

Chapter 8

Computational Molecular Science Research Team

8.1 Members

Takahito Nakajima (Team Leader)
Noriyuki Minezawa (Research Scientist)
Takehiro Yonehara (Research Scientist)
William Dawson (Research Scientist)
Wataru Uemura (Postdoctoral Researcher)
Nobuki Inoue (Postdoctoral Researcher)
Takahide Matsuoka (Postdoctoral Researcher)
Subrata Tewary (Postdoctoral Researcher)
Eisuke Kawashima (Postdoctoral Researcher)
Kizashi Yamaguchi (Visiting Scientist)
Takashi Kawakami (Visiting Scientist)
Marek Janusz Wójcik (Visiting Scientist)

8.2 Overview of Research Activities

8.2.1 Development of Original Molecular Theory

An atomic- and molecular-level understanding of drug actions and the mechanisms of a variety of chemical reactions will provide insight for developing new drugs and materials. Although a number of diverse experimental methods have been developed, it still remains difficult to investigate the state of complex molecules and to follow chemical reactions in detail. Therefore, a theoretical molecular science that can predict the properties and functions of matter at the atomic and molecular levels by means of molecular theoretical calculations is keenly awaited as a replacement for experiment. Theoretical molecular science has recently made great strides due to progress in molecular theory and computer development. However, it is still unsatisfactory for practical applications. Consequently, our main goal is to realize an updated theoretical molecular science by developing a molecular theory and calculation methods to handle large complex molecules with high precision under a variety of conditions. To achieve our aim, we have so far developed several methods of calculation. Examples include a way for resolving a significant problem facing conventional methods of calculation, in which the calculation volume increases dramatically when dealing with larger molecules; a way for improving the precision of calculations in molecular simulations; and a way for high-precision calculation of the properties of molecules containing heavy atoms such as metal atoms.

8.2.2 Quantum Chemistry Software NTChem

Quantum chemistry software comprises immensely useful tools in material and biological science research. Widely diverse programs have been developed in Western countries as Japan has lagged. In fact, only a few programs have been developed in Japan. The mission of our research team is to provide K computer users with a high-performance software for quantum molecular simulation. In the early stage of the K computer project, no quantum chemistry software was available for general purpose and massively parallel computation on the K computer because not every program was designed for use on it. Therefore, we have chosen to develop a new comprehensive ab initio quantum chemistry software locally: NTChem. NTChem is completely new software that implements not only standard quantum chemistry approaches, but also original and improved theoretical methods that we have developed in our research work. The main features of the current version, NTChem2013, are the following:

1. Electronic structure calculation of the ground state of atoms and molecules based on Hartree–Fock (HF) and density functional theory (DFT) methods.
2. Linear-scaling or low-scaling DFT: Gaussian and finite-element Coulomb (GFC) resolution-of-the-identity (RI) DFT, pseudospectral DFT/HF, and dual-level DFT.
3. Low-scaling SCF calculation using diagonalization-free approaches: purification density matrix, pseudo-diagonalization, and quadratic convergence SCF.
4. Excited-state DFT calculation: time-dependent DFT (TDDFT) and transition potential (DFT-TP).
5. Accurate electron correlation methods for ground and excited states: Møller–Plesset perturbation theory, coupled-cluster (CC) theory, and quantum Monte Carlo (QMC) method.
6. Massively parallel computing on the K computer and Intel-based architectures: HF, DFT, resolution-of-the-identity second-order Møller–Plesset (RI-MP2) method, and QMC method.
7. Two-component relativistic electronic structure calculation with spin-orbit (SO) interactions: Douglas–Kroll (DK1, DK2 and DK3), zeroth- and infinite-order regular approximations (ZORA and IORA), and Relativistic scheme for Eliminating Small Components (RESC).
8. Model calculations for large molecular systems: quantum mechanics/molecular mechanics (QM/MM) and Our own N-layered Integrated molecular Orbital and molecular Mechanics (ONIOM).
9. Calculation of solvation effects: COnductor-like Screening MOdel (COSMO) (interfaced with the HONDO program), averaged solvent electrostatic potential/molecular dynamics (ASEP/MD), and QM/MM-MD.
10. Efficient calculation for chemical reaction pathway.
11. Ab initio molecular dynamics calculation.
12. Calculation of magnetic properties: nuclear magnetic resonance (NMR) chemical shifts, magnetizabilities, and electron paramagnetic resonance (EPR) g tensors.
13. Population analysis: Mulliken and natural bond orbital (NBO) analysis (interfaced with NBO 6.0).
14. Orbital interaction analysis: maximally interacting orbital (MIO) and paired interacting orbital (PIO) methods.

8.3 Research Results and Achievements

8.3.1 Trajectory Surface Hopping Molecular Dynamics Simulation by Spin-Flip Time-Dependent Density Functional Theory

This study presents the nonadiabatic molecular dynamics simulation combined with the spin-flip time-dependent density functional theory (SF-TDDFT). In contrast to the conventional single-reference electronic structure methods, which have difficulty in describing the S_0/S_1 conical intersections, the SF-TDDFT can yield the correct topology of crossing points. Thus, one expects that the method can take naturally into account the $S_1 \rightarrow S_0$ nonadiabatic transitions. We adopt Tully’s fewest switch surface hopping algorithm by introducing the

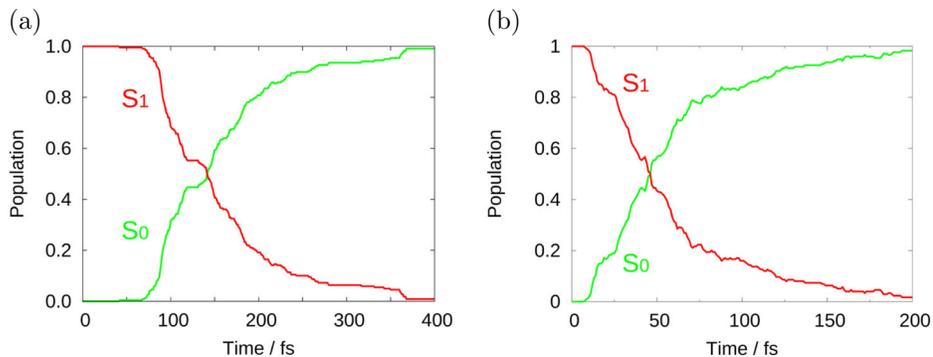


Figure 8.1: Time-dependent decay of (a) *E*-azomethane and (b) methanimine.

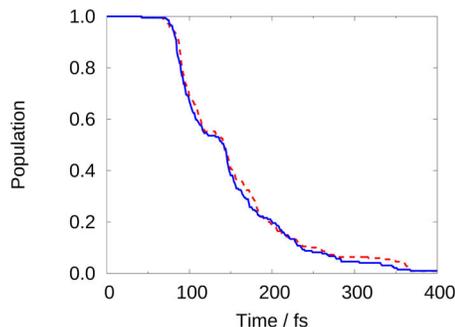


Figure 8.2: Time-dependent population decay of *E*-azomethane in vacuum (red dashed line) and water (blue solid line).

analytic SF-TDDFT nonadiabatic coupling vector. We apply the proposed method to the photoisomerization reactions of *E*-azomethane, methanimine, and ethene molecules and reproduce the results of previous studies based on the multireference methods (Figure 8.1). The proposed approach overcomes the ad hoc treatment of $S_1 \rightarrow S_0$ transition at the single-reference calculation level and affords both the dynamics on the S_1 state and the recovery of the S_0 state with modest computational costs.

8.3.2 Quantum Mechanical/Molecular Mechanical Trajectory Surface Hopping Molecular Dynamics Simulation by Spin-Flip Time-Dependent Density Functional Theory

This study presents the nonadiabatic molecular dynamics simulation in the solution phase using the spin-flip time-dependent density functional theory (SF-TDDFT). Despite the single-reference level of theory, the SF-TDDFT method can generate the correct topology of S_0/S_1 crossing points, thus providing a natural $S_1 \rightarrow S_0$ nonadiabatic transition. We extend the gas-phase trajectory surface hopping simulation with the SF-TDDFT [N. Minezawa and T. Nakajima, *J. Chem. Phys.* 150, 204120 (2019)] to the hybrid quantum mechanical/molecular mechanics (QM/MM) scheme. To this end, we modify the code to evaluate the electrostatic interaction between the quantum and molecular mechanical, respectively QM and MM, atoms and to extract the classical MM energy and forces from the MM program package. We apply the proposed method to the photoisomerization reaction of aqueous *E*-azomethane and anionic green fluorescent protein chromophore in water and compare the results with those of the previous simulation studies based on the multireference methods (Figure 8.2).

8.3.3 Electron dynamics method using a locally projected group diabatic Fock matrix for molecules and aggregates

We propose a method using reduced size of Hilbert space to describe an electron dynamics in molecule and aggregate based on our previous theoretical scheme (Yonehara and Nakajima, 2017). The real-time time-dependent density functional theory is combined with newly introduced projected group diabatic Fock matrix. First, this projection method is applied to a test donor-acceptor dimer, namely, a naphthalene-tetracyanoethylene with

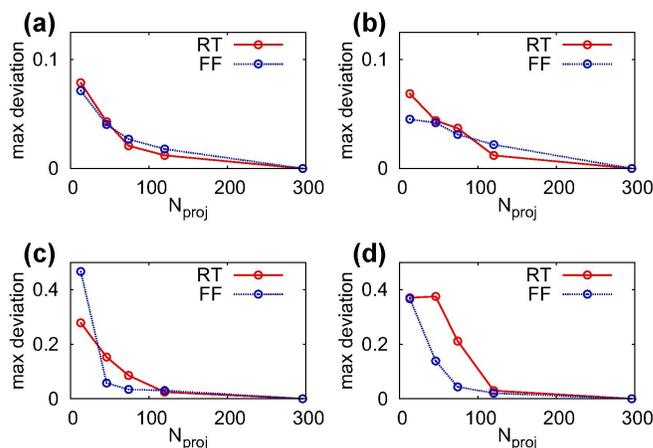


Figure 8.3: N_{proj} dependency of max deviations for donor molecule during dynamics from the results by full orbital calculation. Deviation is expressed as an absolute value but not relative one. The panels of (a), (b), (c) and (d) correspond to the results obtained by electron dynamics calculation using RT-TDDFT scheme considering the self-consistency and frozen Fock approximation.

and without initial local excitations and light fields. Secondly, we calculate an absorption spectrum of five-unit-polythiophene monomer. The importance of feedback of instantaneous density to Fock matrix is also clarified. In all cases, half of the orbitals were safely reduced without loss of accuracy in descriptions of properties (Figure 8.3). The present scheme provides one possible way to investigate and analyze a complex excited electron dynamics in molecular aggregates within a moderate computational cost.

8.3.4 Complexity Reduction in Density Functional Theory Calculations of Large Systems

The goal of this project is to develop new theory, algorithms, and workflows that enable the application of Density Functional Theory (DFT) to large systems. As we detailed in our recent review article, this will allow us to apply DFT to novel fields, such as the development of new enzymes for bioremediation. One of the challenges of large scale DFT is finding ways to extract chemical insight from complex systems. To this end, we have developed and evaluated a systematic complexity reduction framework which can break large systems down into fragments and quantify inter-fragmentation interaction.

To design our complexity reduction framework, we derived two measures from properties of the single particle density matrix: the fragment purity value and the fragment bond order. The purity value is a measure of fragment quality and the bond order of fragment interaction strength. We demonstrated how these two measures can be used to automatically partition a system into fragments at various levels of detail. At a suitable level of granularity, the observables of the full system can be reconstructed as a sum of fragment quantities. We further demonstrated this approach by using the fragment bond order to design QM/MM partitionings that preserve the observables of a target region (multipoles, forces, etc). Together, these two measures can be used to construct graph like views of molecular systems (Figure 8.4).

8.3.5 Molecular Design for Solar Cell Materials

Spiro-OMeTAD, a hole-transporting material (HTM) for perovskite solar cell, is known for its high power conversion efficiency (PCE). Its PCE for $(\text{FAPbI}_3)_{0.92}(\text{MAPbBr}_3)_{0.08}$ perovskite is 23.4%, and the estimated cost for synthesizing Spiro-OMeTAD is approximately \$274/g. Although materials with cheaper synthesizing cost and fairly high PCE have been previously reported (X60: \$120/g, Py-C: \$192/g, etc), more cost-effective materials with higher PCEs are yet to be searched.

We trained deep neural network (DNN) model, which predicts PCEs of HTMs with molecular descriptors provided as inputs. Furthermore, we conducted Bayesian optimization by evaluating the acquisition function with Gaussian process regression (GPR). We employed discrete particle swarm optimization (DPSO) method to optimize the vast chemical space.

We constructed learning models to predict the PCEs of candidate HTMs, generated from fragments of

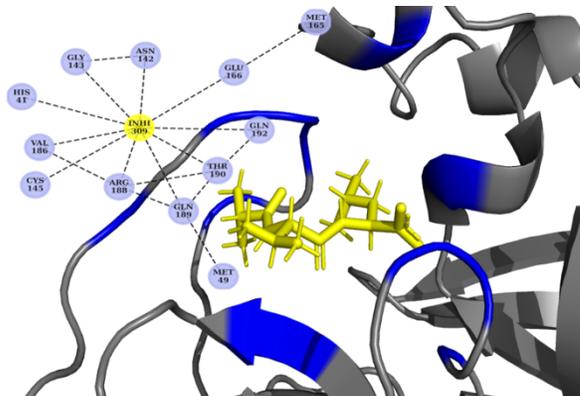


Figure 8.4: A graph view of the binding site of PDB:6lu7 based on the fragment bond order of protein residues with a potential inhibitor.

known HTMs. For training data, we collected 400 entries of experimental data of HTMs for perovskite solar cells. We employed HTMs (170 molecules), active layers (54 compounds, band gaps, valence band maximum, conduction band minimum), dopants, co-dopants, active area, and PCEs as input. 170 HTM molecules have been decomposed into three fragments, and the candidate HTMs have been generated from these fragments. The molecular descriptors have been calculated for each fragment with Mordred. We also prepared quantum descriptors (HOMO, LUMO, total energy, electronic energy, heat of formation, dispersion energy, dipole moment) using NTChem on K-computer. We constructed DNN model, which predicts PCEs, with combination of these descriptors, provided as input. The candidate HTMs have been chosen by GPR model with typical experimental conditions. As virtual experiment, the PCEs of the candidate HTMs have been predicted from the DNN model. The optimum candidates have been selected, while improving the GPR model by repeating the virtual experiment. Global optimization is unrealistic due to the vast chemical space (32,294,400 molecules). We searched for candidate HTMs with DPSO with the PCEs predicted by the constructed models, provided as the objective function. An $n \times N$ bit matrix has been defined as coordinates of the particles, where n is number of fragments species and N is number of fragments contained in candidate HTMs. The coordinates are updated probabilistically. Softmax function has been selected as probability function to select single fragment.

TiO₂ electron-transporting layer and (FAPbI₃)_{0.85}(MAPbBr₃)_{0.15} active layer has been employed as typical experimental conditions. The PCEs of SAF structure HTMs and SFX structure HTMs has been predicted to be high (Figure 8.5). Among the SFX structure HTMs, X55(20.8%) and X26(20.2%) are known to show high PCEs. Therefore, we expect the SFX structure candidate HTMs will also show high PCEs. The predicted PCEs of selected SAF structure candidates are over 20%, which greatly improves from the known SAF structure HTMs, such as SAF-OMe(16.7%) and CW4(16.6%).

8.3.6 Molecular Design for Highly-functional Biopolymer Materials

The functionalities of typical biopolymers, such as poly(lactic acid), PLA, and poly(butylene succinate), PBS, remain equivalent to the ones of general polymers. Highly-functional biopolymer development technology is considered to be a powerful solution for realizing a recycling-oriented society and economic revitalization toward the achievement of SDGs. For developing highly-functional biopolymers, it is necessary to establish the basic technology for estimating the polymer structure from the desired functionality. In order to establish a method for predicting properties from a given structure, we searched for polymers with high heat resistance. The learning model was constructed by using an experimental database and quantum chemistry calculation, and implemented a data-driven search method.

For the glass transition point (T_g) and melting point (T_m), experimental data of 497 types and 421 types of polymers were collected, respectively. Mordred was used to calculate the molecular descriptors determined from the topology and structure of the repeating unit molecules (1501 descriptors for T_g and 1432 descriptors for T_m). Furthermore, the molecule of the repeating unit was decomposed into two fragments, and the molecular descriptors were calculated for each fragment (1347 descriptors for T_g and 1286 descriptors for T_m). Additionally, HOMO, LUMO, total energy, total electron energy, heat of formation, and dipole moment were calculated using NTChem on K-computer and included as quantum descriptors. The base-level learning model was constructed with partial least squares regression (PLSR), support vector machine (SVM), and k -nearest neighbor algorithm

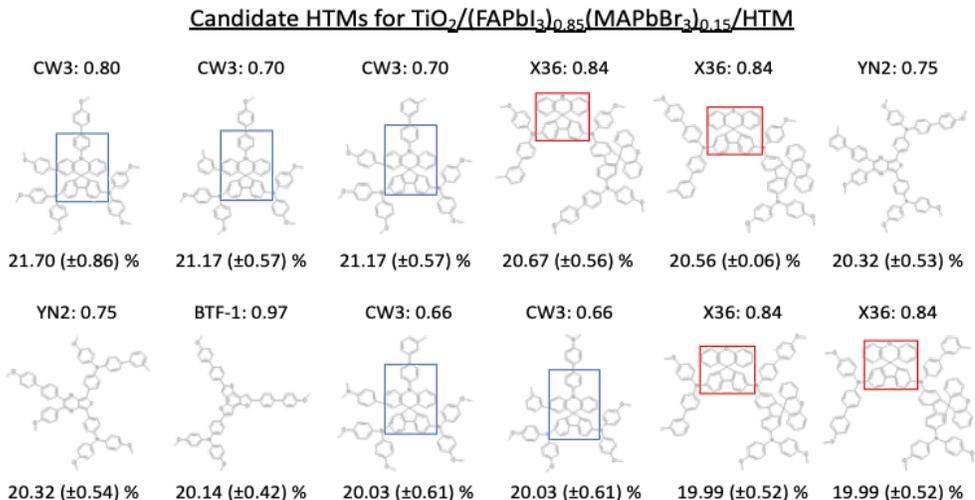


Figure 8.5: Hole-transporting material, HTM, candidates. Blue frames and red frames respectively indicate SAF structures and SFX structures.

(k -NN), and a stacking model was constructed with DNN as meta-level model. We constructed 10 DNN learning models by K -fold cross-validation method. The above descriptors were calculated for all the molecules in the biosynthesizable monomer database (283 types of diols, 71 types of dicarboxylic acids, 307 types of diamines).

T_g and T_m of 57,072 polyesters (combination of 283 diols and 71 dicarboxylic acids) and 18,573 polyamides (combination of 307 diamines and 71 dicarboxylic acids) were predicted from the learning model. As a result, 42 polyesters with predicted T_g of 150°C or higher and 391 polyamides of 200°C or higher were found, and 139 polyesters and 157 polyamides with predicted T_m of 300°C or higher were found (Figure ?? and ??). We found that aromatic polymers are promising as candidates for high heat resistant polymers. Although the biosynthesizable monomer database did not contain the para-dicarboxylic acid, T_g and T_m of the polymer synthesized from the para-dicarboxylic acid were predicted to be high.

8.3.7 Construction of Molecular Orbital Method Consistent with Quantum Electrodynamics

The relativistic molecular orbital (MO) method is commonly used today in the calculation of systems containing heavy elements. However, the relativistic quantum mechanics, which is the basic theory of the relativistic molecular orbital method, is incomplete in the treatment of antiparticles, and this point is the same in the relativistic molecular orbital method. To overcome this problem, the molecular orbital method should be reorganized based on quantum electrodynamics (QED), which is a “true” relativistic extension of quantum mechanics, and the relativistic MO method is extended to the QED-MO method. In this way, by rewriting the expression of quantum electrodynamics into a form that has a high affinity with the MO method, not only the theoretical rigor is increased, but also the quantum electrodynamics calculation of multi-electron systems can be executed by the program of the MO method. Therefore, in this study, we show that frequency-dependent Breit interaction

$$V^{\text{QED}}(r; \omega) = -\frac{\boldsymbol{\alpha} \cdot \boldsymbol{\alpha}}{4\pi\epsilon_0} \frac{e^{i\omega r/c}}{r} - \frac{(\boldsymbol{\alpha} \cdot \nabla)(\boldsymbol{\alpha} \cdot \nabla')}{4\pi\epsilon_0} \frac{1 - e^{i\omega r/c}}{(\omega/c)^2 r} \quad (8.1)$$

is derived from the QED Hamiltonian of based on the first principle, and tried to formulate and implement the QED-MO method using this interaction.

First, we show that the frequency-dependent Breit interaction is derived from the QED Hamiltonian in the first principle. This interaction represents interactions between two electrons by a photon that transports energy ω and can be obtained by applying Schrieffer–Wolff transformation to QED Hamiltonian of interaction representation. By this transformation, we can get the effective Hamiltonian (four-fermion(4f)-QED Hamiltonian)

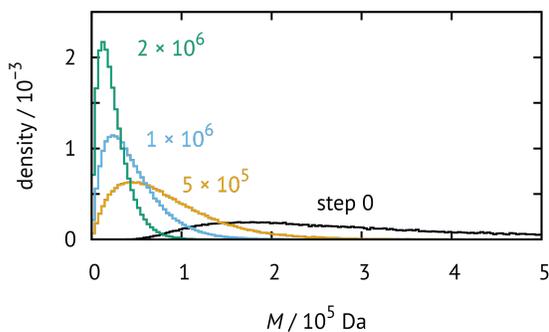


Figure 8.6: Molecular weight distribution of 10^5 poly(lactic acid) molecules under degradation. Molecular weights, M , at initialization follow a log-normal distribution; number- and weight-averaged molecular weights are 2.56×10^5 and 5.22×10^5 , respectively.

which has no photon external lines:

$$\hat{H}_2^{4f\text{-QED}} = \frac{1}{2} \sum_{ijkl} [ij | V^{\text{QED}}(r; \varepsilon_l - \varepsilon_k) | kl] \hat{b}_i^\dagger \hat{b}_k^\dagger \hat{b}_l \hat{b}_j. \quad (8.2)$$

Here, i, j, k and l are the indices of the molecular orbital, and \hat{b}^\dagger and \hat{b} are the electron creation and annihilation operators, respectively. This Hamiltonian depends on the orbital energy, and the Dirac–Hartree–Fock calculation using this Hamiltonian has a large calculation costs. For this reason, we treat it perturbatively, in other words, add the 4f-QED Hamiltonian after determining the orbitals by usual Dirac–Hartree–Fock calculations. 4f-QED Hamiltonian can be basically used as it is for the calculation of electron correlation method by adding it to ordinary two-electron Hamiltonian (Coulomb interaction). That is, the effect of QED can be calculated using the calculation program of the electron correlation method. That is, the QED-MO calculation can be realized if the two-electron integral of frequency-dependent Breit interaction can be evaluated. It is already known that the Obara–Saika method can be applied to two-electron integrals for arbitrary potentials in Cartesian Gaussian basis functions. However, it is necessary to newly create “molecular incomplete gamma function” for each potential. We have succeeded to develop methods for evaluating the “molecular incomplete gamma function” for frequency-dependent Breit interaction, which means that we provided the almost all tools necessary for QED-MO calculations.

8.3.8 Multi-Scale Simulation to Predict Biodegradability of Plastics

Biodegradable plastics are attracting attention to reduce environmental impact and achieve Sustainable Development Goals, SDGs. Biodegradable aliphatic polyesters—polymers or copolymers of hydroxyalkanoates, or copolymers of diols and dicarboxylic acids—eventually decompose into carbon dioxide and water by microbial metabolism. Examples of current commercial biodegradable polyesters are poly(lactic acid), PLA, poly(caprolactone), PCL, and poly(3-hydroxybutyrate-co-3-hydroxyhexanoate), PHBH. Skeleton structures of these materials determine mechanical and thermal properties and biodegradability, which restricts their application.

We are developing theoretical tools to estimate biodegradability of plastics to offer design guideline, by applying materials informatics. Density functional theory, DFT, calculations are employed as implemented in NTChem. Methylmethanoate, $\text{CH}_3\text{COOCH}_3$, was chosen as a model ester to guess structures for others. Reaction pathways of ester hydrolysis under base condition are investigated by nudged elastic band, NEB, and string method. Calculations are performed on HPC clusters including K computer, and activation energies and heats of reaction are obtained. A macroscopic Monte Carlo simulator is also implemented to estimate changes of molecular weight distributions under degradation. At initialization, polymers, whose molecular weights follow a probability distribution such as log-normal and inverse-gamma, are populated to reproduce experimental number- and mass-average molecular weights, and an ester bond is randomly cleaved at each Monte Carlo step (Figure 8.6).

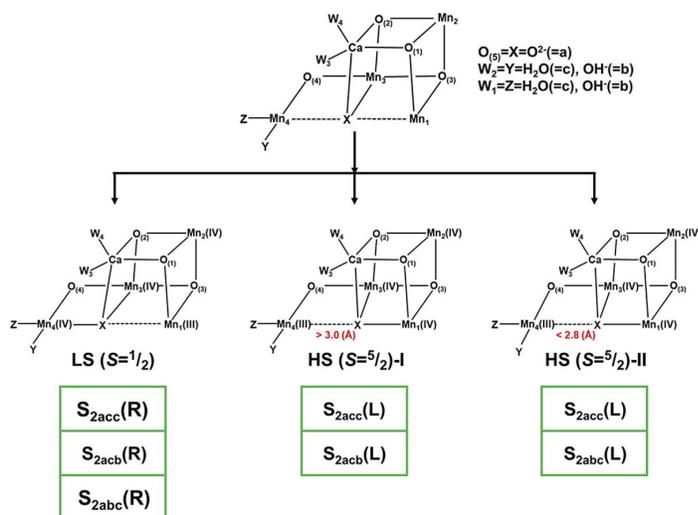


Figure 8.7: Structure of the $CaMn_4O_5$ cluster in oxygen evolving cycle of PSII. Three different S_2 structures, one R-opened LS ($S = 1/2$) and two L-opened HS ($S = 5/2$) structures are optimized by UB3LYP/TZVP methods.

8.3.9 Theoretical and computational investigations of geometrical, electronic and spin structures of the $CaMn_4O_X$ ($X=5, 6$) cluster in the Kok cycle S_i ($i = 0-3$) of oxygen evolving complex of photosystem II

The optimized geometries of the $CaMn_4O_X$ ($X = 5, 6$) cluster in the oxygen evolving complex (OEC) of photosystem II (PSII) by large-scale quantum mechanics (QM) and molecular mechanics (MM) calculations are compared with recent serial femtosecond crystallography (SFX) results for the S_i ($i = 0-3$) states. The valence states of four Mn ions by the QM/MM calculations are also examined in relation to the experimental results by the X-ray emission spectroscopy (XES) for the S_i intermediates. Geometrical and valence structures of right-opened Mn-hydroxide, Mn-oxo and Mn-peroxide intermediates in the S_3 state are investigated in detail in relation to recent SFX and XES experiments for the S_3 state (Figure 8.7). Interplay between theory and experiment indicates that the Mn-oxo intermediate is a new possible candidate for the S_3 state. Implications of the computational results are discussed in relation to possible mechanisms of the oxygen-oxygen bond formation for water oxidation in OEC of PSII.

8.3.10 Domain-Based Local Pair Natural Orbital CCSD(T) Calculations of Strongly Correlated Electron Systems: Examination of Dynamic Equilibrium Models Based on Multiple Intermediates in S_1 State of Photosystem II

Domain-based local pair natural orbital (DLPNO) coupled cluster single and double (CCSD) methods with perturbative triples (T) correction with NormalPNO were used to compute energies for twelve different S_1 structures of the $CaMn_4O_5$ cluster in the oxygen evolving complex (OEC) of photosystem II (PSII). The DLPNO-CCSD(T_0) calculations with TightPNO for the important six structures among them revealed that the right (R)-opened S_{1XYZW} structures were more stable than the corresponding left (L)-opened structures ($X = O_{(5)}$, $Y = W_2$, $Z = W_1$, and $W = O_{(4)}$) of $CaMn_4O_5$. The three different S_1 structures belonging to the R-opened type (S_{1acca} , S_{1bbca} , and S_{1abcb} , where $O^{2-} = a$, $OH^- = b$ and $H_2O = c$) were found nearly degenerated in energy, indicating the possibility of the coexistence of different structures in the S_1 state. The DLPNO-CCSD(T_0) calculations with TightPNO supported the proposal of a dynamic equilibrium model based on the multi-intermediate structures for the S_1 state, which is also in agreement with EPR and other experimental and hybrid DFT computational results. Implications of the computational results are discussed in relation to scope and applicability of NormalPNO and TightPNO for the CCSD(T_0) calculations of strongly correlated electron systems such as 3d transition-metal complexes.

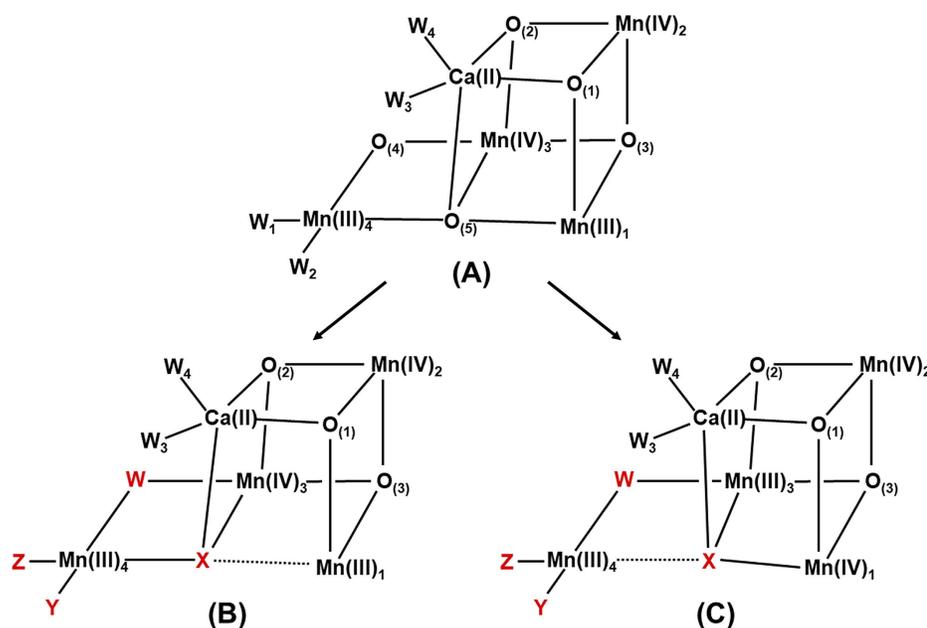


Figure 8.8: (A) Geometrical structure of the CaMn_4O_5 cluster by the high-resolution XRD structure with central symmetry, (B) right (R)-opened optimized S_1 structure and (C) left (L)-opened optimized S_1 structure. The protonation states of water molecule in the S_1 state are denoted as S_{1XYZ} where X, Y and Z are O^{2-} (=a), OH^- (=b) or H_2O (=c). Full geometry optimizations of six different structures; R- and L- S_{1aca} , R- and L- S_{1bba} .

8.3.11 Domain-Based Local Pair Natural Orbital CCSD(T) Calculations of Six Different S_1 Structures of Oxygen Evolving Complex of Photosystem II. Proposal of Multi-Intermediate Models for the S_1 State

Domain-based local pair natural orbital (DLPNO) coupled cluster single and double (CCSD) with triple perturbation (T) correction methods were applied for six different S_1 structures of oxygen evolving complex (OEC) of photosystem II (PSII), showing that right-opened three S_1 structures (Figure 8.8) were nearly degenerated in energy. The DLPNO-CCSD(T0) calculations support proposals of the multi-intermediate models for the S_1 state in accord with the EPR, other experimental and DFT computational results.

8.3.12 Domain-Based Local Pair Natural Orbital CCSD(T) Calculations of Fourteen Different S_2 Intermediates for Water Oxidation in the Kok Cycle of OEC of PSII. Re-visit to One LS-two HS Model for the S_2 State

Domain-based local pair natural orbital (DLPNO) coupled cluster single and double (CCSD) with triple perturbation (T) correction methods were applied for fourteen different S_2 structures of the CaMn_4O_5 cluster in oxygen evolving complex (OEC) of photosystem II (PSII). The DLPNO-CCSD(T0) calculations elucidated that the right (R)-opened S_{2aYZ} structure (a = O^{2-} at the $\text{O}_{(5)}$ site, Y = W2 and Z = W1) with the low spin (LS) ($S = 1/2$, $g = 2$) state and two left (L)-opened S_{2aYZ} structures with the high spin (HS) ($S = 5/2$, $g = 4$; $g > 4$) state were nearly degenerated in energy, supporting previous one LS-two HS model for the S_2 state in compatible with recent EXAFS and EPR results.

8.3.13 Quantum-Mechanical Study of Energies, Structures and Vibrational Spectra of the HF Complexed with Dimethyl Ether

Interaction energies, geometry and vibrational frequencies of the gas-phase HF-dimethyl ether complex were obtained using quantum-chemical methods. Equilibrium and vibrationally averaged geometries, harmonic and anharmonic wavenumbers of the complex were calculated using second-order perturbation theory procedures with B3LYP, B2PLYP-D and MP2 methods with 6-311++G(2df, 2pd) basis set (Figure 8.9). Quantum-mechanical model describing anharmonic-type vibrational couplings within hydrogen bond was used to explain broadening, fine structure and temperature dependence of the F-H stretching IR absorption bands as effect of hydrogen

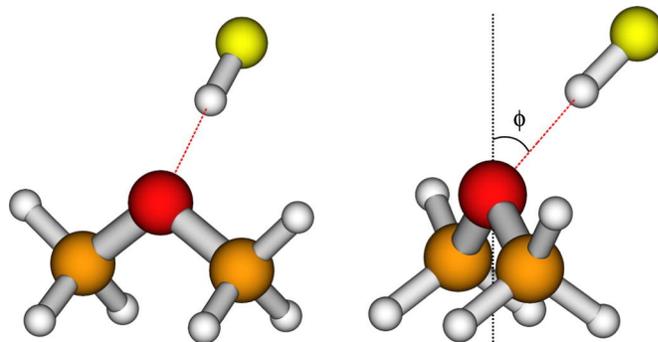


Figure 8.9: Equilibrium structure of the DME:HF complex with the C_s symmetry, optimized at the MP2/6-311++G(2df, 2pd) level, and the definition of the angle ϕ between the O-H direction and the bisector of the C-O-C angle coplanar with this angle.

bond formation. Simulations of the rovibrational structure of the F-H stretching bands were performed for different temperatures. The results were compared with experimental spectra.

8.3.14 IR Spectra of Crystalline Nucleobases. Combination of Periodic Harmonic Calculations with Anharmonic Corrections Based on Finite Models

This work reports a theoretical study of infrared (IR) spectra of four nucleobases (adenine, cytosine, guanine, and thymine) in the crystalline state. The effects responsible for the fine spectral features were revealed, and the nonfundamental bands significantly contributing to the IR fingerprint region were successfully reproduced. Additionally, we compared the fundamental bands simulated for periodic models in harmonic approximation with the results obtained for finite models in anharmonic approximation. On this basis, we concluded that accurate description of the chemical neighborhood is more essential for the IR fingerprint region than the anharmonicity (Figure 8.10). Comparison with previous results indicates that the vibrational properties and the nature of intermolecular interactions of nucleobases in the crystalline state remain similar to those in solution. Therefore, the conclusions obtained for well-defined crystalline structures of nucleobases are general and helpful in understanding the vibrational spectra and properties of nucleobases and their derivatives. Finally, this work evidences that anharmonic force field based on finite models may be applied as an inexpensive correction to the harmonic spectrum of an infinite periodic system.

8.3.15 A Comparison of the Hydrogen Bond Interactions Dynamics in the Guanine and Cytosine Crystals: Ab Initio Molecular Dynamics and Spectroscopic Study

In this work, we present the comparison study of guanine and cytosine crystals based on the hydrogen bond (HB) dynamics. The ab initio molecular dynamics gave us a base for detailed analysis. The analysis of the trajectories by power spectrum generation, as well as the fluctuation of the interaction energies, showed large differences between HB networks in the considered crystals. The charge flow is present in the guanine molecule which forms the flat surfaces in the crystals. In the cytosine zigzag structure, the charge flow is blocked (Figure 8.11). The interaction energy is significantly less stabilizing in the cytosine structure than in the guanine. Finally, the possible influence of charge transfer on the melting temperature has been discussed.

8.3.16 Conversion Reaction of Polyoxometalates from Anderson Structure to Keggin Structure

Aiming at developing synthetic methods of new Keggin-type polyoxometalates (POMs) that can mediate proton-conjugated multi-electron transfer reactions useful for efficient regeneration of fuels, we have investigated the reaction routes from Anderson-type POMs to Keggin-type POMs. The POM systems with various kinds of heteroatoms as their central cations were calculated using the Nudged Elastic Band (NEB) method as well as the first-principles electronic structure method (Figure 8.12). The effects of the heteroatoms on the reaction routes have been discussed.

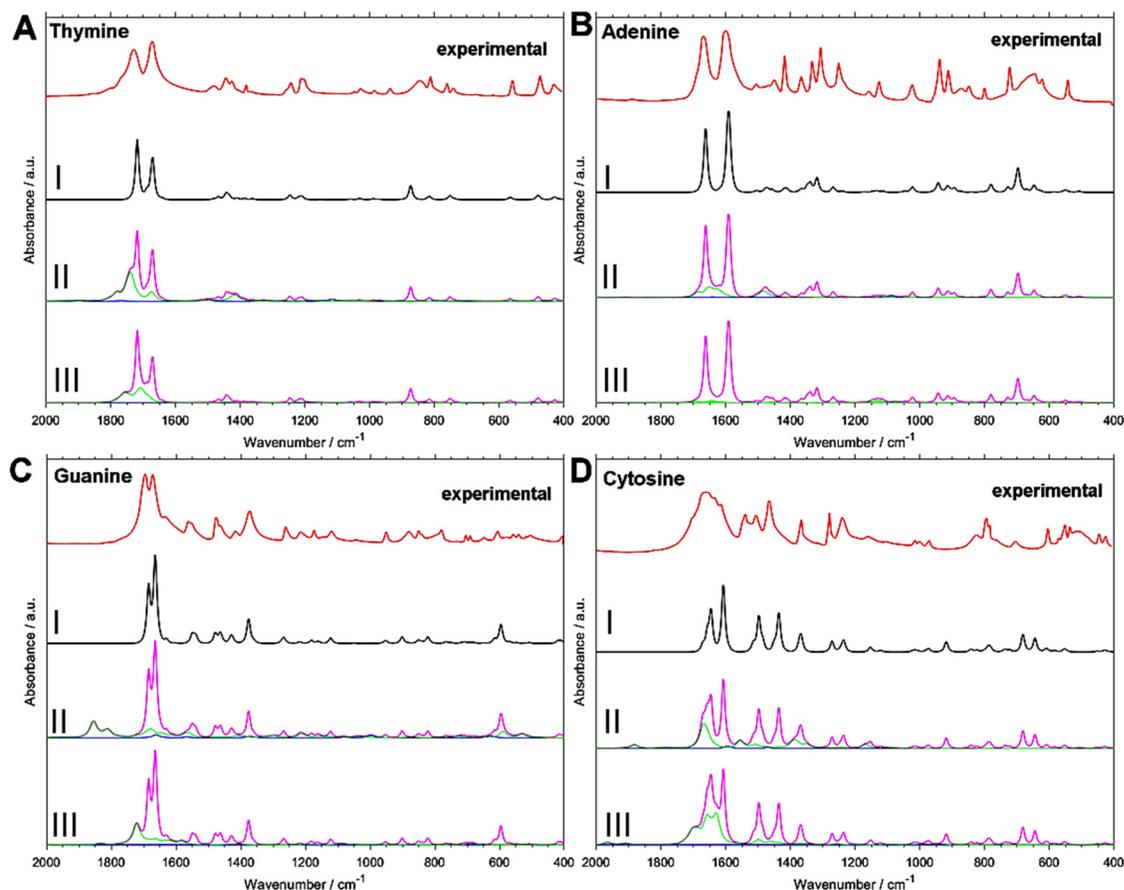


Figure 8.10: Experimental and simulated IR spectra of nucleobases in the fingerprint region ($2000\text{--}400\text{ cm}^{-1}$). The harmonic spectrum (pd./B3LYP/Gatti) in the 3D periodic system (I) and after additional anharmonic correction (II–III); II: DVPT2//B3LYP-GD3BJ/6-31++G(d,p) monomer; III: DVPT2//B3LYP-GD3BJ/6-31G(d,p) dimer. Green line, contribution from binary combinations; blue line, contribution from first overtones.

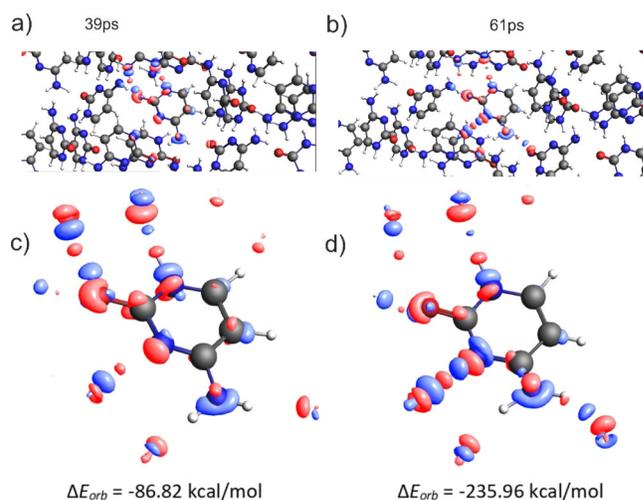


Figure 8.11: Analysis of deformation densities for the cluster of 27 molecules of cytosine. The blue and red colors show the accumulation and depletion of the electron density, respectively. Panel (a) shows the structure after 39 ps of simulation, while panel (b) after 61 ps. Panels (c) and (d) present the magnification of deformation density around the considered molecule in both snapshots, respectively.

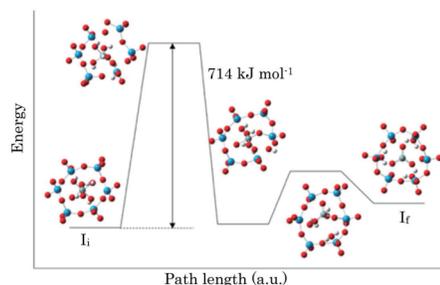


Figure 8.12: Exploration results of the conversion path between Anderson-type and Keggin-type six-membered rings I_i and I_f , for $X = \text{Si}$.

8.4 Schedule and Future Plan

8.4.1 Complexity Reduction in Density Functional Theory Calculations of Large Systems

In the following year, we plan to apply this framework to systems of practical interest including enzymes for bioremediation and important proteins of SARS-Cov-2. We further plan to use this framework in close collaboration with our newly developed version of NTChem aimed at computing large systems. In the previous year, we have worked on improving the robustness and ease of use of this new version of NTChem. Using the fragmentation information generated by our new framework, we hope to create accurate initial guesses for NTChem, further reducing the time to solution of hybrid DFT calculations.

8.4.2 Molecular Design for Solar Cell Materials

We found novel HTMs candidates by employing Bayesian optimization and DPSO. We will further expand our method by including quantum chemical calculation and improve our prediction model. Furthermore, we will improve our method by accelerating the Bayesian optimization process.

8.4.3 Molecular Design for Highly-functional Biopolymer Materials

Currently, we are investigating the experimental feasibility of synthesizing the candidate polymers. We will further improve our database by repeatedly reflecting the experimental results in simulation and informatics.

8.4.4 Construction of Molecular Orbital Method Consistent with Quantum Electrodynamics

The scheme of the QED-MO method developed this study has a high theoretical rigor, but has a demerit that the calculation cost is large. In the future, we will develop a program that can execute this QED-MO method stably and at low cost.

8.4.5 Multi-Scale Simulation to Predict Biodegradability of Plastics

The current MC simulator is not able to consider morphological effects and to evaluate real-time change in molecular weight distribution. We will implement Kinetic Monte Carlo, KMC, simulator to address above, which uses rates calculated by DFT and estimates time change of populations. We are also developing models to predict these values from monomer structures.

8.5 Publications

8.5.1 Articles/Journal

[1] K. Yamaguchi, S. Yamanaka, H. Isobe, M. Shoji, K. Miyagawa, T. Nakajima, T. Kawakami, and M. Okumura, "Theoretical and computational investigations of geometrical, electronic and spin structures of the CaMn_4O_X

- (X=5, 6) cluster in the Kok cycle S_i ($i = 0-3$) of oxygen evolving complex of photosystem II,” *Physiol. Plant.* 166, 44–59 (2019).
- [2] N. Minezawa and T. Nakajima, “Trajectory surface hopping molecular dynamics simulation by spin-flip time-dependent density functional theory,” *J. Chem. Phys.* 150, 204120 (2019).
- [3] L. Boda, M. Boczar, M. Z. Brela, M. J. Wojcik, and T. Nakajima, “Quantum-mechanical study of energies, structures and vibrational spectra of the HF complexed with dimethyl ether,” *Chem. Phys. Lett.* 731, 136590 (2019).
- [4] K. Miyagawa, T. Kawakami, Y. Suzuki, H. Isobe, M. Shoji, S. Yamanaka, M. Okumura, T. Nakajima, and K. Yamaguchi, “Domain-based local pair natural orbital CCSD(T) calculations of strongly correlated electron systems: Examination of dynamic equilibrium models based on multiple intermediates in S_1 state of photosystem II,” *Mol. Phys.* 1666171 (2019).
- [5] K. Miyagawa, T. Kawakami, H. Isobe, M. Shoji, S. Yamanaka, K. Nakatani, M. Okumura, T. Nakajima, and K. Yamaguchi, “Domain-based local pair natural orbital CCSD(T) calculations of six different S_1 structures of oxygen evolving complex of photosystem II. Proposal of multi-intermediate models for the S_1 state,” *Chem. Phys. Lett.* 732, 136660 (2019).
- [6] K. Miyagawa, H. Isobe, T. Kawakami, M. Shoji, S. Yamanaka, M. Okumura, T. Nakajima, and K. Yamaguchi, “Domain-based local pair natural orbital CCSD(T) calculations of fourteen different S_2 intermediates for water oxidation in the Kok cycle of OEC of PSII. Re-visit to one LS-two HS model for the S_2 state,” *Chem. Phys. Lett.* 734, 136731 (2019).
- [7] K. Bec, J. Grabska, M. Czarnecki, C. Huck, M. Wójcik, T. Nakajima, and Y. Ozaki, “IR spectra of crystalline nucleobases. Combination of periodic harmonic calculations with anharmonic corrections based on finite models,” *J. Phys. Chem. B*, 123, 10001–10013 (2019).
- [8] M. Brela, O. Klimas, E. Surmiak, M. Boczar, T. Nakajima, and M. Wójcik, “A comparison of the hydrogen bond interactions dynamics in the guanine and cytosine crystals: Ab initio molecular dynamics and spectroscopic study,” *J. Phys. Chem. A*, 123, 10757–10763 (2019).
- [9] T. Yonehara and T. Nakajima, “Electron dynamics method using a locally projected group diabatic Fock matrix for molecules and aggregates,” *Chem. Phys.* 528, 110508 (2020).
- [10] N. Minezawa and T. Nakajima, “Quantum mechanical/molecular mechanical trajectory surface hopping molecular dynamics simulation by spin-flip time-dependent density functional theory,” *J. Chem. Phys.* 152, 024119 (2020).
- [11] T. Yamakawa, K. Eda, T. Osakai, and T. Nakajima, “Conversion reaction of polyoxometalates from Anderson structure to Keggin structure,” *J. Comput. Chem. Japan* 18, 257–258 (2020).

8.5.2 Posters

- [12] 米原丈博, 中嶋隆人, “局所透熱表示と射影演算子を用いた自己参照型非線形電子動力学の効率的解法,” 第22回理論化学討論会, 札幌, 2019-05-27.
- [13] 井上頌基, 中嶋隆人, “Dirac-Fock 演算子の Douglas-Kroll 変換および IOTC 変換,” 第22回理論化学討論会, 札幌, 2019-05-28.
- [14] 川上貴資, 宮川晃一, 鈴木雄太, 磯辺寛, 庄司光男, 山中秀介, 奥村光隆, 中嶋隆人, 山口兆, “光合成酵素発生錯体 (CaMn_4O_5 クラスタ) の S_1 状態での各構造と電子状態の LPNO-CC 法による解析,” 第13回分子科学討論会, 名古屋, 2019-09-17.
- [15] 神谷宗明, William Dawson, 中嶋隆人, “大規模分子の励起電子状態に対する時間依存密度汎関数法の開発,” 第13回分子科学討論会, 名古屋, 2019-09-17.
- [16] 嶺澤範行, 中嶋隆人, “太陽電池モデル分子の非断熱分子動力学シミュレーション,” 第13回分子科学討論会, 名古屋, 2019-09-18.
- [17] Takehiro Yonehara and Takahito Nakajima, “Computational study on quantum dynamics of excited electrons in molecular aggregates: Toward an efficient control of a conversion from light energy to chemical functionality,” The 2nd R-CCS International Symposium, Kobe, 2020-02-17.
- [18] William Dawson, Luigi Genovese and Takahito Nakajima, “Complexity Reduction in Density Functional Theory Calculations of Large Systems,” The 2nd R-CCS International Symposium, Kobe, 2020-02-17.
- [19] Nobuki Inoue and Takahito Nakajima, “Efficient calculation method considering charge distribution of nuclei in electronic structure theory,” The 2nd R-CCS International Symposium, Kobe, 2020-02-17.
- [20] Eisuke Kawashima and Takahito Nakajima, “Multi-Scale Simulation to Predict Biodegradability of Plastics,” The 2nd R-CCS International Symposium, Kobe, 2020-02-17.

8.5.3 Invited Talks

- [21] 中嶋隆人, “第一原理計算とスーパーコンピュータによる新材料設計,” 神戸大学開拓プロジェクト「階層縦断的アプローチによる革新的光エネルギー変換系の開拓」第1回シンポジウム, 神戸, 2019-04-16.
- [22] 中嶋隆人, “スーパーコンピュータを利用した第一原理シミュレーションによる新材料設計,” 新化学技術推進協会先端化学・材料技術部会講演会, 東京, 2019-05-10.
- [23] 中嶋隆人, “京を利用した第一原理計算とマテリアルズ・インフォマティクスによる新材料設計,” 高分子同友会勉強会, 東京, 2019-07-31.
- [24] T. Nakajima, T. Matsuoka, “Materials design of hole-transporting materials for perovskite solar cells,” Asia-Pacific Conference on Theoretical & Computational Chemistry 2019, Sydney, 2019-10-03.
- [25] 中嶋隆人, “シミュレーションとインフォマティクスの融合による新材料設計 -富岳に向けて-,” 第5回キャタリストインフォマティクスシンポジウム, 東京, 2019-11-12.
- [26] 中嶋隆人, “驚くべきスパコンの世界-その活用と未来展望-,” 量子科学技術研究開発機構学術情報流通講座, 稲毛, 2019-11-19.
- [27] 中嶋隆人, “分析・解析技術としての量子化学シミュレーションとインフォマティクスNMRを中心に(第一原理シミュレーションとマテリアルズ・インフォマティクスによる新材料設計),” 理研シンポジウム第20回 分析・解析技術と化学の最先端, 和光, 2019-12-11.
- [28] 中嶋隆人, “スーパーコンピュータを利用した分子シミュレーションとインフォマティクスによる新材料設計,” “AIと有機合成化学” 第4回公開講演会, 早稲田, 2020-02-03.

8.5.4 Oral Talks

- [29] 川嶋英佑, 中嶋隆人, “Virtual Screening of Non-Fullerene Acceptors for Organic Photovoltaics by Density Functional Theory and Dynamic Monte Carlo Method,” 重点課題5 第2回若手勉強会 (若手研究者によるワークショップ), 犬山, 2019-08-29.
- [30] Takahito Nakajima, “Molecular simulation on the K computer and towards Fugaku,” The 2nd R-CCS International Symposium, Kobe, 2020-02-17.