Theoretical Study of Electron Transfer Mechanism in Biological Systems with a QM (MRSCI+DFT)/MM method

Toshikazu Takada

Research Program for Computational Science, RIKEN

2-1, Hirosawa, Wako, Saitama 351-0198, Japan

E-mail:tz-takada@riken.jp

Abstract. The goal of this project is to understand the charge separation mechanisms in biological systems using the molecular orbital theories. Specially, the charge separation in the photosynthetic reaction center is focused on, since the efficiency in use of the solar energy is extraordinary and the reason for it is still kept unknown. Here, a QM/MM theoretical scheme is employed to take the effects of the surrounding proteins onto the pigments into account. To describe such excited electronic structures, a unified theory by MRSCI and DFT is newly invented. For atoms in the MM space, a new sampling method has also been created, based on the statistical physics. By using these theoretical framework, the excited and positively charged states of the special pair, that is, chlorophyll dimmer are planning to be calculated this year.

Key words: Photosynthesis, Reaction Center, MRSCI, DFT, QM/MM, Excited Electronic Structures

1. Introduction

Photosynthesis has been one of the most challenging research fields for a long time. After the 3D structure of the reaction center of *R. viridis,* where the charge separation occurs by using solar energy, was firstly determined by Deisenhofer et al. [1], many of the detailed mechanisms of the charge separation have been revealed.

However, the reason for its high efficiency in use of the solar energy is still kept unknown and especially the following two scientific interests are still puzzled:

1) Although the pigments in the reaction center are arranged in the C_2 symmetry, the separation occurs only through the right hand side channel.

2) The distance between the special pair and accessory chlorophyll is about 10 A, but the charge transfer occurs within 3 pico seconds from the pair to peophytin through the accessory chlorophyll.

Figure 1 shows the arrangement of the pigments in the reaction center, showing the clear C2 symmetry. From the result in the previous calculation of the special pair, it is speculated that the electrostatic potential near the special pair generates such a field to push electrons from left to right in the center [2]. The electrostatic potential near the special pair is shown in **Figure 2**, which was obtained from the previous calculations. In the present study, the reason for the second puzzle is trying to be answered by carrying out the excited electronic structure calculation of these all pigments with a theoretical framework of QM/MM.
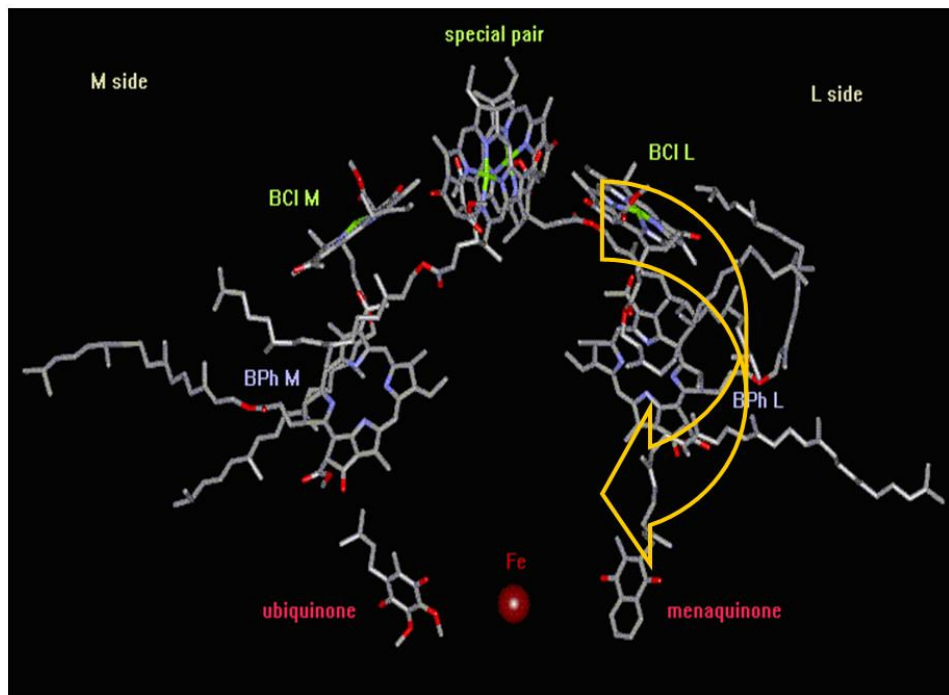
Figure 1. The arrangement of the pigments in the photosynthetic reaction center having the C_2 symmetry.
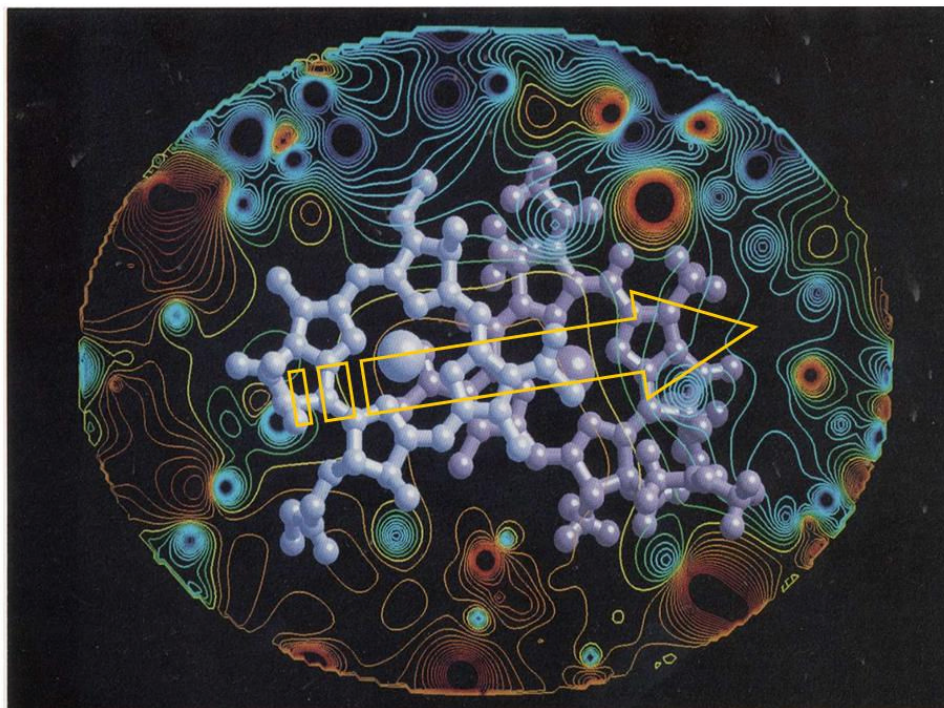

Figure 2. The electrostatic potential near the special pair created by the surrounding proteins. As shown by the arrow, the potential pushes electrons from left to right.

2. Method of Calculations

In the present study, a QM/MM theoretical framework is being adopted in order to take the effects onto the pigment electronic structures by the surrounding proteins, lipids and waters into account. Thus, the QM space which contains the six chlorophyll molecules is described with a hybrid approach of MRSCI and DFT. That is, non-dynamical electronic correlation is considered by the MRSCI scheme, which is a reliable way to describe complicated excited wave functions and dynamical one by the DFT framework. By combining these two theories, correct wave functions of the excited and charge transfer states are expected to obtain. However, one difficulty arises from this approach. That is, the dynamic correlation is doubly counted in both the MRSCI and DFT calculations, which must be subtracted from the total energy. A new hybrid theory has been developed here and the basic equation to remove the double count is derived as [3]:

$$\left\langle \Psi^{exact}\left|\hat{V}_{ee}\right|\Psi^{exact}\right\rangle \cong V_{clmb}\big[\rho(\mathbf{r})\big] + E_X^{MR}\big[\rho(\mathbf{r})\big]$$

$$+\frac{1}{2}\int\frac{d\mathbf{r}_1 d\mathbf{r}_2}{r_{12}}\rho_1(\mathbf{r}_1)\Big[\big(1-e^{-\alpha r_{12}}\big)\rho_C^{MR}(\mathbf{r}_1,\mathbf{r}_2)\Big] + C_{prefactor}E_C\big[\rho(\mathbf{r})\big]$$

There are further difficulties for developing a highly parallelized MRSCI code. One comes out from transformation of two electron integrals to molecular orbital sets on highly parallel computers that have been becoming a major trend for large scale molecular simulations. An efficient algorithm was already invented [4]. By using this algorithm, a CASSCF with 1000 basis functions has been realized in a complete direct fashion on a PC cluster of 128 nodes with a parallel ration of more than 99.9%. The previous benchmark is given in **Table 1**, showing surprising parallel performance on a PC cluster of 128 nodes. Therefore, a question is how to generate σ vectors by combining with this transformation algorithm in way of highly parallelized fashion. Another is diagonalization of CI matrixes, which will be extremely oblong ones. It is well known that Davison's method is really effective to diagonalize such matrixes, when only a few eigenvalues are requested [5].

As electronic structures of the pigments must be strongly influenced by the surrounding environment of proteins, lipids and water molecules, they must be treated by considering these effects. Therefore, a QM/MM theory is employed here, since QM/MM approaches are now recognized to be really efficient for describing large biological molecules like proteins.

Table 1. Speed up of the parallelized two electron integral transformation. This has been measured on a PC cluster of 128 nodes (Pentium4, RedHat Linux 6.2 and LAM/MPI 6.5.4). The left result was for parallelizing with only one of the four atomic orbital indexes and the right for two of them. The basis set used is MIDI-4 and the number of it is 608.

Proc	(*m*, *n*)	Time (sec)	speedup	ratio(%)	(*m*, *n*)	Time (sec)	speedup	ratio(%)
1	1,1	48434.930						
4	4,1	12281.310	3.944	99.525	1,4	12266.150	3.949	99.567
16	16,1	3115.090	15.548	99.806	4,4	3100.000	15.624	99.840
64	64,1	865.877	55.937	99.771	16,4	800.681	60.492	99.908
128	128,1	454.811	106.495	99.841	32,4	419.875	115.356	99.914

3. Results of Calculations

It was confirmed that the Davidson algorithm is quite efficient to solve such large CI matrixes, if computer memory is large enough to store the whole matrixes. Because of the shortage of the memory capacity of 2 GB, the efficiency was benchmarked up to only 30 millions on Intel P4 (2.53gHz). The CPU time for the largest case is found to be less than 10 seconds. From this result, it is expected that the diagonalization of CI matrixes will be negligible in comparison of generating σ vectors.

Development of a parallel code to generate σ vectors is presently underway and it will be soon combined together with the diagonalization package.

Now, the two codes of MRSCI and DFT are trying to be unified into one package and the excited energies of the special pair will be calculated as a probe to search for necessary computer capabilities for completing the six pigment calculations.

As a preparation for calculations of the excited and charge transfer states, molecular dynamical calculations of the pigments including proteins, lipid and water molecules have been performed to obtain their averaged atomic coordinates under influences of these substances. The picture of the reaction center using for the calculations is shown in **Figure 3.**
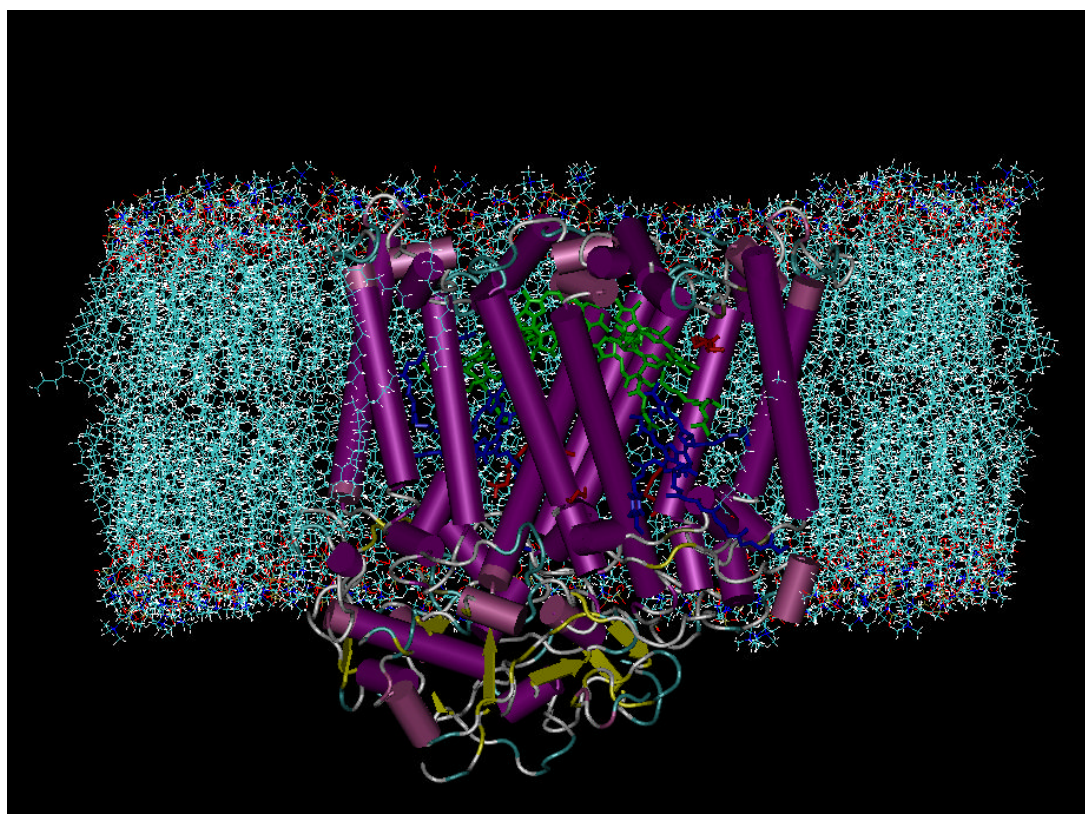

Figure 3. Generated molecular structures for the calculations of the excited and charge separated states of the pigments in order to take the influences from the surrounding substances into account. The rods are protein main chains and the pigments are shown in blue and green.

4. Conclusion

It is expected that, if the detailed mechanism of such an efficient charge separation in the photosynthesis is understood, useful and effective energy devices may be invented by employing a different principle from present silicon solar batteries. However, to achieve this, collaborations with experimentalists in the fields of photosynthesis, solid physics and devices is definitely needed.

This project is originally designed for the next generation supercomputer which is in progress of development at RIKEN with funding from the Japanese government, of which performances will be over peta scale. It is a desire that, through molecular simulations using the supercomputer, such collaboration becomes real in the near future in order to overcome energy problems that the human race will face soon.

References
[1]   J. Deisenhofer et al., Nature **318,** 618 (1985)
[2]   T. Sakuma et al., Int. J. Quant. Chem., **61**, 137(1997)
[3]   S. Yamanaka et al, Int. J. Quant. Chem., **106**, 3312 (2006)
[4]   K Nakata et al., J. Computational and Applied Mathematics**, 149**、351 (2002)
[5]   E. R. Davidson, J. Comput. Phys., **17**, 87 (1975)