QM/MM/3D-RISM study of an electronic structure change of benzimidazole derivative induced by molecular recognition of cucurbit[7]uril

Kodai Kanemaru,[12] Keisuke Shiroshita,[3] Haruyuki Nakano, [3] Norio Yoshida[2*]

[1] Fukui Institute for Fundamental Chemistry, Kyoto University, Takano-Nishibiraki-cho 34-4, Sakyo-ku, Kyoto 606-8103, Japan
[2] Department of Complex Systems Science, Graduate School of Informatics, Nagoya University, Furocho, Chikusa-ku, Nagoya 464-8601, Japan
[3] Department of Chemistry, Graduate School of Science, Kyushu University, Motooka 735, Nishi-ku, Fukuoka 819-0395, Japan

*Corresponding author: noriwo@nagoya-u.jp

Abstract. The quantum mechanics/molecular mechanics three-dimensional reference interaction site model (QM/MM/3D-RISM) method is a three-layer hybrid approach designed to calculate the electronic structure of complex molecules in solution. In this study, the QM/MM/3D-RISM method was employed to investigate electronic structure changes in ligand molecules upon molecular recognition by the macrocyclic host molecule cucurbit[7]uril (CB7). Thiabendazole (TBZ), a benzimidazole derivative, was selected as the ligand, which is known for its significant alterations in properties such as pK_a, solubility, photostability, and fluorescence intensity upon binding to CB7. The calculations elucidated the mechanism underlying the changes in the electronic structure induced by encapsulation, as well as the role of solvation in these transformations.

1. Introduction

Molecular recognition (MR) is a process where ligand molecules bind to specific sites on host molecules, such as proteins, inducing changes in molecular properties. This makes MR a valuable tool for tuning molecular functions. Among various host molecules, macrocyclic compounds like cucurbiturils (CBs) are particularly advantageous due to their lower toxicity, higher binding affinities, and easier synthesis compared to cyclodextrins. Structurally, CBs are oligomers of glycoluril monomers linked by methylene bridges, making them promising candidates for applications like drug delivery and self-healing materials. Benzimidazole derivatives, widely used in pharmaceuticals, show altered physicochemical properties, including pK_a, solubility, and fluorescence, upon encapsulation in CB7 [1].

Theoretical investigations of molecular properties in encapsulated systems require a methodology capable of addressing several challenges. First, since the changes in properties are fundamentally linked to alterations in the electronic structure of the ligand molecule, quantum chemical methods are essential. Second, the size of the macromolecular systems, such as CB7, necessitates an approach that can efficiently handle large molecules. Finally, given that MR occurs in solution, an accurate treatment of solvent effects is crucial.

In this study, the quantum mechanics/molecular mechanics three-dimensional reference interaction-site model (QM/MM/3D-RISM) method—a three-layered hybrid approach—was employed to address these requirements [2]. The QM/MM/3D-RISM method partitions the system into three regions: the chemically active QM region, the inactive MM model region, and the solvent which described by the 3D-RISM theory. In this study, we investigated the electronic structure changes in thiabendazole (TBZ) induced by MR with CB7 in aqueous solution. TBZ, as the ligand, was treated with the QM method, CB7 as the host is assigned to the MM region, and the water solvent was modeled using the 3D-RISM theory.

2. Theory

A brief explanation of the QM/MM/3D-RISM method is given here, as the details can be found in original paper.[2] In the QM/MM/3D-RISM formalism, the system consists of solute ligand and host molecules immersed in solvent at infinite dilution. The ligand electronic Hamiltonian is given by

$$\hat{H}(\boldsymbol{x},\boldsymbol{X}) = \hat{H}_0(\boldsymbol{x}) + \hat{V}_{\mathrm{MM}}(\boldsymbol{X}) + \hat{V}_{\mathrm{3DRISM}}(\boldsymbol{x},\boldsymbol{X})$$

where $\hat{H}_0(\boldsymbol{x})$ is the electronic Hamiltonian of ligand with the atomic coordinate $\boldsymbol{x}$ in isolated system. The potential acting on the ligand electron due to the host molecule $\hat{V}_{\mathrm{MM}}(\boldsymbol{X})$ and the potential acting on the ligand electron by the solvent molecules $\hat{V}_{\mathrm{3DRISM}}(\boldsymbol{x},\boldsymbol{X})$ are given as:

$$\hat{V}_{\mathrm{MM}}(\boldsymbol{X}) = -\sum_{i}^{\mathrm{MM}} \frac{q_i}{|\boldsymbol{X}_i - \boldsymbol{r}_e|}, \quad \hat{V}_{\mathrm{3DRISM}}(\boldsymbol{x},\boldsymbol{X}) = -\sum_{j}^{\mathrm{solvent}} \rho_j \int \frac{q_j^v}{|\boldsymbol{r} - \boldsymbol{r}_e|} g_j(\boldsymbol{r};\boldsymbol{x},\boldsymbol{X}) d\boldsymbol{r}$$

where $\boldsymbol{X}_i$ denotes the coordinate of MM atom i and $\boldsymbol{X}$ is the set of $\boldsymbol{X}_i$. q_i is the point charge on the MM atom i and $\boldsymbol{r}_e$ is the coordinate of electron. ρ_j and q_j^v are the number density and point charge of solvent site j. $g_j(\boldsymbol{r};\boldsymbol{x},\boldsymbol{X})$is the distribution function of solvent site v which is parametrically depends on QM and MM coordinates, $\boldsymbol{x},\boldsymbol{X}$. The distribution function should be obtained by solving the 3D-RISM equation coupled with the closure relation such as Kovalenko-Hirata closure with the solute-solvent interaction potential which depends on the solute electronic wave functions.

3. Computational details

3.1 Parameter preparation

In this study, we employed the QM/MM/3D-RISM method to evaluate the electronic structure of TBZ in solution. It is known that the TBZ takes protonated state, $TBZH^+$, depending pH. Therefore, in this study, we consider both TBZ and $TBZH^+$ as ligands. The initial structure of TBZ, $TBZH^+$, and CB7 were determined by the geometry optimization using Gaussian 16 with B3LYP/6-31G(d) level calculations.[4] The antechamber in AmberTools20 were employed for the parameter assignment.[5] The GAFF parameter set were used for these molecules and the RESP method was used to determine the point charges on solute atoms.[6] The TIP3P forcefield was employed for the solvent water molecules.[7]

3.2 Structure sampling

In prior to the QM/MM/3D-RISM calculation, the QM/MM simulations with explicit TIP3P water were conducted to structure sampling. The PM3 empirical quantum chemical method was applied for ligand TBZ and $TBZH^+$. The 4 ns production NPT MD simulations were conducted for each system with Berendsen temperature and pressure control at 300 K and 1 bar following 20 ps equilibration. The SHAKE method was applied to H atoms and 2 fs time step is used. A cubic solvation box filled with water molecules up to a distance of 15 Å from the solute molecules was prepared. The number of water molecules was about 1700 and 2600 for free ligands and CB7-ligand complex systems, respectively.

3.3 Electronic structure and solvation structure calculation

QM/MM/3D-RISM was performed for the solute structure extracted from the QM/MM simulation and striped all the explicit water at every 10 ps for 4 ns. Thus, 400 structures were used in QM/MM/3D-RISM. The B3LYP/6-31G(d) level was employed for the QM calculation. The density of solvent water was set to 0.99 g/cm^3 and 64 Å × 64 Å × 64 Å solvent box with 0.5 Å grid spacing was employed for

the 3D-RISM calculations. The Kovalenko-Hirata closure was used. All the QM/MM/3D-RISM calculations were conducted the RISMiCal program package combined with GAMESS quantum chemical program package.[8-11]

4. Results and discussion

4.1 Complex structure

Figure 1(a) and (b) depict the superimposed complex structures of TBZ-CB7 and $TBZH^+$-CB7. As shown, both TBZ and $TBZH^+$ are bound to CB7 through their benzimidazole groups, which is consistent with observations from NMR experiments [1]. Specifically, the nitrogen atoms in the benzimidazole ring form hydrogen bonds with the carbonyl oxygen atoms of CB7, highlighting the key interactions responsible for the binding.

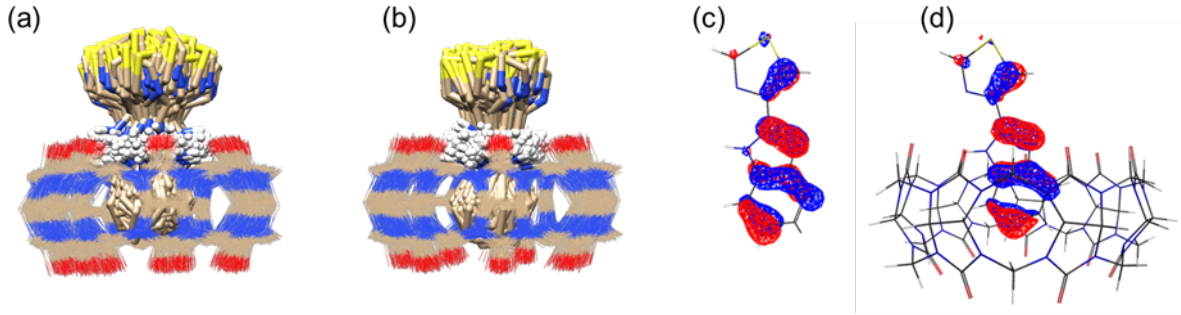

Figure 1. Complex structure of (a)TBZ and (b)$TBZH^+$ with the CB7, respectively. HOMO of TBZ in (c) free and (d) complex form, respectively.

4.2 Electronic structure

In this subsection, we discuss the electronic structure changes resulting from complex formation. Figure 1(c) and (d) compare the highest-occupied molecular orbital (HOMO) of TBZ in its free state and complexed with CB7, based on one of the snapshot structures. The HOMO primarily consists of π-bonds in the aromatic rings. The encapsulation by CB7 appears to have minimal impact on the shape of the HOMO.

Table 1 summarizes the averaged dipole moments and HOMO-LUMO gaps of TBZ and $TBZH^+$ in their free and complex states. For TBZ, the dipole moment undergoes a more significant change upon encapsulation compared to $TBZH^+$. Specifically, the dipole moment of TBZ increases by approximately 1 Debye due to encapsulation. In contrast, the HOMO-LUMO gap is less affected by encapsulation and more influenced by changes in the protonation state. These findings suggest that the changes in the absorption spectrum are likely driven by protonation effects rather than direct interactions with CB7 during encapsulation.

Table 1. The dipole moment and HOMO-LUMO gap of free and complex ligand

	TBZ		$TBZH^+$	
	Free	Complex	Free	Complex
Dipole moment / Debye	4.99	5.96	2.30	2.23
HOMO-LUMO gap/ eV	4.58	4.56	4.37	4.27

4.3 Solvation structure

In addition to the interaction with CB7, changes in the solvation structure may significantly influence the electronic structure of TBZ. Figure 2 illustrates the spatial distribution of water oxygen and hydrogen, represented by red- and gray-colored surfaces, respectively. For free TBZ, a prominent peak of hydrogen-bonded solvent oxygen is observed around the imide hydrogen of the benzimidazole group. Upon binding to CB7, this solvent oxygen is replaced by the carbonyl oxygen of CB7. A distinct peak of solvent hydrogen is also observed near the unprotonated imide group. Interestingly, this hydrogen

peak persists in the complex state, forming a bridge between the nitrogen atom of TBZ and the carbonyl oxygen of CB7. Furthermore, the hydrophobic aromatic rings of TBZ are encapsulated within CB7 and are partially dehydrated compared to their free state. For $TBZH^+$, both imide hydrogens exhibit peaks of hydrogen-bonded water oxygen in the free state, but these peaks disappear upon binding to CB7.

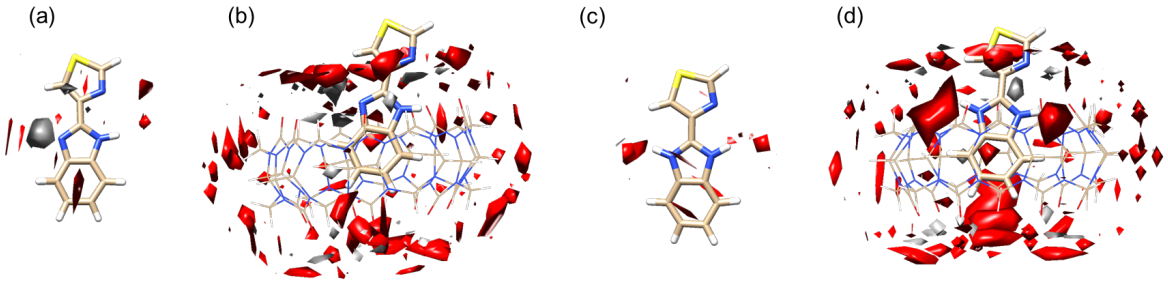

Figure 2. 3D distribution of solvent oxygen (in red, $g(\boldsymbol{r}) = 2.8$) and hydrogen (in gray, $g(\boldsymbol{r}) = 2.3$). (a) Free TBZ, (b) complex, CB7-TBZ, (c) free $TBZH^+$ and (d) complex CB7-TBZ^+

5. Conclusions

In this study, the electronic and solvation structures of thiabendazole (TBZ) and its protonated form ($TBZH^+$) in their free states and in complex with cucurbit[7]uril (CB7) are theoretically investigated using the QM/MM/3D-RISM method. While the results presented here are preliminary, they offer valuable insights into the changes in TBZ properties associated with molecular recognition. Future work will include a more comprehensive evaluation of computational parameters, such as functionals and basis sets, to enable a more detailed investigation.

Acknowledgement

This work was supported by MEXT as "Program for Promoting Researches on the Supercomputer Fugaku"(JPMXP1020230327) and used computational resources of supercomputer Fugaku (hp230212, hp240223) and the MEXT Program: Data Creation and Utilization-Type Material Research and Development Project Grant No. JPMXP1122714694. The computations were also performed using the Research Center for Computational Science, Institute for Molecular Science (23-IMS-C068 and 24-IMS-C066) and MCRP-S at the Center for Computational Sciences, University of Tsukuba. This work was also supported by the JSPS KAKENHI (22H05089, 24K01434, 22H01873, 24H00282) and the Toyota Physical and Chemical Research Institute scholar collaborative research project.

References

[1] Koner A L, Ghosh I, Saleh N and Nau W M 2011 *Can. J. Chem.* **89** 139
[2] Yoshida N, Kiyota Y and Hirata F 2011 *J. Mol. Liquids* **159** 83
[3] Kovalenko A and Hirata F 1999 *J. Chem. Phys.* **110** 10095
[4] Frisch M J et al. Gaussian 16, Rev. C.01, Gaussian, Inc., Wallingford CT, 2019
[5] Case D A et al., Amber 2020
[6] Wang J, Wolf R M, Caldwell J W, Kollman P A and Case D A, 2004 *J. Comput. Chem.* **25** 1157
[7] Jorgensen W L, Chandrasekhar J, Madura J D, Impey R W and Klein M L 1983 *J. Chem. Phys.* **79** 926
[8] Yoshida N and Hirata F 2006 *J. Comput. Chem.* **27** 453
[9] Yoshida N 2014 *J. Chem. Phys.* **140** 214118
[10] Maruyama Y and Yoshida N 2024 *J. Comput. Chem.* **45** 1470
[11] Barca G M J et al. 2020 *J. Chem. Phys.* **152** 154102