JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Letter

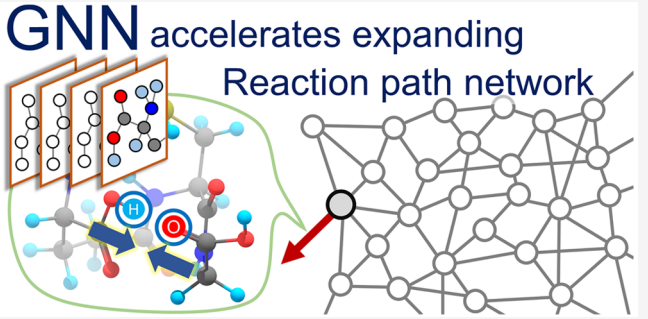
Exploring the Quantum Chemical Energy Landscape with GNN-Guided Artificial Force

Atsuyuki Nakao, Yu Harabuchi, Satoshi Maeda, and Koji Tsuda*

Cite This: *J. Chem. Theory Comput.* 2023, 19, 713−717 Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: Artificial force has been proven useful to get over energy barriers and quickly search a large portion of the energy landscape. This work proposes a method based on graph neural networks to optimize the choice of transformation patterns to examine and accelerate energy landscape exploration. In open search from glutathione, the search efficiency was largely improved in comparison to random selection. We also applied transfer learning from glutathione to tuftsin, resulting in further efficiency gains.



In computational chemistry, chemical reactions are often represented as multiple transitions between equilibrium structures (denoted as EQs) on the quantum chemical energy landscape.[1,2] A reaction path network is defined as a graph where EQs and transitions between them are described as nodes and edges, respectively. Construction and analysis of reaction path networks allow us to predict the products of a reaction and their production pathways from a set of reactants.[3,4] Conversely, the reactants can be predicted from a given set of products by tracing back production pathways.[5]

To construct a reaction path network, various approaches including gradient extremal following,[6] eigenvector following,[7] VADER,[8] anharmonic downward distortion following (ADDF),[9,10] single-/double-ended growing string method with systematic molecular graph transformations,[11,12] freezing string method with the Berny algorithm,[13] nanoreactor,[14] KinBot,[15] Chemoton,[16] and so forth[17−19] have been proposed. Considering the balance of accuracy and computational cost, this study focuses on Monte Carlo search using artificial force and density functional theory (DFT) calculations.[20] It is possible to find neighboring nodes by naive Monte Carlo simulation,[21] but it takes a long time to get over the energy barrier. To cope with the problem, the artificial force induced reaction (AFIR) method employs an artificial force to help the transition beyond a barrier[20] (Figure 1a). The additional force term $V(Q)$ modifies the potential energy surface to encourage transition to another EQ. Monte Carlo simulation based on the modified surface F^{AFIR} provides us a path from the current EQ to another, i.e., AFIR path. A variant of the AFIR algorithm called single-component artificial force induced reaction (SC-AFIR) maintains a list of currently available EQs and choose one based on an energy-based criterion (Figure 1b). Then, two

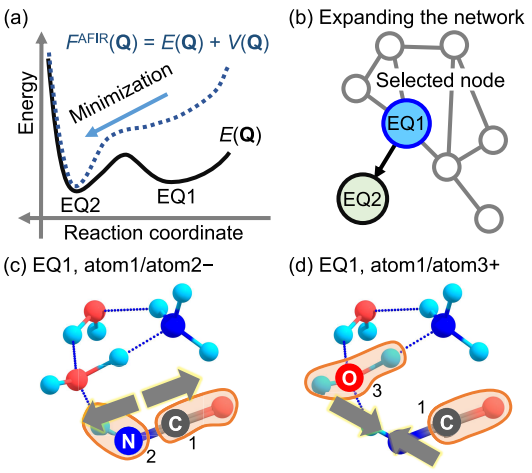
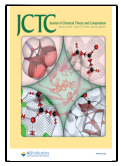


Figure 1. Schematic illustrations of single-component artificial force induced reaction (SC-AFIR) method. (a) AFIR function (F^{AFIR}) consisting of the potential energy surface $E(Q)$ and the force term $V(Q)$ for the reaction coordinate from EQ1 to EQ2. (b) Expansion of the reaction path network by an AFIR-path calculation from a selected node EQ1 to EQ2. (c and d) Fragments around the chosen atom pair pulled apart or pushed closer.

Received: October 25, 2022
Published: January 23, 2023



https://doi.org/10.1021/acs.jctc.2c01061
J. Chem. Theory Comput. 2023, 19, 713−717

atoms in the EQ and the force direction are randomly chosen. The direction can be either positive (push closer, Figure 1d) or negative (pull apart, Figure 1c). An AFIR-path calculation based on DFT is performed with the artifical force applied to the fragments around the chosen atoms. The SC-AFIR method has been proven effective in elucidating various reactions such as *N*-difluoroalkylative dearomatization of pyridines,[22] CO oxidation on Pt(111) surface,[23] and difluoroglycine synthesis.[24]

Despite remarkable applications, the SC-AFIR search still requires an excessively large number of DFT calculations. One reason is that the probability that a new EQ is found is relatively low, because the chosen atoms may not be relevant to any possible reaction. A similar problem is seen in various automated potential energy surface exploration methods. Conventionally, energy-based indices such as the magnitudes of normal-mode eigenvalue and anharmonic downward distortion have been used to rank the importance of paths and improve the probability of finding low-barrier paths.[7,10] However, such energy-based indices tend to guide low-barrier paths such as conformational changes rather than middle- to high-barrier paths accompanying bond reorganizations and thus are not ideal for the present purpose. In this work, we aim to use a deep learning model to optimally choose an *intervention* (i.e., the atom pair and the direction). Given an EQ and an intervention, our graph neural network (GNN) provides a score about the chance of the outcome being a success. We regard the outcome as a success if both of the following criteria are satisfied: (1) A new EQ is found. (2) The new EQ contains a change in 2D structure (e.g., bond formation or break). The second criterion is added to avoid making too many EQs of the same 2D structure and realize quick exploration of the chemical space. Before SC-AFIR is applied to an EQ, all possible interventions are listed and scored by our GNN. The best scoring intervention is chosen for an AFIR-path calculation (Figure 1a,b). In the search, GNN is trained *online*; that is, the training set grows as the search proceeds. In the beginning, GNN conducts un-informed decisions, but it gradually learns from the past experience to make better decisions.

In an exploratory search starting from a tripeptide glutathione, we demonstrate that GNN substantially improves search efficiency over conventional SC-AFIR. In addition, we perform tranfer learning[25] to investigate how well the data obtained in previous searches can improve future search. We took the GNN trained in glutathione search and retrained a part of it in the search from a different peptide, tuftsin. As a result, the search efficiency was improved over the search without transfer learning, showing that knowledge transfer is possible. Our case studies imply that deep learning models that have shown exceptional performances in various tasks[26] can also contribute to reaction analysis and suggest the possibility of further improvement via systematic data collection.

In our model (Figure 2), an EQ is translated to two graphs. One is a fully connected graph, whose node is labeled with a one hot vector of its atom symbol and each edge is labeled, with the following *m*-dimensional vector,

$$e_{ij} = [\exp(-(r_{ij} - \mu_1)^2), \exp(-(r_{ij} - \mu_2)^2), \ldots, \exp(-(r_{ij} - \mu_m)^2)]$$

where r_{ij} is an interatomic distance between atom-*i* and atom-*j* and the cutoff values are defined as $\mu_k = 0.1k$ Å (k = 1, 2, ..., *m*). In the following experiments, *m* is set as 64. The other is a molecular graph whose nodes are labeled with atom symbols and an edge represents the existence of a chemical bond.

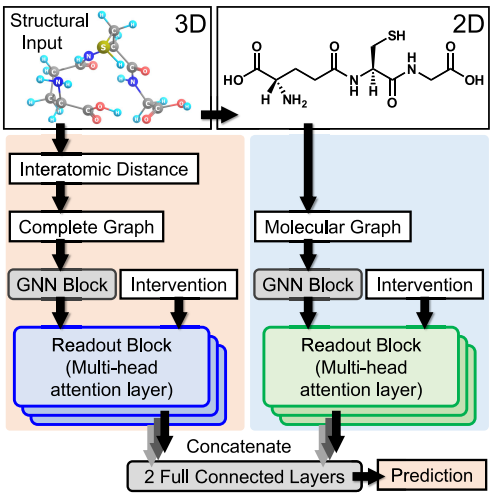


Figure 2. Deep learning model for predicting AFIR outcomes. An EQ geometry is translated to two graphs, i.e., 2D and 3D graphs. A multihead attention layer creates a vector representation. Finally, two vectors are concatenated and fed into a fully connected network to yield a prediction score.

In this network, we are interested in creating a vector representing both the graph and the intervention. Each graph is first processed with 4 layers of graph isomorphism network (GIN),[27] where feature vectors of each node are computed by message-passing. To take the direction into account, an attribute (+1 or −1) is concatenated to all node vectors. Next, a multihead attention layer[28] creates a vector representation specific to the atom pair (i.e., the network pays attention to the atom pair). There are eight independent attention layers, each of which is described as follows. Let $\mathbf{X}$ be the matrix of all feature vectors, and $\mathbf{x}_a$ and $\mathbf{x}_b$ are those of the chosen atom pair, atom-*a* and atom-*b*. Let us define

$$\mathbf{q}_a = \mathbf{x}_a\mathbf{W}_q, \quad \mathbf{q}_b = \mathbf{x}_b\mathbf{W}_q, \quad \mathbf{K} = \mathbf{X}\mathbf{W}_k, \quad \mathbf{V} = \mathbf{X}\mathbf{W}_v$$

where $\mathbf{W}_q$, $\mathbf{W}_k$, and $\mathbf{W}_v$ are trainable parameters for query ($\mathbf{q}$), key ($\mathbf{K}$), and value ($\mathbf{V}$), respectively. The output of the attention layer is computed as

$$\mathbf{z} = \mathrm{softmax}\left(\frac{\mathbf{q}_a\mathbf{K}^{\mathrm{T}}}{\sqrt{n}}\right)\mathbf{V} + \mathrm{softmax}\left(\frac{\mathbf{q}_b\mathbf{K}^{\mathrm{T}}}{\sqrt{n}}\right)\mathbf{V}$$

where *n* is the dimensionality of the query and key, and it is set to eight in this study. The dimensionality of $\mathbf{z}$ is designated as eight, so each graph is translated to a 64-dimensional vector. Finally, the vectors from the two graphs are concatenated and fed into a fully connected network to yield a final prediction score about the outcome. Training was conducted with a Radam optimizer[29] on Pytorch Geometric.[30]

In our implementation of GNN/SC-AFIR, model training, and AFIR-based exploration run in parallel, while sharing a set of training examples. The explorer selects an EQ from the pool of existing EQs by a energy-based criterion and computes the scores with respect to all possible interventions with the current GNN. The top-scoring intervention is chosen and AFIR is carried out to either find a new EQ or fail to do so. This AFIR trial creates a new training example, which is added

https://doi.org/10.1021/acs.jctc.2c01061
J. Chem. Theory Comput. 2023, 19, 713−717

to the training set. The model trainer works continuously in parallel, where GNN is updated with the currently available training set at each epoch.

The SC-AFIR searches are done by GRRM20[31] interfaced with ORCA[32] in a vacuum at the GFN2-xTB level.[33,34] In this study, EQs are selected according to original GRRMs strategy, which is based on Boltzmann distribution. The temperature of the distribution is 10000 K. The algorithm avoiding Hessian calculations is adopted. The model collision energy parameter γ of the AFIR method is set to 200 kJ/mol. To prevent a molecule from going too far from the reaction center, a weak force with $\gamma/(\mathrm{kJ/mol}) = 100/[N(N-1)/2]$ is applied to all atom pairs, where N corresponds to the number of atoms in each system. GNN/SC-AFIR was developed as a stand-alone external code. The GRRM20 program[31] equips SubSelectEQ and SubPathsGen options, and these options allow one to develop an SC-AFIR driver without accessing internal functions of GRRM. The external python code for GNN/SC-AFIR is available at https://github.com/NakaoAtsuyuki/GNN-SC-AFIR.

Glutathione is a tripeptide that works as an antioxidant in plants and animals[35] (Figure 3a). In the search from this

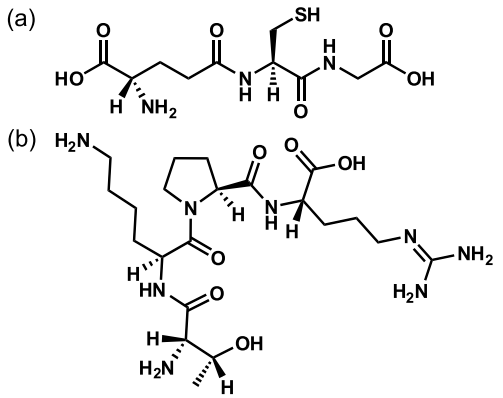


Figure 3. Target molecules: (a) glutathione and (b) tuftsin.

molecule, we compare our model with random selection in efficiency. Figure 4 shows the number of unique 2D structures discovered in the search against the number of AFIR-path calculations during an SC-AFIR search. It is clearly observed that our method is significantly more efficient than a random search. On average over five runs in each method, the success

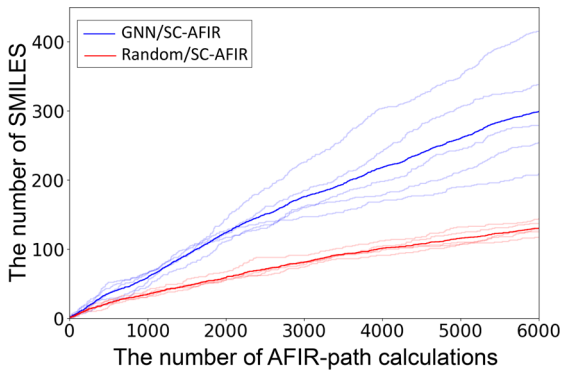


Figure 4. Search efficiency of our model and random selection in the search from glutathione. Light-colored lines represent individual trials, and dark-colored lines represent the mean value of 5 trials.

rate of our model is 18.2%, more than double that of random selection, 8.2%.

Table 1 shows the top-10 frequently selected interventions by our model and random selection. The success rate of the

Table 1. Top-10 Frequently Selected Interventions[a]

(a) Sorted by the number of GNN/SC-AFIR selection

				Selection		Success rate	
entry	atom-*a*	atom-*b*	force	RN	ML	RN	ML
1	O–H	C=O	+	164	1216	20.7 %	33.6 %
2	O–H	C(N)C)C	+	38	1107	18.4 %	30.4 %
3	O–H	O–H	+	91	1099	2.2 %	8.6 %
4	O–H	C–O–H	+	184	699	16.8 %	16.6 %
5	O–H	C–NH–C	+	194	628	24.2 %	33.6 %
6	C–NH–C	C–O–C	+	85	617	16.5 %	20.1 %
7	C–H	O–H	+	633	559	11.4 %	11.6 %
8	O–H	C–O–C	+	97	527	23.7 %	29.4 %
9	C–O–H	C–O–H	+	86	454	9.3 %	4.0 %
10	C=O	C–O–H	+	148	447	10.1 %	11.6 %

(b) Sorted by the number of Random/SC-AFIR selection

				Selection		Success rate	
entry	atom-*a*	atom-*b*	force	RN	ML	RN	ML
1	C–H	C–H	+	940	248	5.4 %	5.2 %
2	C–H	C–O–H	+	877	257	12.3 %	16.3 %
3	C–H	C=O	+	732	186	9.0 %	10.2 %
4	C–H	C–NH–C	+	721	118	8.5 %	5.9 %
5	C–H	N–H	+	703	169	7.7 %	6.5 %
6	C–H	O–H	+	633	559	11.4 %	11.6 %
7	C–H	C(C)C)HN	+	540	88	9.3 %	17.0 %
8	C–H	C–H	−	528	7	2.1 %	0.0 %
9	C–H	CH_2C	+	462	36	6.9 %	0.0 %
10	C–H	C–NH–C	−	413	5	2.2 %	0.0 %

[a]Table a is sorted by the number of GNN/SC-AFIR selections (denoted as ML), and table b is sorted by the number of random/SC-AFIR selections (denoted as RN). Atom-*a* and atom-*b* indicate the vicinity 2D structure around the selected atoms (circled). Force indicates the AFIR force direction, push together (+) or pull apart (−). Success rate represents the rate of reaching EQs having different 2D structures.

atom pairs selected by GNN/SC-AFIR is higher than those by random selection, implying that GNN successfully learned to choose reactive atom pairs. In random selection, a hydrogen connected to a carbon is frequently chosen, simply because such connections are common in molecules. Such hydrogens are often not reactive, and interventions to them are unlikely to change the 2D structure. In GNN/SC-AFIR, the atoms belonging to polar groups are overrepresented. They are highly reactive and likely to lead to new 2D structures. The most frequent pair is the combination of the hydrogen in a carboxy group and the oxygen in a carbonyl group. Considering the fact that a carboxy group is likely to constitute a peptide bond, the atom pair seems to be related to hydrolysis.

Transfer learning is a machine learning methodology of reusing the knowledge obtained from solving a problem to solve a related but different problem.[25] It is often implemented by taking a neural network trained with data about a learning problem and retraining a subset of the parameters with data about a different problem.[36] Here, we take the GNN trained with 60000 examples from the glutathione search and retrain the last part of the GNN, i.e., multihead attention layer and fully connected layers, in searching from a new peptide,

https://doi.org/10.1021/acs.jctc.2c01061
J. Chem. Theory Comput. 2023, 19, 713−717

tuftsin.[37] It is a tetrapeptide Thr-Lys-Pro-Arg related to immunity (Figure 3b). The transfer learning is compared with learning from scratch and random selection. As shown in Figure 5, transfer learning outperformed nontransfer learning

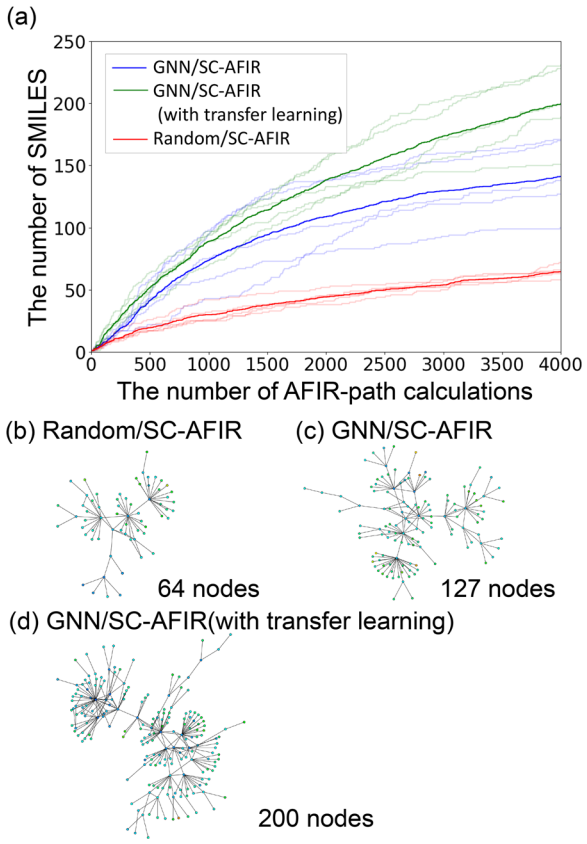


Figure 5. (a) Search efficiency of tuftsin search using transfer learning, nontransfer learning, and random selection. Light-colored lines represent individual trials, and dark-colored lines represent the mean value of 5 trials. (b–d) Constructed reaction path networks by (b) random selection, (c) nontransfer learning, and (d) transfer learning.

significantly, indicating that knowledge obtained in a search can successfully be transferred to a related search. The GRRM input file and the EQs found in the all explorations are available at https://github.com/NakaoAtsuyuki/GNN-SC-AFIR.

In conclusion, we have shown that a graph neural network can be used to accelerate the chemical space search by the SC-AFIR method. Although the guide by GNN has been introduced in the combination with SC-AFIR in this work, it would be possible to combine a similar guide with various other methods such as ones that also choose atom pairs and apply artificial force,[16] ones that choose modes to follow,[6,7,9,10] ones that choose driving coordinates,[8,13] and ones that choose bonds to reorganize.[11,12] Notably, our success in knowledge tranfer is encouraging, because it suggests that the construction of a general purpose model applicable to any reaction (like Alphafold[38]) may be possible. To this aim, however, it is necessary to collect data from as many reactions as possible. Accumulation of such data about reactions would be essential not only to help chemical space search but also to understand the detailed mechanisms of chemical reaction.

AUTHOR INFORMATION

Corresponding Author

Koji Tsuda − *Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa 277-8561, Japan; RIKEN Center for Advanced Intelligence Project, Tokyo 103-0027, Japan; Research and Services Division of Materials Data and Integrated System, National Institute for Materials Science, Tsukuba 305-0047, Japan;* orcid.org/0000-0002-4288-1606; Email: tsuda@k.u-tokyo.ac.jp

Authors

Atsuyuki Nakao − *Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa 277-8561, Japan*

Yu Harabuchi − *Institute for Chemical Reaction Design and Discovery (WPI-ICReDD), Hokkaido University, Sapporo 001-0021, Japan; JST ERATO Maeda Artificial Intelligence for Chemical Reaction Design and Discovery Project, Sapporo 060-0810, Japan; Department of Chemistry, Faculty of Science, Hokkaido University, Sapporo 060-0810, Japan;* orcid.org/0000-0001-8313-3236

Satoshi Maeda − *Institute for Chemical Reaction Design and Discovery (WPI-ICReDD), Hokkaido University, Sapporo 001-0021, Japan; JST ERATO Maeda Artificial Intelligence for Chemical Reaction Design and Discovery Project, Sapporo 060-0810, Japan; Department of Chemistry, Faculty of Science, Hokkaido University, Sapporo 060-0810, Japan;* orcid.org/0000-0001-8822-1147

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.2c01061

Author Contributions

S.M. and K.T. conceived the idea and designed the research. A.N. and Y.H. implemented algorithms and conducted experiments. A.N, Y.H, S.M., and K.T. wrote the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Ryo Tamura and Andrejs Tucs for fruitful discussions. This work is supported by JST ERATO JPMJER1903.

REFERENCES

(1) Schlegel, H. B. Exploring potential energy surfaces for chemical reactions: an overview of some practical methods. *J. Comput. Chem.* **2003**, *24*, 1514−1527.

(2) Wales, D. J.*Energy landscapes: With applications to clusters, biomolecules and glasses*; Cambridge University Press, 2004.

(3) Maeda, S.; Morokuma, K. Finding Reaction Pathways of Type A + B → X: Toward Systematic Prediction of Reaction Mechanisms. *J. Chem. Theory Comput.* **2011**, *7*, 2335−2345.

(4) Unsleber, J. P.; Reiher, M. The exploration of chemical reaction networks. *Annu. Rev. Phys. Chem.* **2020**, *71*, 121−142.

(5) Sumiya, Y.; Harabuchi, Y.; Nagata, Y.; Maeda, S. Quantum Chemical Calculations to Trace Back Reaction Paths for the Prediction of Reactants. *JACS Au* **2022**, *2*, 1181−1188.

(6) Bondensgard, K.; Jensen, F. Gradient extremal bifurcation and turning points: An application to the H_2CO potential energy surface. *J. Chem. Phys.* **1996**, *104*, 8025−8031.

(7) Doye, J.; Wales, D. Surveying a potential energy surface by eigenvector-following Applications to global optimization and the structural transformations of clusters. *Z. Phys. D* **1997**, *40*, 194−197.

(8) Černohorský, M.; Kettou, S.; Koča, J. *VADER*: New Software for Exploring Interconversions on Potential Energy Surfaces. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 705−712.
(9) Maeda, S.; Ohno, K. Global Mapping of Equilibrium and Transition Structures on Potential Energy Surfaces by the Scaled Hypersphere Search Method: Applications to ab Initio Surfaces of Formaldehyde and Propyne Molecules. *J. Phys. Chem. A* **2005**, *109*, 5742−5753.
(10) Maeda, S.; Taketsugu, T.; Morokuma, K.; Ohno, K. Anharmonic Downward Distortion Following for Automated Exploration of Quantum Chemical Potential Energy Surfaces. *Bull. Chem. Soc. Jpn.* **2014**, *87*, 1315−1334.
(11) Zimmerman, P. M. Automated discovery of chemically reasonable elementary reaction steps. *J. Comput. Chem.* **2013**, *34*, 1385−1392.
(12) Zimmerman, P. M. Single-ended transition state finding with the growing string method. *J. Comput. Chem.* **2015**, *36*, 601−611.
(13) Suleimanov, Y. V.; Green, W. H. Automated Discovery of Elementary Chemical Reaction Steps Using Freezing String and Berny Optimization Methods. *J. Chem. Theory Comput.* **2015**, *11*, 4248−4259.
(14) Wang, L.-P.; McGibbon, R. T.; Pande, V. S.; Martinez, T. J. Automated Discovery and Refinement of Reactive Molecular Dynamics Pathways. *J. Chem. Theory Comput.* **2016**, *12*, 638−649.
(15) Van de Vijver, R.; Zador, J. KinBot: Automated stationary point search on potential energy surfaces. *Comput. Phys. Commun.* **2020**, *248*, 106947.
(16) Unsleber, J. P.; Grimmel, S. A.; Reiher, M. Chemoton 2.0: Autonomous Exploration of Chemical Reaction Networks. *J. Chem. Theory Comput.* **2022**, *18*, 5393−5409.
(17) Maeda, S.; Ohno, K.; Morokuma, K. Systematic exploration of the mechanism of chemical reactions: the global reaction route mapping (GRRM) strategy using the ADDF and AFIR methods. *Phys. Chem. Chem. Phys.* **2013**, *15*, 3683−3701.
(18) Dewyer, A. L.; Argüelles, A. J.; Zimmerman, P. M. Methods for exploring reaction space in molecular systems. *WIREs Comput. Mol. Sci.* **2018**, *8*, e1354.
(19) Simm, G. N.; Vaucher, A. C.; Reiher, M. Exploration of reaction pathways and chemical transformation networks. *J. Phys. Chem. A* **2019**, *123*, 385−399.
(20) Maeda, S.; Harabuchi, Y. Exploring paths of chemical transformations in molecular and periodic systems: An approach utilizing force. *WIREs Comput. Mol. Sci.* **2021**, *11*, e1538.
(21) Frenkel, D.; Smit, B.*Understanding molecular simulation: From algorithms to applications*; Elsevier: Amsterdam, 2001; Vol. *1*. DOI: 10.1016/B978-0-12-267351-1.X5000-7.
(22) Hayashi, H.; Katsuyama, H.; Takano, H.; Harabuchi, Y.; Maeda, S.; Mita, T. In silico reaction screening with difluorocarbene for N-difluoroalkylative dearomatization of pyridines. *Nat. Synth.* **2022**, *1*, 804−814.
(23) Sugiyama, K.; Sumiya, Y.; Takagi, M.; Saita, K.; Maeda, S. Understanding CO oxidation on the Pt (111) surface based on a reaction route network. *Phys. Chem. Chem. Phys.* **2019**, *21*, 14366−14375.
(24) Mita, T.; Harabuchi, Y.; Maeda, S. Discovery of a synthesis method for a difluoroglycine derivative based on a path generated by quantum chemical calculations. *Chem. Sci.* **2020**, *11*, 7569−7577.
(25) Zhuang, F.; Qi, Z.; Duan, K.; Xi, D.; Zhu, Y.; Zhu, H.; Xiong, H.; He, Q. A comprehensive survey on transfer learning. *Proceedings of the IEEE* **2021**, *109*, 43−76.
(26) LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436−444.
(27) Hu, W.; Liu, B.; Gomes, J.; Zitnik, M.; Liang, P.; Pande, V.; Leskovec, J.Strategies for pre-training graph neural networks. *arXiv Preprint (Machine Learning)*, 2019. arXiv:1905.12265. https://arxiv.org/abs/1905.12265.
(28) Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, Ł; Polosukhin, I.Attention is all you need. *NIPS '17: Proceedings of the 31st International Conference on Neural Information Processing Systems*; MIT Press, 2017; Vol. *30*, pp 6000−6010.
(29) Liu, L.; Jiang, H.; He, P.; Chen, W.; Liu, X.; Gao, J.; Han, J.On the variance of the adaptive learning rate and beyond. *arXiv Preprint (Machine Learning)*, 2019. arXiv:1908.03265. https://arxiv.org/abs/1908.03265.
(30) Fey, M.; Lenssen, J. E.Fast graph representation learning with PyTorch Geometric. *arXiv Preprint (Machine Learning)*, 2019. arXiv:1903.02428. https://arxiv.org/abs/1903.02428
(31) Maeda, S.; Harabuchi, Y.; Sumiya, Y.; Takagi, M.; Suzuki, K.; Sugiyama, K.; Ono, Y.; Hatanaka, M.; Osada, Y.; Taketsugu, T.; Morokuma, K.; Ohno, K.*GRRM20*; Hokkaido University: Sapporo, Japan2020.
(32) Neese, F. The ORCA program system. *WIREs Comput. Mol. Sci.* **2012**, *2*, 73−78.
(33) Grimme, S.; Bannwarth, C.; Shushkov, P. A robust and accurate tight-binding quantum chemical method for structures, vibrational frequencies, and noncovalent interactions of large molecular systems parametrized for all spd-block elements (Z= 1−86). *J. Chem. Theory Comput.* **2017**, *13*, 1989−2009.
(34) Bannwarth, C.; Caldeweyher, E.; Ehlert, S.; Hansen, A.; Pracht, P.; Seibert, J.; Spicher, S.; Grimme, S. Extended tight-binding quantum chemistry methods. *WIREs Comput. Mol. Sci.* **2021**, *11*, e1493.
(35) Wu, G.; Fang, Y.-Z.; Yang, S.; Lupton, J. R.; Turner, N. D. Glutathione metabolism and its implications for health. *J. Nutr.* **2004**, *134*, 489−492.
(36) Yamada, H.; Liu, C.; Wu, S.; Koyama, Y.; Ju, S.; Shiomi, J.; Morikawa, J.; Yoshida, R. Predicting materials properties with little data using shotgun transfer learning. *ACS Cent. Sci.* **2019**, *5*, 1717−1730.
(37) Siebert, A.; Gensicka-Kowalewska, M.; Cholewinski, G.; Dzierzbicka, K. Tuftsin-Properties and analogs. *Curr. Med. Chem.* **2017**, *24*, 3711−3727.
(38) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583−589.

https://doi.org/10.1021/acs.jctc.2c01061
J. Chem. Theory Comput. 2023, 19, 713−717