in Figure 1. We assumed that the densities and spin densities of the full model are the reference values. We took the differences between the values calculated for models 1–3 and the reference values as the errors due to the modellings. Then, we listed the errors of densities on the Fe ion, the N atoms of the porphyrin ring, and the S ion of the Cys residue in Table 1, and the errors of spin densities for the same set of atoms Table 2. For the complete data of the errors on densities and spin densities, please refer to Tables S1 and S2, respectively. We can see from tables Tables 1 and 2, that the errors due to the simplest model (Model 1), are much larger than other two models. Furthermore, in model 1, the MM charges do not improve the results in general. On the other hand, the QM cluster models based on the models 2 and 3 reduce the errors on both densities and spin densities drastically. In particular, the errors of the model 3 with QM cluster model are less than 0.005, which indicates that model 3 is nearly equivalent to the full model. The reader might notice that there remain slightly larger errors in the results of the QM/MM model, based on model 3. These results are consistent with two well-known facts of the QM cluster and the QM/MM approaches [5]: First, the QM/MM results show more slower convergence behavior than the QM cluster results as the QM size increases. Second, the inclusion of the environmental effects, by a polarized model or a MM model, requires special care for the modeling and the parameters used. These are the reasons for some cases, where the QM cluster model outperforms the QM/MM model. In this case, the chains of the Cys residue lie parallel to the porphyrin plane, and the point charges of the QM/MM model that mimic the CH<sub>3</sub> beyond C(48), and the CH<sub>3</sub>CO beyond N(52) might slightly induce the polarization effects on the reaction center in model 3. We would like to emphasize the fact that model 3, by employing

the QM cluster model, yields the results that are nearly equivalent to those of the full model. This fact implies that the boundary of the QM model according to the criteria,  $|\delta\rho(I)/\delta v(\text{RC})| < 0.01$  and  $|\delta\rho_z(I)/\delta v_m(\text{RC})| < 0.01$ , as suggested above is the appropriate boundary for calculations of the reaction center of P450.

**Table 1.** The errors of the calculated values, obtained by the quantum mechanics (QM) cluster and the quantum mechanics/molecular mechanics (QM/MM) models and the reference values on densities ( $\Delta\rho$ ) of the Fe ion, the N atoms of the porphyrin ring, and the S ion of the Cys residue. The reference values are those of the full model shown in Figure 1.

Atom	Model 1		Model 2		Model 3	
	QM	QM/MM	QM	QM/MM	QM	QM/MM
Fe	0.069	0.099	−0.006	−0.013	0.001	0.014
N	0.025	0.034	0.001	0.001	0.000	−0.001
N	0.020	0.029	0.000	0.000	0.000	0.000
N	0.017	0.028	0.000	0.001	0.000	−0.001
N	0.020	0.027	0.004	0.003	0.000	−0.003
S	0.082	−0.018	−0.006	0.033	−0.001	−0.015

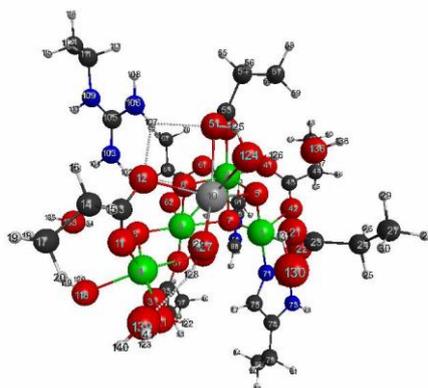
**Table 2.** The errors of the calculated values obtained by the QM cluster and the QM/MM models and the reference values on spin densities ( $\Delta\rho_z$ ) of the Fe ion, the N atoms of the porphyrin ring, and the S ion of the Cys residue. The reference values are those of the full model shown in Figure 1.

Atom	Model 1		Model 2		Model 3	
	QM	QM/MM	QM	QM/MM	QM	QM/MM
Fe	0.023	−0.070	0.007	0.017	−0.002	−0.016
N	0.015	0.002	0.002	0.005	−0.001	−0.004
N	0.018	0.004	0.000	0.002	0.000	−0.002
N	0.013	−0.002	0.000	0.001	0.000	−0.001
N	0.025	0.014	−0.001	0.000	0.000	−0.001
S	−0.113	0.024	−0.018	−0.039	0.003	0.035

### 3.2. Oxygen Evolving Complex of Photosystem II.

#### 3.2.1. The Target System

Next, we will examine the case of polynuclear complexes embedded in proteins. Di- $\mu$ -oxo bridged polynuclear Mn complexes, embedded in proteins, are also widely seen structural units for the reaction centers of biocatalyses, such as catalase [46], superoxide dismutase [47], and the oxygen evolving complex (OEC) in photosystem II [48–51]. Here, we choose the OEC geometry, that is extracted from the X-ray structure of the dark stable PS-II (pdb entry: 5b66) [48], of which the geometry is shown in Figure 11. The same picture, with the larger size, is shown in Figure S8 for checking the numberings of the atoms. In accordance with previous research on OEC [50,51], we assumed that the ground-state electronic configuration of the core of OEC in the dark stable ( $S_1$ ) state is Mn(III)Mn(IV)Mn(IV)Mn(III)Ca(II), which are bridged by di- $\mu$ -oxo ions ( $O^{2-}$ ) and carboxylates. All oxygen atoms that are not the parts of neither the cluster nor amino residues are assumed to be those of water molecules. Other modeling details are described in Materials and Methods section.



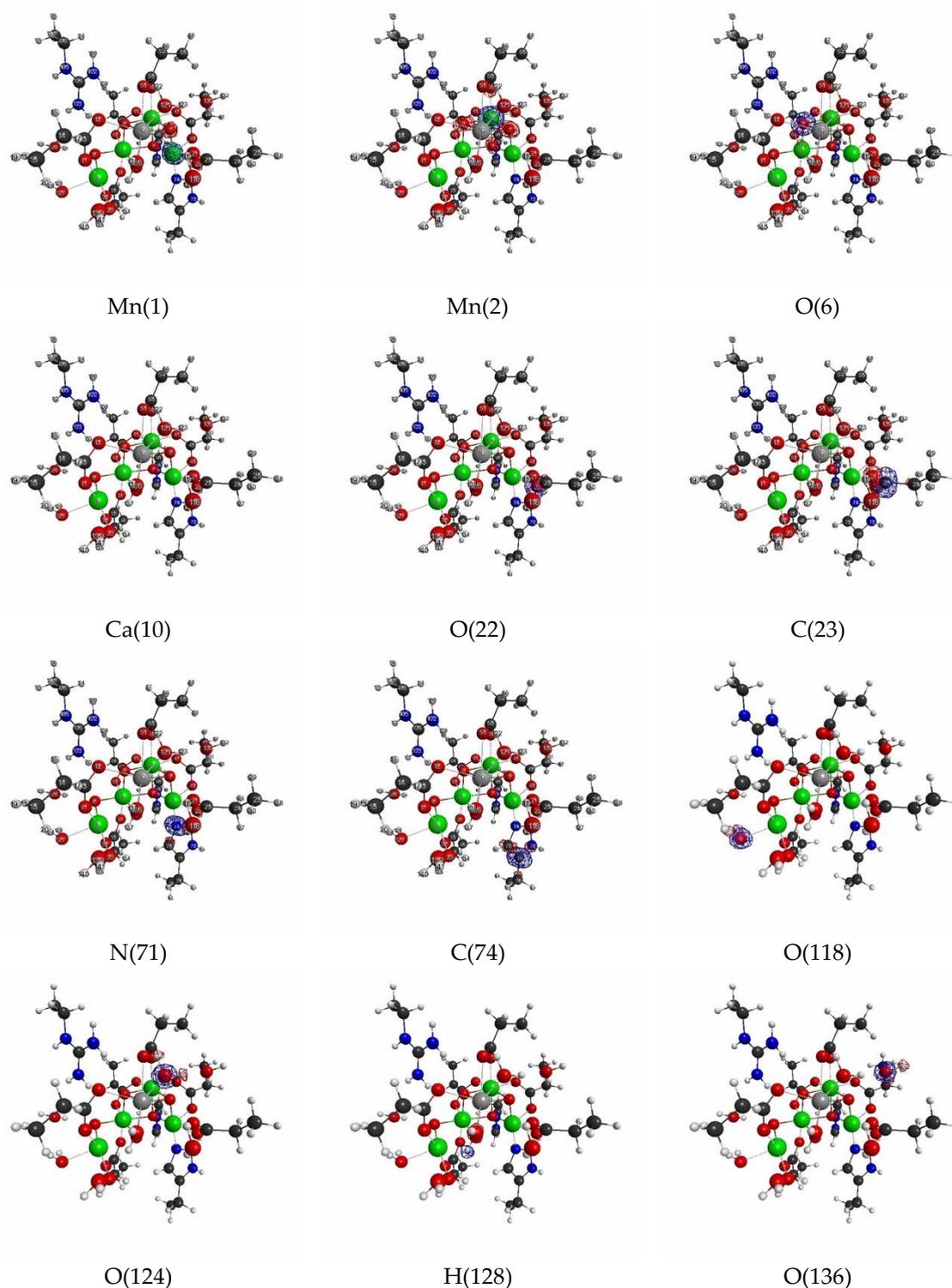
**Figure 11.** Geometry of the active site of oxygen evolving complex (OEC) (pdb entry:5b66). The numbering of the atoms is used for the following figures to show linear response functions. The larger picture is shown in Figure S8.

### 3.2.2. Linear Response Functions of Electron Density of OEC

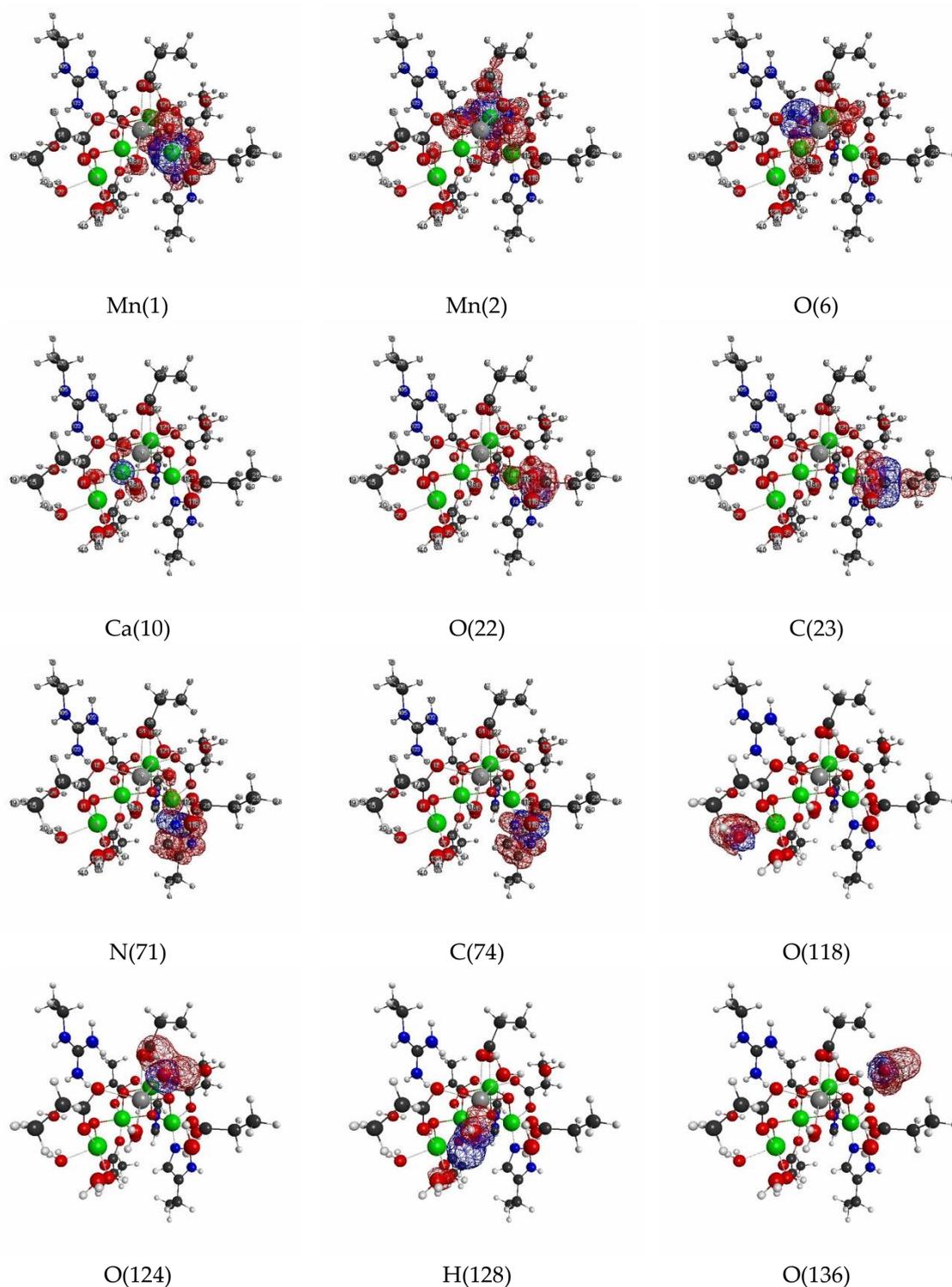
In Figures 12 and 13, we show the selected essential isosurfaces of LRFs for our discussion. Again, the thresholds for the isosurfaces are set to be  $\delta\rho(\mathbf{r})/\delta v(I) = \pm 0.1$  and  $\delta\rho(\mathbf{r})/\delta v(I) = \pm 0.01$ , where the variable  $I$  indicates the atom to which the virtual perturbation is applied. We only show the selected isosurfaces, due to the limits of space, but more comprehensive figures of the isosurfaces are shown in Figures S9 and S10 for  $\delta\rho(\mathbf{r})/\delta v(I) = \pm 0.1$ , and  $\delta\rho(\mathbf{r})/\delta v(I) = \pm 0.01$ , respectively. We can see from Figure 12 that even if the perturbation is applied to either the Mn, or the O ions, in the core cluster ( $\delta\rho(\mathbf{r})/\delta v(\text{Mn}(2))$  and ( $\delta\rho(\mathbf{r})/\delta v(\text{O}(6))$ ), the density fluctuation propagates at most to the nearest ions for the threshold,  $\delta\rho(\mathbf{r})/\delta v(I) = \pm 0.1$ . This implies that even the partitioning of the core cluster is possible at this approximation level. This is a very rough approximation, but the speculated mono- and di-nuclear models of OEC in the 1990s can be considered examples of such models ( $\delta\rho(\mathbf{r})/\delta v(I) = \pm 0.1$ ) [52], which were inevitable back then, because even the number of ions consisting the core cluster of the OEC was not completely determined before 2010 [48]. On the other hand, as shown in Figure 10, the isosurface for  $\delta\rho(\mathbf{r})/\delta v(\text{Mn}(2)) = \pm 0.01$  spreads beyond C(55), the carbon atom of the carboxylate that bridges between Mn(2) and Ca(10), and reaches at the next sp<sup>3</sup> junction, C(54). This implies that the approximation of glutamic acid and aspartic acid residues, as formic acids, is not appropriate. In fact, many researchers model glutamic acid or aspartic acid residues as acetic acids [5,50,51]. Note that the isosurface for  $\delta\rho(\mathbf{r})/\delta v(\text{Mn}(3)) = \pm 0.001$  are slightly delocalized to the two sp<sup>3</sup> junctions away, C(47) from the C(43)O(41)O(42) carboxylate, as shown in Figure S10. Thus, “the propionic acid” approximation for the carboxylate type of residues is appropriate, if we employ the criterion,  $\delta\rho(\mathbf{r})/\delta v(I) = \pm 0.001$ . Conversely, we can see from  $\delta\rho(\mathbf{r})/\delta v(\text{C}(23))$ , where O(22) of the carboxylate, C(23)O(22)O(24), ligates to Mn(1),  $\delta\rho(\mathbf{r})/\delta v(\text{C}(23))$  is delocalized over Mn(1) as shown in Figure 13. In contrast, the coupling between His residue and Mn ion is smaller than that between Glu or Asp residue and Mn ion as shown in  $\delta\rho(\mathbf{r})/\delta v(\text{N}(71))$  and  $\delta\rho(\mathbf{r})/\delta v(\text{C}(74))$  in Figure 13.

As for the interaction between water molecules and the OEC cluster, there are three possibilities: (i) Mn ion and water molecule, (ii) Ca ion and water molecule, and (iii) oxygen ion and water molecule. However, in this structure, there is no type (iii) of the interactions. From the our LRF analysis, type (i) interaction is detected by the threshold,  $\delta\rho(\mathbf{r})/\delta v(I) = \pm 0.01$ , as shown in  $\delta\rho(\mathbf{r})/\delta v(\text{O}(118))$ , but type (ii) is not as shown in  $\delta\rho(\mathbf{r})/\delta v(124)$  in Figure 13. Of course, the later interaction is detected by the threshold,  $\delta\rho(\mathbf{r})/\delta v(\text{O}(124)) = \pm 0.001$  as shown in Figure S10. Nevertheless, whether a hydrogen bond between water molecules is detected or not depends on the positions and orientations of the two water molecules: See  $\delta\rho(\mathbf{r})/\delta v(128)$  and  $\delta\rho(\mathbf{r})/\delta v(136)$  in Figure 13. Further the extensive hydrogen bonding network over many water molecules and other anionic or cationic parts of residues and cluster are detected by the threshold,  $\delta\rho(\mathbf{r})/\delta v(I) = \pm 0.001$ : see, for instance  $\delta\rho(\mathbf{r})/\delta v(\text{O}(127))$  in Figure S10.

Thus, if the QM/MM modeling is to search for the reaction mechanism, involving movements of water molecules near the cluster, it is better to include all water molecules in the QM region.



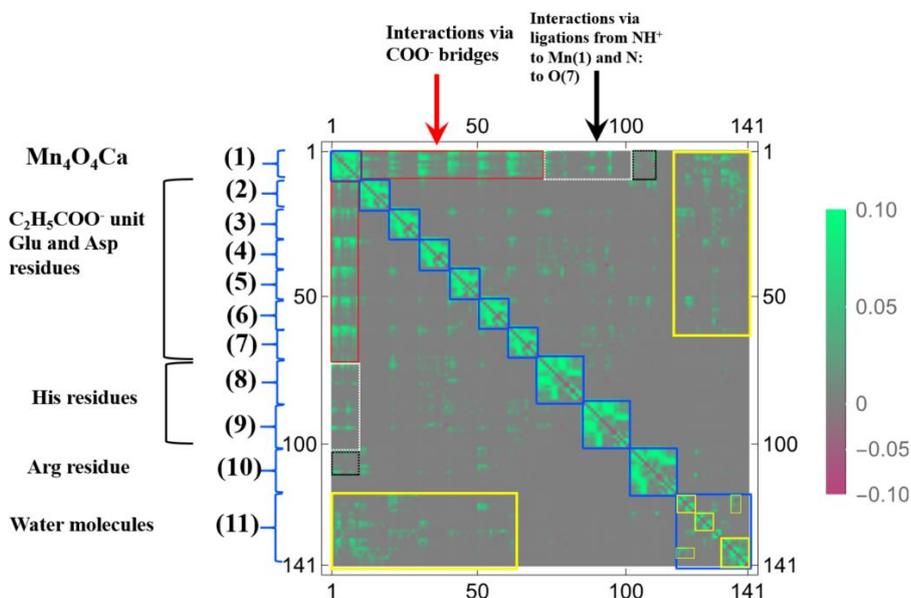
**Figure 12.** Isosurfaces of linear response functions of density for the perturbations on atomic sites of OEC (pdb entry:5b66). The threshold of  $\delta\rho(r)/\delta v(I)$  is  $\pm 0.1$  (+: blue, -: red). The atomic symbols and those numbers in parentheses below the figures indicate the atoms (I) to which the virtual perturbations are applied.



**Figure 13.** Isosurfaces of linear response functions of density for the perturbations on atomic sites of OEC (pdb entry:5b66). The threshold of  $\delta\rho(\mathbf{r})/\delta v(I)$  is  $\pm 0.01$  (+: blue, -: red). The atomic symbols and those numbers in parentheses below the figures indicate the atoms (I) to which the virtual perturbations are applied.

The more compact representation of Figures S9–S11 is shown in Figure 14 as a heat-map representation of the matrix,  $\{\delta\rho(I)/\delta v(J)\}_{I,J=1\sim 141}$ . Obviously we can see from this figure that there are 11 diagonal sub-blocks, corresponding to the  $\text{Mn}_4\text{O}_5\text{Ca}$  core (1), the amino residues (2)~(10), and the group of water molecules (11), which are encircled by blue squares. As for the interactions

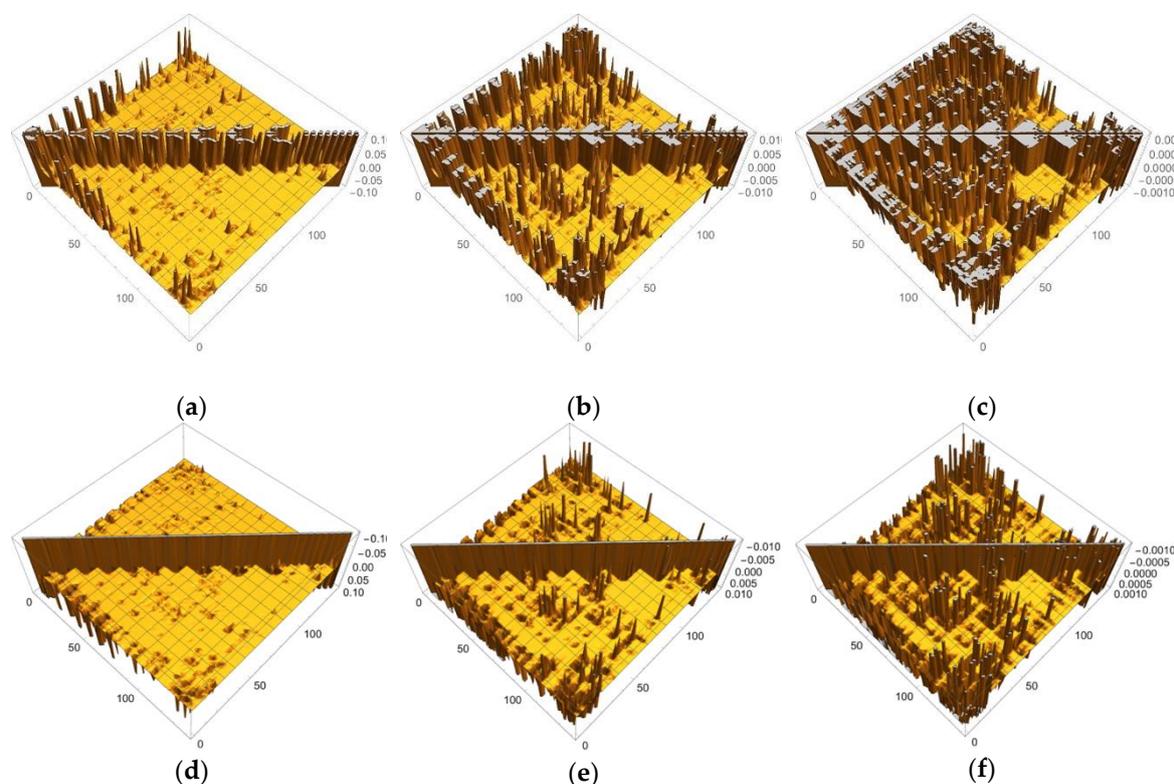
among the sub-blocks, we first noticed the periodic structures encircled in the red rectangles. These are the coupling units where carboxylates bridge Mn and Mn, or Mn and Ca, in the  $\text{Mn}_4\text{O}_5\text{Ca}$  core. Among them, the off-diagonal part between (1) and (3) corresponds to  $\delta\rho(r)/\delta v(\text{C}(23))$ , shown in Figure 13. Also the interactions between the  $\text{Mn}_4\text{O}_5\text{Ca}$  core and the His residues via the ligations of  $\text{NH}^+$  to Mn(1) and N to O(7), are observed in the white rectangles. In addition, the interactions between the Arg residue and the  $\text{Mn}_4\text{O}_5\text{Ca}$  core are also observed, but the magnitudes of them are considerably weaker, as shown in the dotted black rectangles. In the sub-block consisting of water molecules (11), there are several hydrogen bonding networks, which are indicated by the small yellow rectangles and squares. In addition, there are other hydrogen bonding networks in the off-diagonal region that are encircled by yellow rectangles, which includes the ligations of the water molecules to Mn and to Ca, respectively as  $\delta\rho(r)/\delta v(\text{O}(118))$  and  $\delta\rho(r)/\delta v(\text{O}(124))$  shown in Figure 13.



**Figure 14.** The heat map for the matrix representation of condensed version of linear response functions of electron density,  $\{\delta\rho(I)/\delta v(J)\}_{I,J=1\sim 141}$ , for OEC. The number of the axes corresponds to the numbering of the atoms indicated in Figure 11. The functional units and the couplings among units are also indicated (see text for the details).

In order to discuss our results from the viewpoint of the QM/MM modeling more clearly, the sections of the LRF of the density with six thresholds are summarized in the matrix forms, as shown in Figure 15a–f. First, we can see from Figure 15d–f that the increases of electron density, due to attractive perturbations ( $\delta\rho(I)/\delta v(J) < 0$ ), are considerably localized in the diagonal region,  $\{\delta\rho(I)/\delta v(I)\}_{I=1\sim 141}$ . Next, we again confirm from Figure 15a–c that there are distinct sub-blocks in the diagonal region, which correspond to the  $\text{Mn}_4\text{O}_5\text{Ca}$  core, 9 amino acid residues, and the group of the water molecules. The important point is that the number of off-diagonal parts increases as the threshold becomes tighter from Figure 15a ( $\delta\rho(I)/\delta v(J) < 0.1$ ) to Figure 15c ( $\delta\rho(I)/\delta v(J) < 0.001$ ). As described above, the criterion,  $\delta\rho(I)/\delta v(J) < 0.1$ , detects the strong couplings such as covalent and ionic bonds: We should note that the  $\text{sp}^3$  junctions of the residues clearly cut the islands at this level and that the weak ionic bonds such as hydrogen bonds is difficult to detect. In other words, the QM/MM modeling under this criterion is very rough. On the other hand, the model that is constructed according to the guideline,  $\delta\rho(I)/\delta v(J) < 0.01$ , the core part is described with the correct surrounding effects such as hydrogen bonds and the hyperconjugation effects. In fact, almost all models that are employed by most QM/MM researchers seem to satisfy this criterion [5–17,50,51]. As shown Figure 15c, the criterion,  $\delta\rho(I)/\delta v(J) < 0.001$ , leads to a greater number of interaction islands, which interconnect among units, and so, in other words, it requires a large QM region. In fact, the

extent to which the LRF of density spreads over is shown in Figure S11. If such an accurate QM computational method is available and is needed to describe the electronic structure of the core part of an enzyme, we had better employ this criterion. However, we do not recommend this accurate criterion for actual QM/MM calculations using DFT because, even in the pure QM calculations, the approximated exchange-correlation terms always leads to errors on the description of the density.



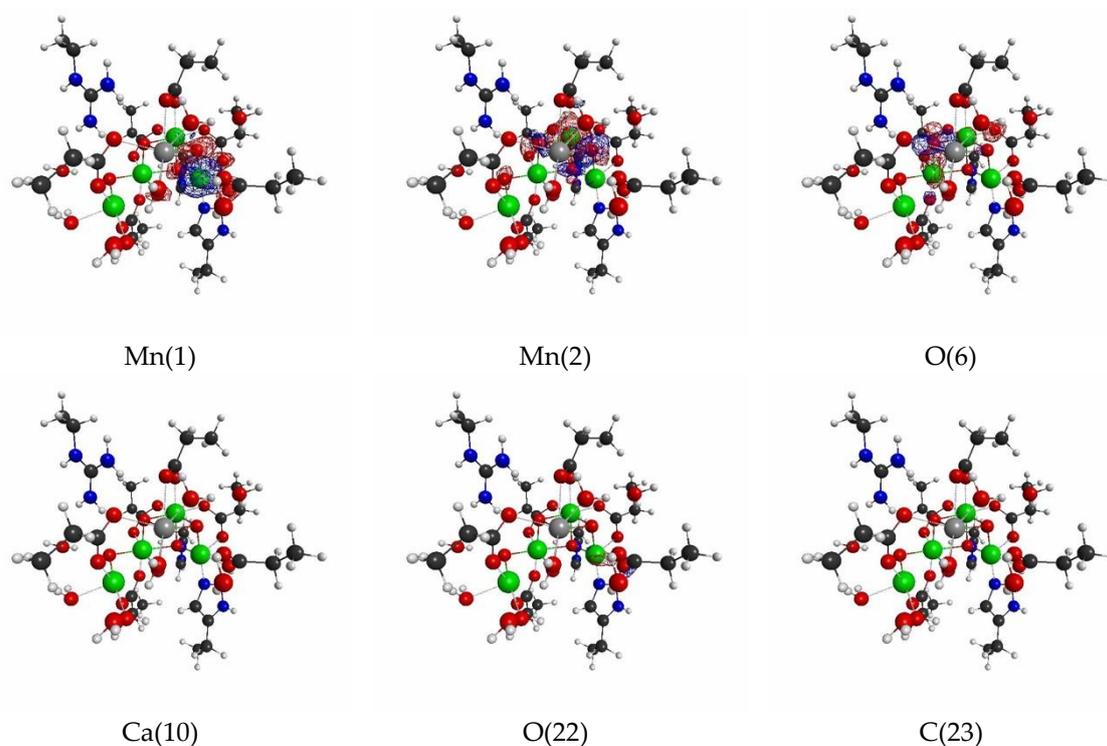
**Figure 15.** Matrix representation of condensed version of linear response functions of electron density for OEC with the six cutting thresholds. (a)  $\delta\rho(I)/\delta v(J) < 0.1$ . (b)  $\delta\rho(I)/\delta v(J) < 0.01$ . (c)  $\delta\rho(I)/\delta v(J) < 0.001$ , (d)  $\delta\rho(I)/\delta v(J) > -0.1$ . (e)  $\delta\rho(I)/\delta v(J) > -0.01$ . (f)  $\delta\rho(I)/\delta v(J) > -0.001$ .

### 3.2.3. Linear Response Functions of Spin Density of OEC

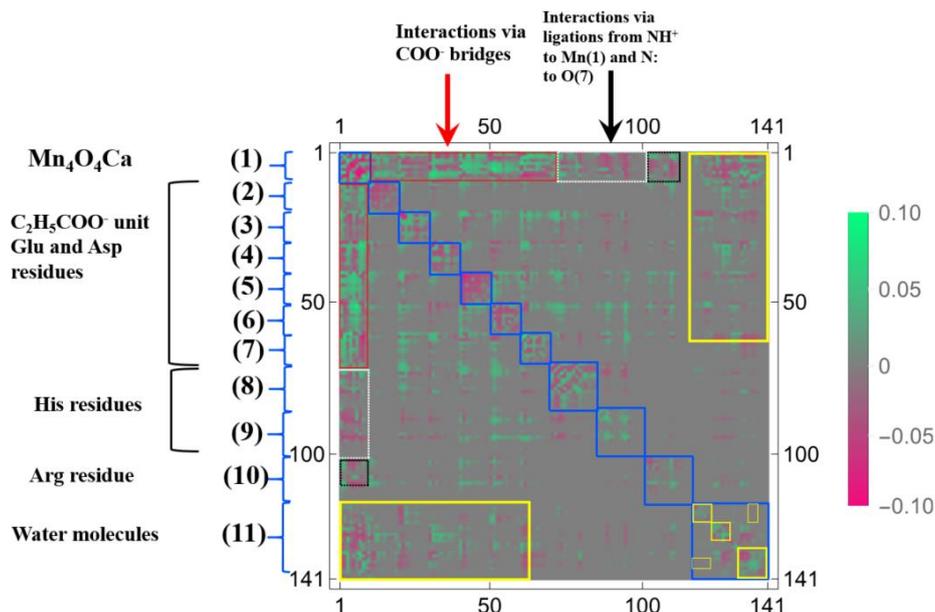
Next, we show the results of LRFs of spin density for the OEC. In fact, it is known that the ground-state of the dark stable state has anti-ferromagnetic coupling, consisting of two Mn(III) ions and Mn(IV) ions, that are coupled with carboxylate and di- $\mu$ -oxo ions bridges, which require the correct treatment of the spin state. Figure 16 shows selected important isosurfaces of LRFs of spin density for the threshold,  $\delta\rho_z(\mathbf{r})/\delta v_m(I) = \pm 0.01$ . Before discussing the results shown in Figure 16, we should note that almost all LRFs of spin density for the threshold,  $\rho_z(\mathbf{r})/\delta v_m(I) = \pm 0.1$ , become negligible as shown in Figure S12: Only the perturbations to Mn ions and bridged O ions yield the apparent fluctuations of spin density. As shown in Figure 16 ( $\delta\rho_z(\mathbf{r})/\delta v_m(\text{Mn}(1))$ ) and  $\delta\rho_z(\mathbf{r})/\delta v_m(\text{Mn}(2))$ ), the isosurfaces for  $\delta\rho_z(\mathbf{r})/\delta v_m(I) = \pm 0.01$  spread over the core and bridged units. Not surprisingly, the perturbations to atoms of water molecules (for instance, O(118), O(124), O(136)) do not emerge spin fluctuations, because there are negligibly spin polarized. However, it is unexpected that, at this threshold, all the isosurfaces for the perturbations to the N(71) and C(74) atoms of His that ligates to Mn(1), and to O(118) and O(121) atoms of H<sub>2</sub>O that ligates to Mn(4), are negligible. Further, we can see from Figure 16a–c that the LRF of spin density for OEC is considerably localized, compared with that of density shown in Figure 13a–c.

All other non-negligible isosurfaces for  $\delta\rho_z(\mathbf{r})/\delta v_m(I) = \pm 0.01$  and  $\delta\rho_z(\mathbf{r})/\delta v_m(I) = \pm 0.001$  are shown in Figures S13 and S14, respectively. As an overview, we plot a heat-map representation of the

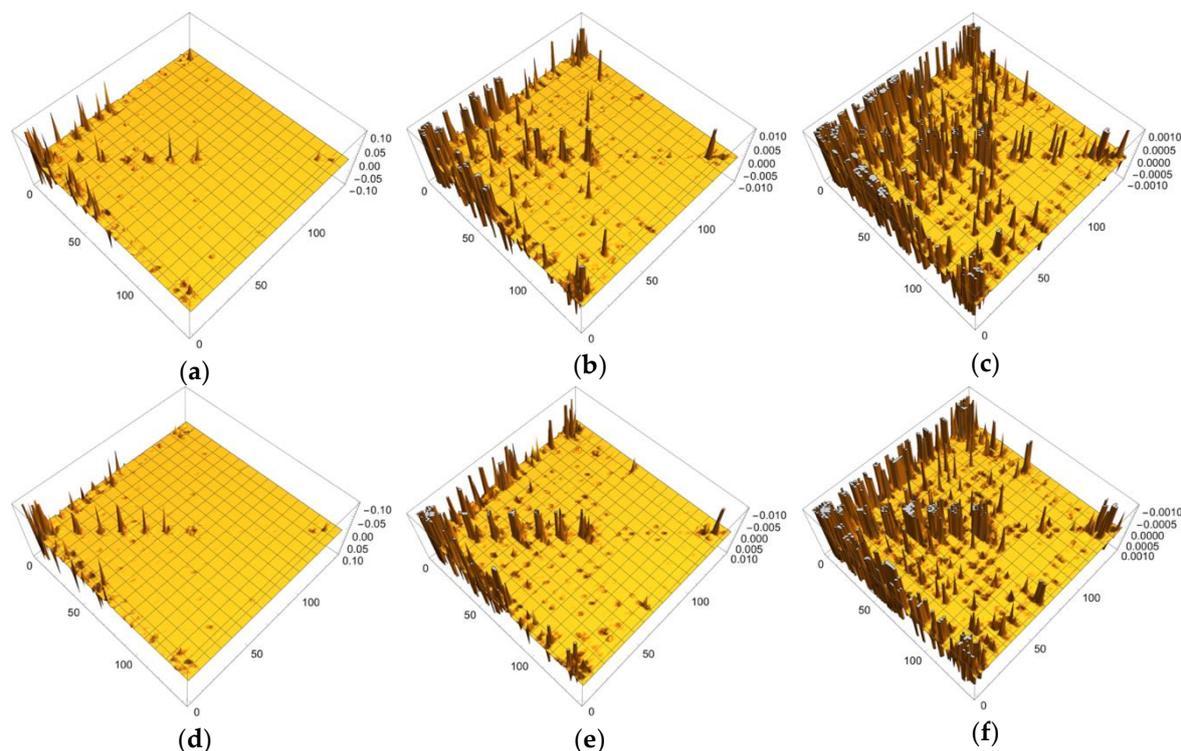
matrix,  $\{\delta\rho_z(I)/\delta v_m(J)\}_{I,J=1\sim 141}$  in Figure 17. In addition, we plot the sections for the six thresholds in Figure 18a–f. We can see that the gray region in the heap-map for  $\{\delta\rho_z(I)/\delta v_m(J)\}_{I,J=1\sim 141}$  shown in Figure 17 is larger than that in the heat-map for  $\{\delta\rho(I)/\delta v(J)\}_{I,J=1\sim 141}$  shown in Figure 14. Since the gray region corresponds to the region where there is nearly no response, i.e.,  $\delta\rho_z(I)/\delta v_m(J) \cong 0.0$ , the almost all region of the OEC model is not sensitive to the virtual magnetic perturbations. Also as shown in Figure 18a,d, the sections do not exist for both  $\delta\rho_z(I)/\delta v_m(J) < 0.1$  and  $\delta\rho_z(I)/\delta v_m(J) > -0.1$ , being consistent with the above-mentioned fact that almost all LRFs of spin density for the threshold,  $\rho_z(\mathbf{r})/\delta v_m(I) = \pm 0.1$ , become negligible, as shown in Figure S11. However as the threshold becomes tighter, the number of sections increase, as shown in Figure 18b,c,e,f. These results are consistent with the fact that the isosurfaces for  $\delta\rho_z(\mathbf{r})/\delta v_m(I) = \pm 0.01$  spread over the Mn ions in core and bridged carboxylate units as described above (see figures  $\delta\rho_z(\mathbf{r})/\delta v_m(I) = \pm 0.01$ ,  $I = \text{Mn}(1), \text{Mn}(2), \text{O}(6), \text{O}(22), \text{C}(23)$  in Figure 16). Furthermore, as shown in Figure 18, for the threshold,  $\delta\rho_z(I)/\delta v_m(J) < 0.001$  and  $\delta\rho_z(I)/\delta v_m(J) > -0.001$ , the sections appear in the off-diagonal regions, (i) between the  $\text{Mn}_4\text{O}_5\text{Ca}$  core and the His residue, (ii) the  $\text{Mn}_4\text{O}_5\text{Ca}$  core and the Arg residue, and (iii)  $\text{Mn}_4\text{O}_5\text{Ca}$  core and the group of water molecules. These interactions, together with the intra residue and the intra water cluster interactions are shown as dull green and dull red in Figure 17.



**Figure 16.** Isosurfaces of linear response functions of spin density for the perturbations on atomic sites of OEC (pdb entry:5b66). The threshold of  $\delta\rho_z(\mathbf{r})/\delta v_m(I)$  is  $\pm 0.01$  (+: blue, -: red). The atomic symbols and those numbers in parentheses below the figures indicate the atoms ( $I$ ) to which the virtual perturbations are applied.



**Figure 17.** The heat map for the matrix representation of condensed version of linear response functions of spin density,  $\{\delta\rho_z(I)/\delta v_m(J)\}_{I,J=1\sim 141}$ , for OEC. The number of the axes corresponds to the numbering of the atoms indicated in Figure 11. The functional units and the couplings among units are also indicated (see text for the details).



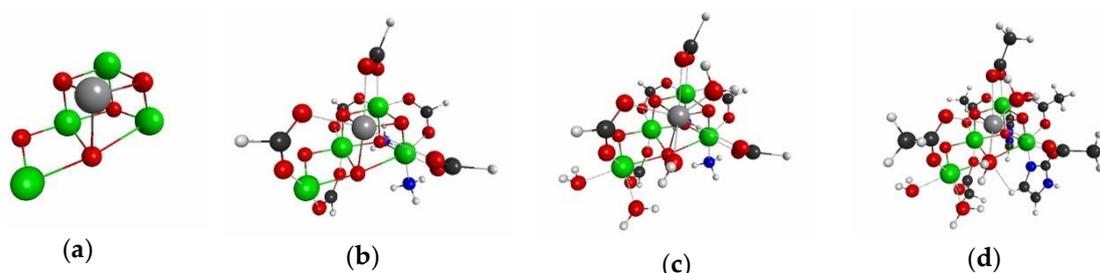
**Figure 18.** Matrix representation of condensed version of linear response functions of spin density for OEC with the six cutting thresholds. (a)  $\delta\rho_z(I)/\delta v_m(J) < 0.1$ . (b)  $\delta\rho_z(I)/\delta v_m(J) < 0.01$ . (c)  $\delta\rho_z(I)/\delta v_m(J) < 0.001$ , (d)  $\delta\rho_z(I)/\delta v_m(J) > -0.1$ . (e)  $\delta\rho_z(I)/\delta v_m(J) > -0.01$ . (f)  $\delta\rho_z(I)/\delta v_m(J) > -0.001$ .

If we consider the threshold,  $\delta\rho_z(r)/\delta v_m(I) = \pm 0.01$  is appropriate for the guideline to construct the QM/MM model; the small QM model consist of the  $\text{Mn}_4\text{O}_5\text{Ca}$  core and the carboxylates, that ligate to the  $\text{Mn}_4\text{O}_5\text{Ca}$  core enough to describe the magnetic systems embedded in non-magnetic environments such as proteins and waters. One of the theoretical reasons in raising such a question,

is that the degree of freedom of charges basically does not couple with that of the spins. As most of the biological environments are affected electrostatically, and not magnetically, a large QM region does not always improve the descriptions of the magnetic systems of the active site. In fact, *ab initio* DFT calculations showed that the counter anions in molecular solids, including di-nuclear Mn complexes do not affect the intra-molecular magnetic interactions of the Mn complexes considerably [53–55]. However, we should note that our guidelines given by Equation (17) are not only of the spin density, but also of the density. In addition, we should note that the energy resolution to correctly estimate the relative stability of degenerate states for the Mn complexes, such as OEC, is less than that of the so-called “chemical accuracy”(1 kcal/mol). Because the magnitude of the magnetic interactions between the Mn ions bridged with carboxylate and di- $\mu$ -oxo ions are usually in the order of  $100\text{ cm}^{-1}$  or less ( $<0.3\text{ kcal/mol}$ )[53], which is less than one-tenth of the binding energy of a usual hydrogen bond. Thus in order to determine the spin state of the ground state among many degenerate states, we usually need a far more accurate method to describe the electrostatic interactions correctly. Therefore small electrostatic interactions might affect the relative stabilities among degenerate spin states, due to the higher order effects on the electronic structure of the active site. This implies that the criterion of the LRF of spin density must not be  $\rho_z(r)/\delta v_m(I) = \pm 0.01$ , but  $\delta\rho_z(r)/\delta v_m(I) = \pm 0.001$  or tighter criterions for describing nearly degenerate spin states. However, we would like to emphasize again that, for estimating the correct stable spin states, the problem remains rather in the approximation of the exchange-correlation functional we used than how to model the system by the QM/MM method [53,55]. In the next section, we examine the specific spin state and oxidation state that were proposed to be as the ground state of the OEC core [50,51] to evade the issue, such as the performance of the XC functional for estimating the relative stability among many degenerate states. In other words, we focus our attention only on the errors of densities and spin densities from the full QM model, which is given by Figure 11 in this study.

### 3.2.4. QM Cluster and QM/MM Calculations for Several Models of OEC

We now show the actual QM/MM calculation results for several models shown in Figure 19. Here we assumed that the reaction center (RC) of OEC is the  $\text{Mn}_4\text{O}_5\text{Ca}$  core, which is considered as model 1 shown in Figure 19a. In model 2, presented in Figure 19b, we approximate all the residues of glutamic acids and aspartic acids residues as formic acids, the His residue as  $\text{NH}_3$ , and the protonated His residue as  $\text{NH}_4^+$ . Model 3 further includes water molecules that ligate to the Ca ion or the Mn ion in the  $\text{Mn}_4\text{O}_5\text{Ca}$  core, as shown in Figure 19c, resulting in the coordination saturation model. Figure 19d shows Model 4, which suits the criteria  $|\delta\rho(I)/\delta v(\text{RC})| < 0.01$  and  $|\delta\rho_z(I)/\delta v_m(\text{RC})| < 0.01$ . In all of the models, we replaced the original parts in the full model by capped H atoms. We computed both the QM cluster model and the QM/MM model with the link atom method for the models 1–4 [3–5]. The details of the calculations are described in the Materials and Methods section.



**Figure 19.** Geometries of the three models for the active site of OEC: (a) model 1, (b) model 2, (c) model 3, (d) model 4. We assumed that the full model is the model shown in Figure 11.

Then, we listed the errors of densities and spin densities of the  $\text{Mn}_4\text{O}_5\text{Ca}$  core in Table 1, and Table 2, respectively. For complete data of the errors on densities and spin densities, please refer to Tables S3 and S4, respectively. We can see from tables Table 3 Table 4 that the errors due to the simplest

model (Model 1) are too large to discuss the oxidation state and spin state: The maximum errors are 1.158 (Ca(10)) for densities and  $-0.858$  (Mn(2)) for the spin densities. Such large errors are not improved by the QM/MM model for Model 1. Introducing all the carboxylates that bridge among the Mn ions and the Ca ion improve the results considerably, in particular—the spin densities. The errors of spin densities and electron densities reduce to, at most  $\pm 0.2$ , and  $\pm 0.4$ , respectively. In Model 2, the effective point charges of the surrounding atoms do not improve the accuracy of the QM cluster model. The acceptable accuracy is obtained with Model3, of which the magnitudes of all errors reduce to less than 0.1. Furthermore, Model 4 yields quite accurate results compared with the full model. In addition, the QM/MM model based on Model 4 out-perform the QM cluster model based on Model 4. In particular, almost all of the errors of  $\Delta\rho_z$  of the  $\text{Mn}_4\text{O}_5\text{Ca}$  core are less than 0.01, implying that the QM/MM model based on Model4 is nearly equivalent to the full QM model. We can also see from Table S3 that for Model 4, that artificial polarizations occur on the peripheral atoms (see the densities of the terminal carbons of amino residues,  $\{\text{C}(x) \mid x = 14,243,444, 54,647,490\}$  and the terminal hydrogen atoms of water molecules  $\{\text{H}(x) \mid x = 120,126,128,129\}$ ); the atoms in the inner region are nearly completely protected from the artificial electric fields and the edge effects (the QM cluster) because of the screening effects due the electrons in the outer regions. Since Model 4 is the model that satisfies our guidelines given by Equation (17), this result exemplifies the validity of our method to determine the QM region in the QM/MM calculation for the reaction center embedded in proteins.

**Table 3.** The errors of the calculated values obtained by the QM cluster and the QM/MM models from the reference values on densities ( $\Delta\rho$ ) of the  $\text{Mn}_4\text{O}_5\text{Ca}$  core. The reference values are those of the full model shown in Figure 11.

Atom	Model 1		Model 2		Model 3		Model 4	
	QM	QM/MM	QM	QM/MM	QM	QM/MM	QM	QM/MM
Mn(1)	0.941	1.069	0.051	0.032	0.041	0.025	0.004	0.005
Mn(2)	0.835	1.050	0.036	0.026	0.030	0.021	0.017	0.014
Mn(3)	0.571	0.687	0.041	0.013	0.024	$-0.001$	0.019	0.001
Mn(4)	1.071	1.134	0.237	0.365	0.026	0.043	0.009	0.024
O(5)	0.185	0.192	$-0.003$	0.024	$-0.001$	0.025	$-0.013$	0.005
O(6)	0.230	0.158	0.036	$-0.006$	0.042	$-0.003$	0.037	$-0.014$
O(7)	0.236	0.033	$-0.015$	$-0.034$	$-0.019$	$-0.036$	$-0.008$	$-0.007$
O(8)	0.380	0.178	0.071	0.021	0.021	$-0.002$	0.011	$-0.016$
O(9)	0.224	0.153	$-0.004$	0.002	$-0.009$	0.002	$-0.011$	0.004
Ca(10)	1.158	1.175	0.364	0.412	0.022	0.004	0.014	$-0.005$

**Table 4.** The errors of the calculated values obtained by the QM cluster and the QM/MM models from the reference values on densities ( $\Delta\rho_z$ ) of the  $\text{Mn}_4\text{O}_5\text{Ca}$  core. The reference values are those of the full model shown in Figure 11.

Atom	Model 1		Model 2		Model 3		Model 4	
	QM	QM/MM	QM	QM/MM	QM	QM/MM	QM	QM/MM
Mn(1)	0.321	0.305	$-0.012$	$-0.015$	$-0.012$	$-0.015$	$-0.003$	$-0.002$
Mn(2)	$-0.858$	$-0.634$	$-0.012$	$-0.026$	$-0.026$	$-0.033$	$-0.020$	$-0.004$
Mn(3)	0.267	0.272	$-0.010$	0.027	$-0.027$	0.015	$-0.033$	0.008
Mn(4)	$-0.630$	$-0.507$	$-0.162$	$-0.141$	$-0.007$	$-0.001$	$-0.006$	0.000
O(5)	0.112	0.148	$-0.005$	0.004	$-0.002$	0.005	0.000	0.002
O(6)	0.224	0.157	0.012	0.003	0.033	0.014	0.031	0.004
O(7)	0.167	0.116	$-0.017$	$-0.015$	$-0.008$	$-0.011$	$-0.004$	$-0.001$
O(8)	0.627	0.306	0.092	0.036	0.019	$-0.006$	0.021	0.001
O(9)	$-0.111$	$-0.048$	$-0.040$	$-0.060$	0.021	$-0.016$	0.026	$-0.011$
Ca(10)	$-0.003$	$-0.001$	0.002	0.000	0.001	0.000	0.001	0.000

## 4. Materials and Methods

### 4.1. Numerical Details for Computations of “Condensed” Versions of the Linear Response Function

Here we describe the numerical details to compute condensed versions of the linear response function (LRF). The purpose is to obtain the density change at  $\mathbf{r}$  due to the perturbation that is applied to the J-th atom,  $\delta\rho(\mathbf{r})/\delta v(J)$ ,

$$\frac{\delta\rho(\mathbf{r})}{\delta v(J)} = \int_{J\text{-th atom}} d\mathbf{r}' \frac{\delta\rho(\mathbf{r})}{\delta v(\mathbf{r}')} \quad (18)$$

and the density change at the I-th atom due to the perturbation that is applied to the J-th atom,  $\delta\rho(I)/\delta v(J)$ ,

$$\frac{\delta\rho(I)}{\delta v(J)} = \int_{I\text{-th atom}} d\mathbf{r} \int_{J\text{-th atom}} d\mathbf{r}' \frac{\delta\rho(\mathbf{r})}{\delta v(\mathbf{r}')}. \quad (19)$$

Here the point is the numerical integration over all space  $\{\mathbf{r}\}$ , and  $\{\mathbf{r}'\}$ . We employ the Wigner-Seitz cell for the domain of integration for the I-th (and J-th) atom, where the cell for a specific atom in the molecule can be defined as a region encircled by all perpendicular bisectors with neighboring atoms. Then the 50 point Euler-Maclaurin quadrature and the 302 point Lebedev quadrature are used for radial, and spherical integrations [56,57], respectively. This is a similar quadrature scheme to that used in Kohn-Sham DFT code used in quantum chemistry program [58], but the fuzzy cell scheme is not used. The above numerical scheme is also used for the LRF of spin density.

### 4.2. Computational Details of Transition Metal Complexes

We used the B3LYP functional for the exchange-correlation (XC) function for all spin-unrestricted DFT calculations. The basis sets used in the computations are follows. The Wachter+f basis is used for Fe and Mn for all transition metal complexes [59,60]. The 6-31G\*\* basis are used for other atoms of P450. For C, N, O, H atoms of reaction center of PS II, 6-31G\* basis is used. The missing protons in pdb files are added and the partial optimization calculations are performed using similar XC and basis sets with fixing the positions of all heavy atoms. In order to facilitate the convergences of self-consistent and geometry optimization partial procedures, we used Gaussian 09 Rev. C first [60,61], and then using molecular orbitals obtained after diagonalization of the Kohn-Sham Hamiltonian constructed from the converged molecular orbitals. We performed linear response computations by using the modified version of GAMESS [41]. We employed MacMolPlt [62] to visualize the equi-valued surfaces of LRF, and Mathematica Ver. 11 [63] for visualizing the matrix form of LRFs, respectively. For the QM/MM calculations described in Sections 3.1.4 and 3.2.4, we employed the standard approach of the QM/MM method with the link-atom method [3–5]: i.e., we approximated the MM parts except the frontier MM atoms by the point charges that are determined by the electro-static potential fitting (ESP) method [64]. The positions of the capped H atoms are partially optimized. For the computation of the partially optimization of the capped H atoms and the computation of the ESP charges of the MM parts of P450 and OEC models, we used Gaussian 09 Rev. C [60].

## 5. Conclusions

We computed the linear response functions (LRFs) of mono- and poly-nuclear transition metal complexes, which have open-shell transition metal ions. From our computations, it is concluded that the order of the LRF values for the virtual perturbations, that are applied to the surrounding atoms, i.e.,  $\delta\rho(\mathbf{r})/\delta v(J)$  and  $\delta\rho_z(\mathbf{r})/\delta v_m(J)$ , can be used to determine the approximation level of the QM/MM modeling for such systems. The modeling that satisfies  $|\delta\rho(\mathbf{r})/\delta v(J)|$  (and  $|\delta\rho_z(\mathbf{r})/\delta v_m(J)|$ )  $< 0.1$  is a very rough model. The model covers the environments that are strongly coupled with the active spot, but weak couplings, due to hydrogen bonding and hyper-conjugations are missed. The residues that directly ligate to the active site should be included in the model. Even the partitioning of the poly-nuclear transition metal cores is often possible if we employed this criterion.

This was often used for biochemical systems until the 1990s [52]. If the model satisfies the criterion,  $|\delta\rho(\mathbf{r})/\delta v(J)|$  (and  $|\delta\rho_z(\mathbf{r})/\delta v_m(J)|$ )  $< 0.01$ , it is consistent with almost all QM/MM modeling [5,50,51]. In order to satisfy the guidelines given by Equation (17), i.e.,  $|\delta\rho(\mathbf{r})/\delta v(J)| < 0.01$  and  $|\delta\rho_z(\mathbf{r})/\delta v_m(J)| < 0.01$ , the next surrounding environments beyond the weakly coupled interconnections, such as hydrogen bonds and hyperconjugations over  $sp^3$  junction, should be included in the QM region. In fact, our QM cluster and QM/MM models, based on these guidelines, yield accurate densities and spin densities of the reaction centers of P450 and OEC (see Sections 3.1.4 and 3.2.4), compared with the full QM model.

Of course, the use of  $\delta\rho(\mathbf{r})/\delta v(J)$  and  $\delta\rho_z(\mathbf{r})/\delta v_m(J)$  implies that the above suggestions would become meaningless if we developed an accurate methods to reduce  $\delta v$  and  $\delta v_m$  to nearly zero: What we should treat with the QM approach would only be the actual spot that is responsible for the QM phenomenon. However, it would be very difficult to develop such a “nearly exact” approximation, by which the environmental effects, except the active spot, can be re-normalized in some effective fields. In addition, before we proceed to such an accurate level, we need to compare the errors due to QM/MM modeling with those resulting from the choice of the XC functional or the basis set that we used.

Comparing the LRF of density and the LRF of spin density, the conclusion is that the density of the active site is more sensitive than the spin density of the active site for the effects from the perturbations of surrounding environments. However, this does not imply that the small QM model is enough for an accurate description of the transition metal complexes, having open-shell transition metal ions embedded in proteins. To obtain accurate electron densities and spin densities for such systems, the model is required to satisfy not only the condition for electron densities, but also that for spin densities. In fact, our QM cluster and QM/MM calculations showed that the models which satisfy both  $|\delta\rho(\mathbf{r})/\delta v(J)| < 0.01$  and  $|\delta\rho_z(\mathbf{r})/\delta v_m(J)| < 0.01$  yields accurate results that are nearly equivalent to those of the full QM model. Furthermore, in order to determine the spin-state of the ground-state correctly, more accurate computations would be needed. This would require both a large QM region to satisfy the criterion,  $|\delta\rho_z(\mathbf{r})/\delta v_m(J)| < 0.001$ , and an accurate QM method.

**Supplementary Materials:** The following are available online, Figures S1–S14; Tables S1–S4.

**Author Contributions:** Theory and methodology, S.Y., C.K.K., T.M.; software, S.Y., J.O., T.M., C.K.K.; validation, T.K., M.O., K.Y.; formal analysis, S.Y.K. and C.K.K.; investigation, J.O., T.M., and C.K.K.; data curation, J.O., T.M., and C.K.K.; writing—original draft preparation, S.Y., and C.K.K.; writing—review and editing, S.Y.; visualization, J.O., T.M., C.K.K., and Y.M.; supervision, M.O., K.Y., and S.Y.; project administration, S.Y.

**Funding:** This research received no external funding.

**Acknowledgments:** S.Y. acknowledges Professor Haruki Nakamura for his valuable discussion.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Warshel, A.; Karplus, M. Calculation of ground and excited state potential surfaces of conjugated molecules. I. Formulation and parametrization. *J. Am. Chem. Soc.* **1972**, *94*, 5612–5625. [[CrossRef](#)]
2. Warshel, A.; Levitt, M. Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* **1976**, *103*, 227–249. [[CrossRef](#)]
3. Vreven, T.; Morokuma, K.; Farkas, O.; Schlegel, H.B.; Frisch, M.J. Geometry optimization with QM/MM, ONIOM, and other combined methods. I. Microiterations and constraints. *J. Comp. Chem.* **2003**, *24*, 760–769. [[CrossRef](#)] [[PubMed](#)]
4. Senn, H.M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem. Int. Ed.* **2009**, *48*, 1198–1229. [[CrossRef](#)] [[PubMed](#)]
5. Ahmadi, S.; Herrera, L.B.; Chehelamirani, M.; Hostas, J.; Jalife, S.; Salahub, D.R. Multiscale modeling of enzymes: QM-cluster, QM/MM, and QM/MM/MD: A tutorial review. *Int. J. Quantum Chem.* **2018**, *118*, e25558. [[CrossRef](#)]

6. Bash, P.A.; Field, M.J.; Davenport, R.C.; Petsko, G.A.; Ringe, D.; Karplus, M. Computer Simulation and Analysis of the Reaction Pathway of Triosephosphate Isomerase. *Biochemistry* **1991**, *30*, 5826–5832. [[CrossRef](#)] [[PubMed](#)]
7. Liao, R.-Z.; Thiel, W. Convergence in the QM-only and QM/MM modeling of enzymatic reactions: A case study for acetylene hydratase. *J. Comput. Chem.* **2013**, *34*, 2389–2397. [[CrossRef](#)] [[PubMed](#)]
8. Hu, L.; Söderhjelm, P.; Ryde, U. On the Convergence of QM/MM Energies. *J. Chem. Theory Comput.* **2011**, *7*, 761–777. [[CrossRef](#)] [[PubMed](#)]
9. Kulik, H.J.; Zhang, J.; Klinman, J.P.; Martínez, T.J. How Large Should the QM Region Be in QM/MM Calculations? The Case of Catechol O-Methyltransferase. *J. Phys. Chem. B* **2016**, *120*, 11381–11394. [[CrossRef](#)] [[PubMed](#)]
10. Karelina, M.; Kulik, H.J. Systematic Quantum Mechanical Region Determination in QM/MM Simulation. *J. Chem. Theory Comput.* **2017**, *13*, 563–576. [[CrossRef](#)] [[PubMed](#)]
11. Sumner, S.; Söderhjelm, P.; Ryde, U. Effect of Geometry Optimizations on QM-Cluster and QM/MM Studies of Reaction Energies in Proteins. *J. Chem. Theory Comput.* **2013**, *9*, 4205–4214. [[CrossRef](#)] [[PubMed](#)]
12. Jindal, G.; Warshel, A. Exploring the Dependence of QM/MM Calculations of Enzyme Catalysis on the Size of the QM Region. *J. Phys. Chem. B* **2016**, *120*, 9913–9921. [[CrossRef](#)] [[PubMed](#)]
13. Benediktsson, B.; Bjornsson, R. QM/MM Study of the Nitrogenase MoFe Protein Resting State: Broken-Symmetry States, Protonation States, and QM Region Convergence in the FeMoco Active Site. *Inorg. Chem.* **2017**, *56*, 13417–13429. [[CrossRef](#)] [[PubMed](#)]
14. Das, S.; Nam, K.; Major, D.T. Rapid Convergence of Energy and Free Energy Profiles with Quantum Mechanical Size in Quantum Mechanical–Molecular Mechanical Simulations of Proton Transfer in DNA. *J. Chem. Theory Comput.* **2018**, *14*, 1695–1705. [[CrossRef](#)] [[PubMed](#)]
15. Yang, Z.; Mehmood, R.; Wang, M.; Qi, H.W.; Steeves, A.H.; Kulik, H.J. Revealing quantum mechanical effects in enzyme catalysis with large-scale electronic structure simulation. *React. Chem. Eng.* **2019**, *4*, 298–315. [[CrossRef](#)]
16. Senet, P.J. Nonlinear electronic responses, Fukui functions and hardnesses as functionals of the ground state electronic density. *J. Chem. Phys.* **1996**, *105*, 6471–6489. [[CrossRef](#)]
17. Contreras, R.R.; Fuentealba, P.; Galván, M.; Pérez, P. A direct evaluation of regional Fukui functions in molecules. *Chem. Phys. Lett.* **1999**, *304*, 405–413. [[CrossRef](#)]
18. Yamanaka, S.; Yonezawa, Y.; Nakata, K.; Nishihara, S.; Okumura, M.; Takada, T.; Yamaguchi, K.; Nakamura, H. Locality and nonlocality of electronic structures of molecular systems: Toward QM/MM and QM/QM approaches. *AIP Conf. Proc.* **2012**, *1504*, 916–919.
19. Ueda, K.; Yamanaka, S.; Nakata, K.; Ehara, M.; Okumura, M.; Yamaguchi, K.; Nakamura, H. Linear response function approach for the boundary problem of QM/MM methods. *Int. J. Quantum Chem.* **2013**, *113*, 336–341. [[CrossRef](#)]
20. Mitsuta, Y.; Yamanaka, S.; Yamaguchi, K.; Okumura, M.; Nakamura, H. Theoretical Investigation on Nearsightedness of Finite Model and Molecular Systems based on Linear response analysis. *Molecules* **2014**, *19*, 13358–13373. [[CrossRef](#)] [[PubMed](#)]
21. Mitsuta, Y.; Yamanaka, S.; Saito, T.; Kawakami, K.; Yamaguchi, K.; Okumura, M.; Nakamura, H. Nearsightedness-related indices of finite systems based on linear response function: One-dimensional cases. *Mol. Phys.* **2016**, *114*, 380–388. [[CrossRef](#)]
22. Berkowitz, M.; Parr, R.G. Molecular hardness and softness, local hardness and softness, hardness and softness kernels, and relations among these quantities. *J. Chem. Phys.* **1988**, *88*, 2554–2557. [[CrossRef](#)]
23. Geerlings, P.; Proft, F.D.; Langenaeker, W. Conceptual density functional theory. *Chem. Rev.* **2003**, *103*, 1793–1873. [[CrossRef](#)] [[PubMed](#)]
24. Geerlings, P.; Flas, S.; Boisdenghien, Z.; De Proft, F. Conceptual DFT: Chemistry from the linear response function. *Chem. Soc. Rev.* **2014**, *43*, 4989–5008. [[CrossRef](#)] [[PubMed](#)]
25. Hohenberg, P.; Kohn, W. Inhomogeneous electron gas. *Phys. Rev.* **1964**, *136*, B864–B871. [[CrossRef](#)]
26. Kohn, W.; Sham, L.-J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **1965**, *140*, A1133–A1138. [[CrossRef](#)]
27. Prodan, E.; Kohn, W. Nearsightedness of electronic matter. *Proc. Natl. Acad. Sci. USA* **2005**, *112*, 11635–11638. [[CrossRef](#)] [[PubMed](#)]

28. *Handbook of Metalloprotein*; Messerschmidt, A.; Huber, R.; Poulos, T.; Wieghardt, K. (Eds.) John Wiley & Sons: New Jersey, NJ, USA, 2001; Volume 1.
29. Nakamura, A.; Ueyama, N.; Yamaguchi, K. *Organometallic Conjugation: Structures, Reactions and Functions of d-d and d- $\pi$  Conjugated Systems*; Springer: Berlin/Heidelberg, Germany; New York, NY, USA, 2003.
30. Itoh, K.; Kinoshita, M. (Eds.) *Molecular Magnetism*; Gordon & Breach: New York, NY, USA, 2000.
31. Parr, R.G.; Yang, W. Density Functional Theory of Atoms and Molecules. In *Horizons of Quantum Chemistry*; Springer, Dordrecht: New York, NY, USA, 1989.
32. Perdew, J.P.; Ernzerhof, M.; Burke, K.; Savin, A. On-top pair-density interpretation of spin density functional theory, with applications to magnetism. *Int. J. Quantum Chem.* **1997**, *61*, 197–205. [[CrossRef](#)]
33. Takeda, R.; Yamanaka, S.; Yamaguchi, K. CAS-DFT based on odd-electron density and radical density. *Chem. Phys. Lett.* **2002**, *366*, 321–328. [[CrossRef](#)]
34. Yamaguchi, K. *Self-Consistent Field Theory and Applications*; Carbo, R., Klobukowski, M., Eds.; Elsevier: Amsterdam, The Netherlands, 1990; pp. 727–757.
35. Yamanaka, S.; Kanda, K.; Saito, T.; Umena, Y.; Kawakami, K.; Shen, J.-R.; Kamiya, N.; Nakamura, H.; Yamaguchi, K. Electronic and spin structures of the  $\text{CaMn}_4\text{O}_5(\text{H}_2\text{O})_4$  cluster in OEC of PS II refined to 1.9 Å X-ray resolution. *Adv. Quantum Chem.* **2012**, *64*, 121–187.
36. Feynmann, R.P. Forces in molecules. *Phys. Rev.* **1939**, *56*, 340–343. [[CrossRef](#)]
37. Fias, S.; Boisdenghien, Z.; Proft, F.D.; Geerlings, P. The spin polarized linear response from density functional theory: Theory and application to atoms. *J. Chem. Phys.* **2014**, *141*, 184107. [[CrossRef](#)] [[PubMed](#)]
38. Yamanaka, S.; Kawakami, T.; Nagao, H.; Yamaguchi, K. Effective exchange integrals for open-shell species by density functional methods. *Chem. Phys. Lett.* **1994**, *231*, 25–33. [[CrossRef](#)]
39. Yamaguchi, K. General Molecular Orbital Theories of organic reaction mechanisms. *Chem. Phys.* **1978**, *29*, 117–139. [[CrossRef](#)]
40. Kishi, R.; Ochi, S.; Izumi, S.; Makino, A.; Nagami, T.; Fujiyoshi, J.; Nakano, N. Diradical Character Tuning for the Third-Order Nonlinear Optical Properties of Quinoidal Oligothiophenes by Introducing Thiophene-S,S-dioxide Rings. *Chem. Eur. J.* **2016**, *22*, 1493–1500. [[CrossRef](#)] [[PubMed](#)]
41. Schmidt, M.W.; Baldrige, K.K.; Boatz, J.A.; Elbert, S.T.; Gordon, M.S.; Jensen, J.H.; Koseki, S.; Matsunaga, N.; Nguyen, K.A.; Su, S.; et al. General atomic and molecular electronic structure system. *J. Comput. Chem.* **1993**, *14*, 1347–1363. [[CrossRef](#)]
42. Becke, A.D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652. [[CrossRef](#)]
43. Heme, T.L. Enzyme Structure and Function, Poulos. *Chem. Rev.* **2014**, *114*, 3919–3962.
44. Gober, J.G.; Ghodse, S.V.; Bogart, J.W.; Wever, W.J.; Watkins, R.R.; Brustad, E.M.; Bowers, A.A. P450-Mediated Non-natural Cyclopropanation of Dehydroalanine-Containing Thiopeptides. *ACS Chem. Biol.* **2017**, *12*, 1726–1731. [[CrossRef](#)] [[PubMed](#)]
45. Yamaguchi, K.; Kawakami, T.; Yamaki, D. Theory of Molecular Magnetism. In *Molecular Magnetism, Molecular Magnetism*; Gordon & Breach: New York, NY, USA, 2000; pp. 9–48.
46. Whittaker, J.W. Non-heme manganese catalase—the ‘other’ catalase. *Arch. Biochem. Biophys.* **2012**, *525*, 111–120. [[CrossRef](#)] [[PubMed](#)]
47. Holley, A.K.; Baktthavathalu, V.; Velez-Roman, J.M.; St. Clair, D.K. SOD Manganese Superoxide Dismutase: Guardian of the Powerhouse. *Int. J. Mol. Sci.* **2011**, *12*, 7114–7162. [[CrossRef](#)] [[PubMed](#)]
48. Umena, Y.; Kawakami, K.; Shen, J.R.; Kamiya, N. Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å. *Nature* **2011**, *473*, 55–60. [[CrossRef](#)] [[PubMed](#)]
49. Tanaka, A.; Fukushima, Y.; Kamiya, N. Two Different Structures of the Oxygen-Evolving Complex in the Same Polypeptide Frameworks of Photosystem II. *J. Am. Chem. Soc.* **2017**, *139*, 1718–1721, A chain (pdb entry: 5b66). [[CrossRef](#)] [[PubMed](#)]
50. Yamaguchi, K.; Shoji, M.; Isobe, H.; Yamanaka, S.; Umena, Y.; Kawakami, K.; Kamiya, N. On the guiding principles for understanding of geometrical structures of the  $\text{CaMn}_4\text{O}_5$  cluster in oxygen-evolving complex of photosystem II. Proposal of estimation formula of structural deformations via the Jahn–Teller effects. *Mol. Phys.* **2016**, *115*, 636–666. [[CrossRef](#)]

51. Shoji, M.; Isobe, H.; Tanaka, A.; Fukushima, Y.; Kawakami, K.; Umena, Y.; Kamiya, N.; Nakajima, T.; Yamaguchi, K. Understanding Two Different Structures in the Dark Stable State of the Oxygen-Evolving Complex of Photosystem II: Applicability of the Jahn–Teller Deformation Formula. *ChemPhotoChem* **2018**, *2*, 257–270. [[CrossRef](#)] [[PubMed](#)]
52. Siegbahn, P.E.M.; Blomberg, M.R.A. Density functional theory of biologically relevant metal centers. *Annu. Rev. Phys. Chem.* **1999**, *50*, 221–249. [[CrossRef](#)] [[PubMed](#)]
53. Yamanaka, S.; Kanda, K.; Saito, T.; Kitagawa, Y.; Kawakami, T.; Ehara, M.; Okumura, M.; Nakamura, H.; Yamaguchi, K. Does B3LYP correctly describe magnetism of manganese complexes. *Chem. Phys. Lett.* **2012**, *519–520*, 134–140. [[CrossRef](#)]
54. Eduardo, M.; Sproviero, E.M.; Gascon, J.A.; McEvoy, J.P.; Brudvig, G.W.; Batista, V.S. Characterization of synthetic oxomanganese complexes and the inorganic core of the O<sub>2</sub>-evolving complex in photosystemII: Evaluation of the DFT/B3LYP level of theory. *J. Inorg. Biochem.* **2006**, *100*, 786–800.
55. Yamanaka, S.; Kanda, K.; Saito, T.; Kitagawa, Y.; Kawakami, T.; Ehara, M.; Okumura, M.; Nakamura, H.; Yamaguchi, K. Density Functional Study of Manganese Complexes: Protonation Effects on Geometry and Magnetism. In *Progress in Theoretical Chemistry and Physics*; Kiyoshi Nishikawa, K., Maruani, J., Brändas, E.J., Delgado-Barrio, G., Piecuch, P., Eds.; Springer Business Media Dordrecht: New York, NY, USA, 2013; pp. 21–26.
56. Smith, F.J. Quadrature Methods based on the Euler-Maclaurin Formula and on the Clenshaw-Curtis Method of Integration. *Numer. Math.* **1965**, *7*, 406–411. [[CrossRef](#)]
57. Lebedev, V.I. Spherical quadrature formulas exact to orders 25–29. *Siberian Math. J.* **1977**, *18*, 99–107. [[CrossRef](#)]
58. Becke, A. A multicenter numerical integration scheme for polyatomic molecules. *J. Chem. Phys.* **1988**, *88*, 2547–2553. [[CrossRef](#)]
59. Wachters, J.H. Gaussian basis set for molecular wavefunctions containing third-row atoms. *J. Chem. Phys.* **1970**, *52*, 1033–1036. [[CrossRef](#)]
60. Bauschlicher, C.W., Jr.; Langhoff, S.R.; Barnes, L.A. Theoretical studies of the first-and second-row transition-metal methyls and their positive ions. *J. Chem. Phys.* **1989**, *91*, 2399–2411. [[CrossRef](#)]
61. Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Petersson, G.A.; Nakatsuji, H.; et al. GAUSSIAN 09, Revision C 0.1. Available online: [http://gaussian.com/g09\\_c01/](http://gaussian.com/g09_c01/) (accessed on 25 February 2018).
62. Bode, B.M.; Gordon, M.S. MacMolPlt: A graphical user interface for GAMESS. *J. Mol. Graphics Mod.* **1998**, *16*, 133–138. [[CrossRef](#)]
63. Mathematica Ver. 11, Wolfram Research: Champaign, IL, USA. Available online: <http://www.wolfram.com/mathematica> (accessed on 30 December 2018).
64. Singh, U.C.; Kollman, P.A. An approach to computing electrostatic charges for molecules. *J. Comput. Chem.* **1984**, *5*, 129–145. [[CrossRef](#)]

**Sample Availability:** Samples of the compounds not available.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).