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
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
2025	Professor, Keio University
2025	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
CHIBA, Fumika

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. Ion Energy Shift (IES) Method for Exploring the Reaction Coordinate of Nonradiative Decay from Highly Excited State of Lanthanide Complexes

The exploration of minimum energy crossing points (MEXs) between potential energy surfaces (PESs) is essential for deepening our understanding of nonradiative decay mechanisms and plays a key role in the design of photofunctional materials. However, in lanthanide (Ln^{3+}) complexes, the presence of open-shell 4f electrons leads to quasi-degenerate electronic states, making MEX searches particularly challenging. To address this, the energy shift (ES) method has been used, in which the PES of a 4f–4f excited state is approximated by vertically shifting the ground-state PES by the corresponding 4f–4f excitation energy. This approach eliminates the need for explicit treatment of $4f^N$ electrons, enabling the use of large-core relativistic effective core potential (RECP) and conventional DFT methods to compute relevant states and locate key MEXs. While the ES method has greatly contributed to understanding and designing materials based on 4f–4f luminescence, it cannot be applied to 4f–5d or charge-transfer excited states (*i.e.*, $4f^{N-1}X$ states). To overcome this limitation, we proposed a new approximation, the ion energy shift (IES) method, which described the $4f^{N-1}X$ excited state using the large-core RECP for Ln^{4+} and shifted its PES to match the target excitation energy. The validity of the IES method was demonstrated by comparing its results with those of the multistate second-order perturbation theory (MSCASPT2). We applied it to elucidate the origin of the different excited-state lifetimes of hydrated Ce^{3+} complexes with and without coordination of a carboxylate ligand.

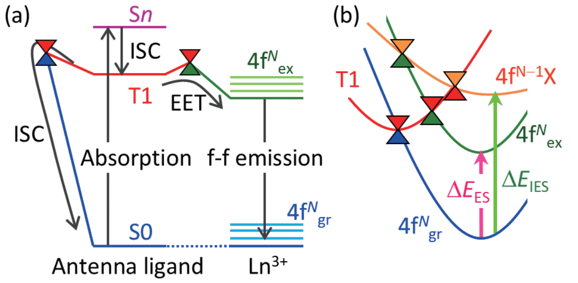

Figure 1. The Jablonski diagram of Ln^{3+} complex (a) and Schematic illustration of the ES and IES methods (b). The PES of the $4f^N$ excited states is described by shifting that of the ground state with the energy shift value of ΔE_{ES}. The PES of the $4f^{N-1}X$ excited state, in which $4f^{N-1}$ electrons were included in the large-core RECP for Ln^{4+}, is shifted to reproduce its excitation energy (ΔE_{IES}).

2. Database of Cerium Complexes for Excited State Analysis and Data-Driven Ligand Design

Since the early 2010s, the concept of data-driven materials discovery, commonly referred to as materials informatics, has gained significant traction. In response, a variety of databases compiling quantum mechanical (QM) calculations of materials have been developed. Large-scale QM datasets for molecules have also been constructed; however, due to the extreme diversity of chemical space, the field remains in a developing stage with respect to data diversity and extensibility. In particular, Ln complexes, which have broad industrial applications, pose challenges for establishing design guidelines because of their distinct structures, bonding characteristics, and physical properties. There is thus an urgent need for a database that captures the structural and electronic diversity of Ln complexes. To address this, we constructed a database of experimentally reported Ln complexes, including their optimized geometries and ground- and excited-state properties computed using DFT and TDDFT methods. In this initial study, we focus on Ce^{3+} complexes, which have recently attracted attention as photocatalysts. We collected 1635 Ce^{3+} complexes from the Cambridge Structural Database and identified key substructures that enhance excitation energy and oscillator strength by building machine learning models.

3. Mechanistic Study of C-Glycosylation Catalyzed by GgCGT Enzyme

C-glycosyltransferases have garnered attention owing to their ability to synthesize C-glycosides with high conversion and selectivity in one-pot reactions. Their potential in rational enzyme engineering makes them valuable for the synthesis of diverse C-glycosides. However, the detailed reaction mechanism remains unclear. To address this, we investigated the C-glycosylation of a polyphenol (phloretin) catalyzed by the glycosyltransferase GgCGT in the presence of the coenzyme UDP-glucose. Using the DFT calculations on a cluster model, we identified the most favorable pathway for C-glycosylation. The reaction proceeds *via* an initial proton transfer from phloretin to UDP-glucose, followed by the nucleophilic attack of phloretin on the glucose moiety and subsequent dissociation of UDP in an S_N2-like manner. The S_N2 step yields non-aromatic intermediate, which can be rapidly converted to C-glycoside even without an enzymatic environment. The key residue that facilitates the rate-determining S_N2 step is His-27, which stabilizes phloretin *via* hydrogen bonding. Additionally, to clarify why alternative products such as O-glycosides are not formed, we also investigated the O-glycosylation pathway. Our calculations revealed that O-glycosylation was promoted by proton transfer from UDP-glucose, like C-glycosylation, but was suppressed by structural fixation due to hydrogen bonding among phloretin, glucose, and GgCGT.

References

1) D. Terada, T. Inagaki and M. Hatanaka, *RSC Adv.* **15**, 28592–28600 (2025).
2) T. Yoshimura, H. Kato, S. Oikawa, T. Inagaki, S. Asano, T. Sugawara, T. Miyao, T. Matsubara, H. Ajiro, M. Fujii, Y. Ohnishi and M. Hatanaka, *Digital Discovery* **4**, 195–203 (2025).
3) R. Somaki, T. Inagaki and M. Hatanaka, *Mol. Inf.* **44**, e202400041 (2025).
4) H. Yamada, S. Takasuka, S. Oikawa, Y. Harashima, T. Takayama, A. Nag, A. Wakiuchi, T. Ando, T. Sugawara, M. Hatanaka, T. Miyao, T. Matsubara, Y. Ohnishi, H. Ajiro and M. Fujii, *ACS Appl. Eng. Mater.* **3**, 2402–2415 (2025).