Integration of Quantum Chemistry and Machine Learning for a Deeper Understanding and Rational Design of Functional Materials

Division of Advanced Molecular Science
(Department of Theoretical and Computational Molecular Science, Theoretical Molecular Science II)

HATANAKA, Miho
Associate Professor
(Cross Appointment)
[hatanaka@ims.ac.jp]

Education
2002 B.S. Keio University
2008 Ph.D. Keio University

Professional Employment
2009 Research Associate (fixed term/research incentive), Keio University
2011 Fukui fellow, Kyoto University
2015 Assistant Professor, Kindai university
2015 Japan Science and Technology Agency (JST) PRESTO Researcher
2017 Associate Professor, Nara Institute of Science and Technology
2020 Associate Professor, Keio University
2024 Associate Professor (Cross Appointment), Institute for Molecular Science

Awards
2017 11th PCCP Prize
2019 12th Young Scientist Award of the Japan Society for Molecular Science
2021 The Chemical Society of Japan Award for Outstanding Young Women Chemists for 2021
2021 MEXT National Institute of Science and Technology Policy (NISTEP), NISTEP Selection (The Researchers with Nice Step) 2021

Member
Secretary
MASUDA, Michiko

Keywords Materials Informatics, Lanthanide, Computational Chemistry

We develop computational methods, apply them to elucidate mechanisms, and design materials using data science techniques such as machine learning (ML). Currently, our research focuses on the following subjects:

(1) Lanthanide Photofunctional Materials

Lanthanide (Ln) luminescence, originating from 4f–4f transitions, has been applied in various optical materials. While the 4f–4f emission wavelengths of Ln complexes are almost independent of the surroundings, the emission intensities are heavily influenced, making the design of appropriate ligand crucial. To get deeper insights into Ln luminescence, we proposed the energy shift method, which has contributed to the understanding and design of various Ln materials and has also been applied to non-Ln systems. However, it could not handle 4f–5d and charge transfer excited states, which are particularly important for cerium and europium complexes. To address this limitation, we are developing an updated method. In addition, we have constructed a database of cerium complexes containing 1.7k geometries and electronic properties. We plan to make this database public and further develop it into a platform for optical function prediction *via* ML.

(2) Transition Metal Catalysts

Transition metal (TM) complexes play a crucial role in organic synthesis, catalyzing a wide variety of chemical reactions. The catalytic abilities of TM complexes can be finely tuned by manipulating ligand-induced electronic and steric effects, which are key to controlling the reactivity and selectivity of specific reactions. In recent years, however, chemical products are being produced in smaller quantities and with greater variety, reducing the time available to study the synthesis conditions for each product. To address this challenge, we are constructing a database of TM complexes with various organophosphorus ligands, ranging from monodentate to multidentate skeletons. For descriptors applicable to different skeletons, we propose using the reaction energies of elementary reactions, such as oxidative addition and transmetallation. Our goal is to accelerate catalyst design by developing an ML model where our descriptors serve as explanatory variables and experimental catalytic activity results are used as objective variables.

Selected Publications

- S. Gocho, H. Nakamura, Q. Gao, T. Kobayashi, T. Inagaki and *M. Hatanaka, "Excited State Calculations Using Variational Quantum Eigensolver with Spin-Restricted Ansätze and Automatically-Adjusted Constraints," *npj Comput. Mater.* **9**, 13 (2023).
- X.-F. Wei, T. Wakaki, T. Itoh, H.-L. Li, T. Yoshimura, A. Miyazaki, K. Oisaki, *M. Hatanaka, *Y. Shimizu and *M. Kanai, "Catalytic Regio- and Enantioselective Proton Migration from Skipped Enynes to Allenes," *Chem* **5**, 585 (2019).
- *M. Hatanaka, Y. Hirai, Y. Kitagawa, T. Nakanishi, *Y. Hasegawa and *K. Morokuma, "Organic Linkers Control the Thermosensitivity of the Emission Intensities from Tb(III) and Eu(III) in a Chameleon Polymer," *Chem. Sci.* **8**, 423 (2017).

1. Geometry Exploration Using Bayesian Optimization

Conventional molecular geometry searches on a potential energy surface utilize energies and energy gradients from quantum chemical calculations. However, replacing energy calculations with noisy quantum computer measurements generates errors in the energies, which makes geometry optimization using the energy gradient difficult. One gradient-free optimization method that can potentially solve this problem is Bayesian optimization. To use Bayesian optimization in geometry search, a suitable acquisition function must be defined. In this study, we propose a strategy for geometry searches using Bayesian optimization and examine the appropriate acquisition functions to explore the global minimum (GM) and the most stable conical intersection (CI). The acquisition functions for the GM and CI searches were defined as the probability improvements of the energy with opposite sign and the cost function C in eq (1), respectively.

$$C(\boldsymbol{x}) = -\left\{\frac{E_{S0}(\boldsymbol{x}) + E_{S1}(\boldsymbol{x})}{2} + \frac{\{E_{S0}(\boldsymbol{x}) - E_{S1}(\boldsymbol{x})\}^2}{\alpha}\right\} \quad \cdots (1)$$

Here, $E_{Sn}(x)$ and α are the electronic energy of the n-the singlet state at the geometry x and the hyperparameter, respectively. We applied our strategy to two molecules, formaldehyde and ethylene, and found the GM and the most stable CI geometries with high accuracy for both molecules.

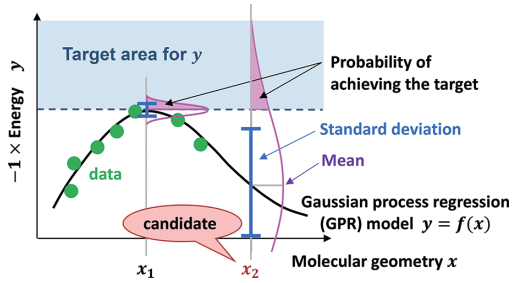

Figure 1. Schematic illustration of global minimum search using Bayesian optimization.

2. Copolymer Descriptor Database and Its Application to Machine Learning

Polymer informatics, which applies data-driven science to polymers, has attracted significant research interest. However, developing suitable descriptors for polymers, particularly copolymers, to enable machine learning models with limited data sets remains a challenge. To address this, we computed a range of parameters, including reaction energies and activation barriers of elementary reactions in the early stage of radical polymerization, for 2,500 radical–monomer pairs derived from 50 commercially available monomers. We then created an open-access database called the "Copolymer Descriptor Database (CopDDB)." Additionally, we constructed machine learning models using these descriptors as explanatory variables and various physical properties—such as the reactivity ratios, monomer conversion rates, monomer composition ratios, and molecular weights—as objective variables and achieved high predictive accuracies. We also conducted process optimization for the copolymerization of untested monomers. Typical Bayesian optimization requires an initial training set of target molecules, which entails significant experimental cost. However, our approach overcomes this challenge by using data from previously tested molecules that do not contain the target molecules for the initial training dataset. These models achieved high predictive accuracy, demonstrating the potential of our descriptors to advance polymer informatics.

3. Mechanism of Formation of Thermally Stable Cyclotrisilanyl Radicals

A halogen-substituted cyclic disilene compound, bromocyclotrisilene, $Si_3Br(Eind)_3$ (**3a**) bearing the fused-ring bulky Eind groups, was synthesized as an extraordinarily air-stable compound by our collaborators (Prof. Matsuo group at Kindai university). Even though **3a** is quite air-stable both in solutions and in the solid state, its Si–Br bond is reactive under reducing conditions. The further treatment of **3a** with Li metal leads to the formation of room-temperature thermally-stable silicon homologues of the cyclopropyl radical, *i.e.*, the cyclotrisilanyl radicals (**6a**), via intramolecular C–H bond activation in a transient silicon homologue of the cyclopropenyl radical, *i.e.*, the cyclotrisilenyl radical, $[Si_3(Eind)_3]^{\bullet}$ (**5a**). To elucidate the formation mechanism of **6a** from **5a**, we performed the reaction path search using the artificial force induced reaction (AFIR) method at the ONIOM (B3LYP-D3/PM6-D3) level. The reaction started from the approach of silyl radical center of **5a** to one of the H atoms in the Eind group, which afforded carbon-centered radical specie. The carbon-centered radical attacked Si_3, which induced the formation of C–Si and formed a ring-opened radical species. Then silylene center approaches the Si–H bond, which induces the formation of the Si–Si and Si–H bonds and results in the formation of the stable product **6a**. To the best of our knowledge, derivatives of the cyclopropyl radical have not yet been isolated and remained elusive. The cyclotrisilanyl radicals **6a** are the first room-temperature thermally stable homologues of the cyclopropyl radical, whose structural features and electronic properties have been experimentally and theoretically investigated.

References

1) R. Ohno, K. Ota, N. Nishimura, K. Taniguchi, S. Kurokawa, T. Wakabayashi, M. Hatanaka, A. Rosas-Sánchez, D. Hashizume and T. Matsuo, *J. Am. Chem. Soc.* **146**, 24911–24924 (2024).
2) K. Muto, M. Hatanaka, F. Kakiuchi and T. Kochi, *Organometallics* **43**, 1528–1535 (2024).
3) S. Kanno, H. Nakamura, T. Kobayashi, S. Gocho, M. Hatanaka, N. Yamamoto and Q. Gao, *npj Quantum Inf.* **10**, 56 (2024).
4) A. Wakiuchi, S. Jasial, S. Asano, R. Hashizume, M. Hatanaka, Y. Ohnishi, T. Matsubara, H. Ajiro, T. Sugawara, M. Fujii and T. Miyao, *MRS Commun.* **14**, 439–444 (2024).
5) K. Muto, M. Hatanaka, F. Kakiuchi and T. Kochi, *J. Org. Chem.* **89**, 4712–4721 (2024).
6) K. Terasaka, T. Kamoshida, T. Ichikawa, T. Yokoyama, M. Shibuta, M. Hatanaka and A. Nakajima, *J. Am. Chem. Soc.* **146**, 9605–9613 (2024).
7) T. Inoue, M. Hatanaka and A. Nakajima, *J. Am. Chem. Soc.* **145**, 23088–23097 (2023).