ACS OMEGA

Open Access

http://pubs.acs.org/journal/acsodf

Article

Enhancing the Search Performance of Bayesian Optimization by Creating Different Descriptor Datasets Using Density Functional Theory

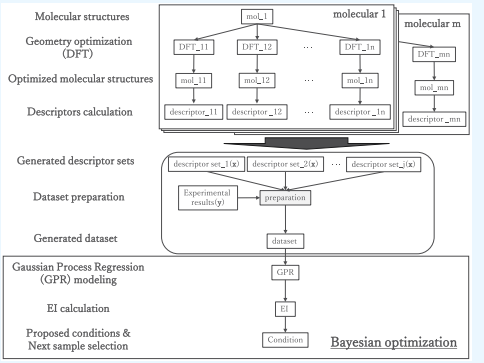
Toshiharu Morishita and Hiromasa Kaneko*

Cite This: *ACS Omega* 2023, 8, 33032−33038

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Descriptors calculated from molecular structure information can be used as explanatory variables in Bayesian optimization (BO). Even though structural and descriptor information can be obtained from various databases for general compounds, information on highly confidential compounds such as pharmaceutical intermediates and active pharmaceutical ingredients cannot be retrieved from these databases. In particular, determining the stable structure and electronic state of a compound via quantum chemical calculations from descriptor information requires considerable computational time. Although descriptor information can be obtained using density functional theory (DFT), which has a relatively light computational load, only conventional combinations of basis sets and functionals can be selected before experiments instead of the best ones. Few studies have discussed these effects on the search performance of BO, and good search performance is highly dependent on the application. Therefore, we developed a method to improve the search performance of BO by using descriptors computed from several combinations of basis sets and functionals. The dataset obtained from averaging multiple descriptor sets exhibited better BO search performance than that of a single descriptor dataset. In addition, the more descriptor sets used for averaging, the better the search performance. This method has a relatively small computational load and can be easily used by those who are unfamiliar with quantum chemical calculations.



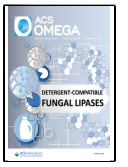
INTRODUCTION

The search for an optimal solution is a crucial area of research in machine learning, neural networks, robotics, aerospace engineering, and the design of experiments (DoE). In this context, Bayesian optimization (BO)[1,2] has been widely explored as a solution. For example, BO is used to improve search performance and conduct multi-objective optimization to accelerate the development of new materials.[3,4] It was previously used to determine the Pareto optimal solution for the optimal design of chemical reactors using computational fluid dynamics (CFD), which reduced the number of CFD calculations. Moreover, it was used to minimize power consumption and maximize gas retention in a stirred tank reactor.[5] Recently, constrained BO has been applied to reduce the generation of invalid molecular structures while using the latent space of variational autoencoders.[6] Additionally, it has been applied to identify the most stable molecular conformers. Finding low-energy molecular conformers is challenging because of the high dimensionality of the search space and the computational cost of accurate quantum chemical methods for determining conformer structures and energies. However, combining BO with density functional theory (DFT) reduces the cost of quantum chemical calculations by ∼90%.[7] Therefore, adaptive experimental design methods using BO have been studied and applied in various fields.

However, there are few examples of its application in the optimization of reaction conditions. In 2021, a framework for Bayesian reaction optimization and an open-source software tool that allows chemists to easily integrate optimization algorithms into their laboratory practices were developed.[13] In that study, a large benchmark dataset for palladium-catalyzed direct arylation was collected through high-throughput experimentation (HTE) to compare the number of experiments required to reach the optimal solution. BO optimized the reaction conditions faster than experts in organic synthesis, and the descriptors obtained from the molecular structures via DFT calculations exhibited a higher search performance than that obtained when the compounds were treated by one-hot-encoding (OHE).[8,9]

Even though structural and descriptor information can be obtained from various databases for general compounds, highly confidential compounds such as pharmaceutical intermediates and active pharmaceutical ingredients cannot be retrieved or

Received: July 7, 2023
Accepted: August 14, 2023
Published: August 28, 2023



obtained from databases. In particular, when the stable structure and electronic state of a compound must be determined by quantum chemical calculations, obtaining descriptor information requires considerable computational time. Approximate calculations led to reduced accuracy, and new methods have been developed to reduce these errors.[17] Although descriptor information could be obtained using DFT, which has a relatively low computational load, selecting the best basis sets and functionals before conducting experiments is difficult, as only the typical conventional combinations are selected. Few studies have discussed these effects on the search performance of BO, and a good search performance depends on the application.

Therefore, in this study, we developed a method to improve the search performance of BO by using descriptors computed by adopting several combinations of basis sets and functionals. Descriptors were computed for the two reactions studied in Shields et al.,[8] direct arylation and Suzuki–Miyaura coupling reactions, and the search performance in BO was compared using datasets created from each set of descriptors. Subsequently, we identified the best method to select the descriptor sets for BO and confirmed its effect on search performance in BO.

METHODS

Bayesian Optimization. The Gaussian process regression (GPR) model was used for the BO. The Gaussian process GP($\mu(\boldsymbol{x})$, $k(\boldsymbol{x},\boldsymbol{x}')$) represents a distribution of the functions characterized by a prior mean $\mu(\boldsymbol{x})$ and a kernel function $k(\boldsymbol{x},\boldsymbol{x}')$. Matern52 was selected as the kernel function $k(\boldsymbol{x},\boldsymbol{x}')$. Previously optimized hyperparameters[8] were used for the GPR and BO in this study.

$$f \sim \mathrm{GP}(\mu(\boldsymbol{x}), k(\boldsymbol{x}, \boldsymbol{x}')) \tag{1}$$

$$k(r) = \alpha\left(1 + \frac{\sqrt{5}r}{l} + \frac{5r^2}{3l^2}\right)e^{-\sqrt{5}r/l} \tag{2}$$

$$r = |\boldsymbol{x} - \boldsymbol{x}'| \tag{3}$$

where r is the distance between the experimental conditions, α is the output scale parameter, and l is the length scale parameter. The Gaussian process posterior distribution mean μ under the experimental condition $\boldsymbol{x}$ was expressed as follows:

$$\mu(\boldsymbol{x}) = k(\boldsymbol{x})^T(K + \sigma_n^2 I)^{-1}\boldsymbol{y} \tag{4}$$

where $k(\boldsymbol{x})$ is the covariance vector between the experimental condition $\boldsymbol{x}$ and the training conditions, K is the covariance matrix between all training conditions, σn^2 is the variance of the estimated noise, I is the identity matrix, and $\boldsymbol{y}$ is the vector of responses corresponding to the training data. The variance in the posterior distribution of the Gaussian process under experimental condition $\boldsymbol{x}$ was expressed as follows:

$$\sigma^2(\boldsymbol{x}) = k(\boldsymbol{x}, \boldsymbol{x}) - k(\boldsymbol{x})^T(K_\theta + \sigma_n^2 I)^{-1}k(\boldsymbol{x}) \tag{5}$$

Predictions are normally distributed in GPR; therefore, the hyperparameters can be calculated using the maximum likelihood estimation. The hyperparameters were determined such that the following log-likelihood function was maximized:

$$\log p(\boldsymbol{y}|\theta) = -\frac{1}{2}\boldsymbol{y}^T(K_\theta + \sigma_n^2 I)^{-1}\boldsymbol{y} - \frac{1}{2}\log|K_\theta + \sigma_n^2 I| - \frac{n}{2}\log 2\pi \tag{6}$$

The expected improvement (EI), which is the expected value of $I(\boldsymbol{x})$, is generally selected as the acquisition function in BO. The improvement $I(\boldsymbol{x})$ represents an increase in the objective function $f(\boldsymbol{x})$ relative to the best outcome f^+.

$$I(\boldsymbol{x}) = \begin{cases} f(\boldsymbol{x}) - f^+ - \delta & f(\boldsymbol{x}) > f^+ \\ 0 & f(\boldsymbol{x}) < f^+ \end{cases} \tag{7}$$

The expectation value of $I(\boldsymbol{x})$, $EI(\boldsymbol{x})$, for a given experimental condition $\boldsymbol{x}$ has the following form:

$$EI(\boldsymbol{x}) = \begin{cases} I(\boldsymbol{x})\Phi\left(\frac{I(\boldsymbol{x})}{\sigma(\boldsymbol{x})}\right) + \sigma(\boldsymbol{x})\varphi\left(\frac{I(\boldsymbol{x})}{\sigma(\boldsymbol{x})}\right) & \sigma(\boldsymbol{x}) > \delta \\ 0 & \sigma(\boldsymbol{x}) < \delta \end{cases} \tag{8}$$

where $I(\boldsymbol{x})$ is the improvement in the surrogate mean prediction $\mu(\boldsymbol{x})$ diminished by δ, an empirical exploration parameter, which was set to 0.01, a commonly used value; $\sigma(\boldsymbol{x})$ is the surrogate standard deviation; and Φ and φ are the cumulative distribution function and probability density function of the standard normal distribution, respectively.

The following experimental conditions were selected based on the value of $EI(\boldsymbol{x})$. Considering that the experimental conditions are expressed as a combination of various compounds, the chemical space is finite, and $\boldsymbol{x}$ can be selected such that the expected value of $EI(\boldsymbol{x})$ is the highest.

$$\arg\max_{\boldsymbol{x}\in X} EI(\boldsymbol{x}) \tag{9}$$

If the experiments are conducted in parallel, the Kriging Believer algorithm[14] can be used to iteratively compute $\boldsymbol{x}$, for which $EI(\boldsymbol{x})$ is the maximum. This was achieved by adding the Gaussian process posterior mean $\mu(\boldsymbol{x})$ to the known data and updating the GPR model. Hence, we used the following procedure for calculating the BO, and the corresponding flowchart is presented in Figure 1

1. The experimental space was defined (e.g., solvent, ligand, and temperature), and the initial samples $\boldsymbol{x}$ were selected. If the information on samples $\boldsymbol{x}$ and the corresponding $\boldsymbol{y}$ was already available, Step 3 was directly conducted.
2. The experiments were performed based on the selected samples $\boldsymbol{x}$.
3. A GPR model was built using the information on samples $\boldsymbol{x}$ and the corresponding $\boldsymbol{y}$, and the EI values of all samples were calculated. All $\boldsymbol{x}$ and $\boldsymbol{y}$ were normalized and used in the calculation.
4. The sample with the highest EI was selected for subsequent experiments. If the experiments are conducted in parallel, the calculations are repeated.
5. Experiments were conducted based on the selected samples.
6. Steps 3–5 were repeated until $\boldsymbol{y}$ reached the target value.

Dataset Creation. A group of multiple descriptors was created from the structural information of compounds obtained by DFT calculations, and a dataset was generated by averaging the descriptor sets. The obtained dataset was then

https://doi.org/10.1021/acsomega.3c04891
ACS Omega 2023, 8, 33032–33038

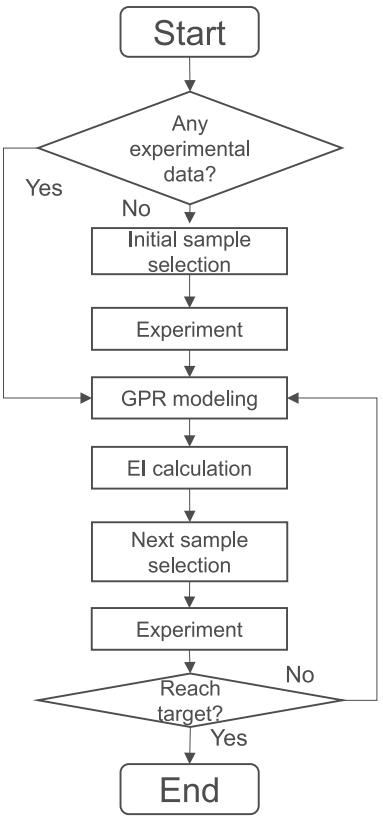


Figure 1. Procedure for Bayesian optimization (BO).

used to build a GPR model and perform BO. The calculation procedure is illustrated in Figure 2 and described below:

1. The molecular structures (MOL files, etc.) were prepared for m compounds.
2. DFT calculations were performed with n combinations of various basis sets and functionals using Gaussian quantum chemical calculation software.[15]
3. Zero-, one-, two-, and three-dimensional descriptors were computed using the Mordred[10,12] and Codessa[16] descriptor computation methods.
4. The k descriptor sets obtained from the DFT calculations were averaged and combined with a continuous-valued explanatory variable $\mathbf{x}$ and an objective variable $\mathbf{y}$ to create a dataset. The number of descriptor sets (k) depends on the number of variables, compounds, basis sets, functionals, and the method of combination.
5. A GPR model was constructed using the created dataset, and the condition with the highest acquisition function value in the BO was selected as the next experimental condition.

RESULTS AND DISCUSSION

Datasets. The search performances of the BO of the direct palladium-catalyzed arylation (Scheme 1) and Suzuki–Miyaura coupling (Scheme 2) reported by Shields et al.[8] were verified using our strategy.

In the case of the direct palladium-catalyzed arylation (Reaction A), the experimental conditions were fixed for equivalent reaction substances, catalysts, and ligands, and 1728 combinations of three reaction temperatures, three substance concentrations, twelve ligands, three bases, and four solvents were used (Table S1). Only ten conditions resulted in yields of ≥95%, accounting for 0.58% of the dataset, and only seven conditions exhibited yields of ≥98%, accounting for 0.41% of the dataset. Because ligands, bases, and solvents are categorical rather than quantitative variables, we used Mordred[10,11] to convert the molecular structures of the MOL files into zero-,

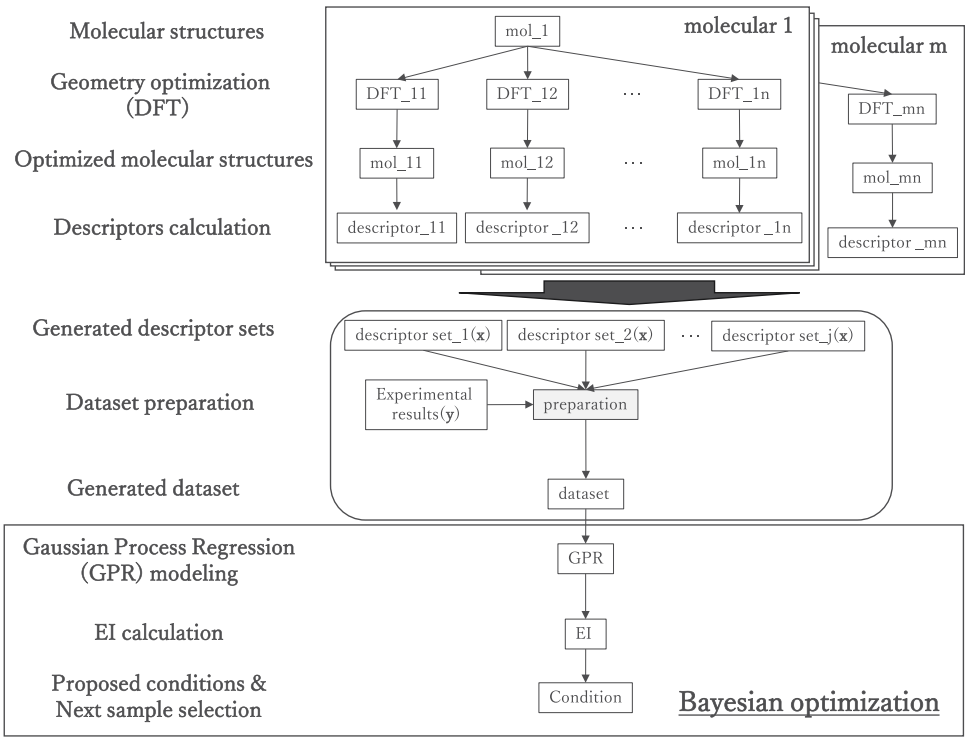
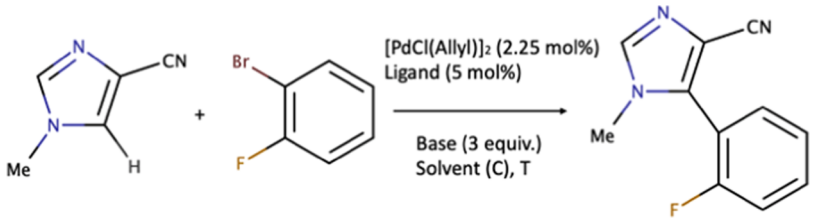


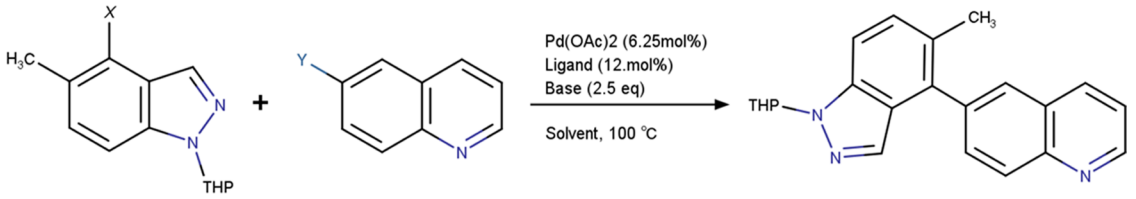
Figure 2. Procedure for creating datasets for BO.

https://doi.org/10.1021/acsomega.3c04891
ACS Omega 2023, 8, 33032−33038

Scheme 1. Reaction A: Direct Pd-Catalyzed Arylation



Scheme 2. Reaction B: Suzuki−Miyaura Coupling



one-, and two-dimensional molecule descriptors. Additionally, the molecular structures calculated by DFT using the quantum chemical calculation software Gaussian 16 W and Gauss View 6.0.16[15] were converted into three-dimensional molecular descriptors using the molecular descriptor calculation software Codessa 3.3.[16] All resulting descriptors were used for the calculation. In the DFT calculations, 10 basis sets (STO-3G, 3−21G, 3−21Gd, 6-31G, 6-31Gd, 6-31Gdp, 6-31G+d, 6-311G, 6-311Gd, and 6-311Gdp) and the B3LYP functional, commonly used for organic synthesis calculations, were used. The descriptors obtained were reduced to 20 variables by principal component analysis. The cumulative contribution ratio was >99%.

In the case of the Suzuki−Miyaura coupling (Reaction B), the experimental conditions were fixed for equivalent reaction substances, bases, solvents, catalysts, and ligands, and 3696 combinations of four electrophiles, three nucleophiles, eleven ligands, seven bases, and four solvents were used (Table S4). Only 71 conditions exhibited yields of ≥95%, accounting for 1.92% of the dataset, whereas only 10 conditions exhibited yields of ≥98%, accounting for 0.27% of the dataset. Because electrophiles, nucleophiles, ligands, bases, and solvents are categorical variables rather than quantitative variables, they were converted into descriptors for the calculations. The descriptor calculation method and postprocessing were the same as those for Reaction A.

Search Performance Benchmarking. BO was conducted for Reactions A and B under various conditions to compare search performances. The reactions, number of descriptors, target yields, and number of proposed experiments (per round) for each case are listed in Table 1. In Cases 1, 2, and 4, we used all of the descriptors calculated from the molecular information obtained by DFT calculations using all 10 basis sets. In Case 3, for the categorical variables ligand, solvent, and base, we used descriptors calculated from the molecular information obtained by DFT calculations with the basis sets STO-3G, 3−21G, and 6-31G. In Cases 5 and 6, for the categorical variables electrophile, nucleophile, ligand, base, and solvent, we used descriptors calculated from the molecular information obtained by DFT calculations with the basis sets STO-3G and 6-31G. The initial conditions were determined randomly, and more than 10,000 calculations were performed for each case to eliminate the influence of random numbers. The mean and standard deviation of the number of rounds required to reach the target yield were compared. However, no relationship was identified between the choice of basis sets, functionals, and search performance in BO (Tables S7−S12), suggesting that more rigorous DFT calculations would not necessarily have resulted in better search performance. These results were used as a benchmark for the evaluation of the dataset creation and for the search performance when a single descriptor set was used.

Table 1. Reactions A and B under Different Conditions for Comparing Their Search Performances in BO

case no.	1	2	3	4	5	6
reaction	A	A	A	B	B	B
number of descriptor sets	10	10	27	10	32	32
target yield	98	95	98	95	98	95
number of proposed experiments	5	5	5	5	10	5

Comparison of the Methods Used for Creating Descriptor Datasets. To confirm the influence of the descriptor dataset on the search performance of BO, we compared the search performance using seven methods: one where a new dataset is generated by averaging descriptor sets (the method proposed in this study), another where all of the descriptors are used as a single dataset, and the other five in which datasets are created by selecting sets of descriptors to be used in BO (random selection, similarity of descriptor sets: D-optimality and mean distance from the center of gravity, prediction performance of the GPR model, and acquisition function EI value of BO). The calculation procedure for all of the methods was almost the same as the abovementioned dataset creation method (Figure 2), except for the dataset selection criteria. In the descriptor set selection method, the descriptor sets used in the BO were reselected for each calculation to create the dataset. When multiple descriptor sets were selected (ensemble), they were selected in the order of the highest or in the order of the lowest of the respective evaluation indices. Details of the five methods are presented below.

- Random selection of a descriptor set: A uniform random number was used to randomly select a group of descriptors for BO.

https://doi.org/10.1021/acsomega.3c04891
ACS Omega 2023, 8, 33032−33038

- Selection based on diversity (D-optimality and mean distance from the center of gravity): The poor search performance of BO can be attributed to its inability to escape a local optimum. Therefore, we selected descriptors that were as dissimilar as possible to obtain diverse conditions. First, to evaluate the similarity of the descriptor sets, we vectorized (flattened) the descriptor sets and combined them into a single dataset. The average distance from the center of gravity and the D-optimality of each created descriptor set were determined. The greater the average distance from the center of gravity or the D-optimality, the more dissimilar the descriptor sets were considered.
- Selection based on the prediction performance of the GPR model: Assuming that a higher generalization performance of GPR will lead to a proportionally high search performance of BO, we cross-validated the GPR models constructed in BO, evaluated their generalization performance using mean squared error (MSE), and selected the model (descriptor set) with the lowest MSE for a conditional search. Because this method requires validation for all GPR models and the computational load is significantly large (Figures S1 and S2), the number of folds for cross-validation was set to the number of proposed experiments (per round). Because cross-validation cannot be performed when the number of proposed experiments is one, the number of proposed experiments should be at least five.
- Selection based on expected improvement: Assuming that the search performance of BO would be higher for a condition with a high expected improvement interval (EI), all of the acquisition function values (EI) calculated for the multiple BOs were stored, and the models with the highest EI values were adopted to propose the next experimental condition. However, the computational load becomes relatively large owing to the necessity of a GPR model and EI calculations for all of the created descriptor datasets (Figures S1 and S2).

For the six cases listed in Table 1, we used twelve different dataset creation methods (Table 2) to check for differences in their search performances. The basis function, functional, number of datasets, termination condition, and number of proposed experiments (per round) for each case are listed in Table 1. The initial conditions were determined randomly, and more than 10,000 calculations were performed for each case to minimize the effects of the initial conditions and random numbers as much as possible. Two types of evaluations were conducted to compare the methods for selecting descriptors for BO: one using a single descriptor set and the other using an ensemble of multiple (five) descriptor sets. In Table 2, the numbers listed after the method name indicate the number of descriptor sets selected.

Table 2. Selection of the Dataset Creation Method

method name	number of selected descriptor sets	explanation
Rand1	1	random selection of the descriptor set
Rand5	5	
dopt1	1	descriptor set selection based on diversity (D-optimality)
dopt5	5	
dist1	1	descriptor set selection based on diversity (average distance from the center of gravity)
dist5	5	
CV1	1	descriptor set selection based on the prediction performance of the GPR model
CV5	5	
EI1	1	descriptor set selection based on expected improvement
EI5	5	
all	1	all of the descriptor sets combined into a single descriptor set
ave	1	average of all of the descriptor sets used as a single descriptor set (proposed method)

Figures 3 and S3−S7 show the search results of the BO performed using different dataset creation methods. The

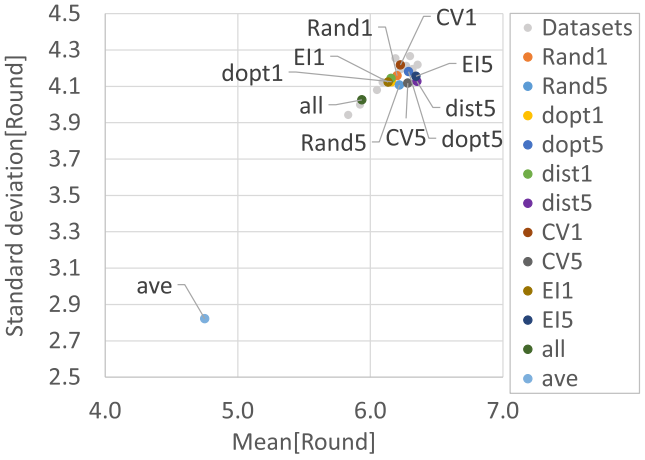


Figure 3. Mean and standard deviation of the number of experimental rounds for Case 1. Case descriptions are summarized in Table 1.

horizontal axes in the figures indicate the average number of rounds required to reach the target yield, and the vertical axes indicate the standard deviation. The notations in the legend are the names of the dataset creation methods, and the datasets indicate the search results for the BO performed using the previously established benchmark descriptor set. The method using the average of all descriptor sets (ave) obtained in the DFT calculations showed the highest search performance among all cases, and both the mean and standard deviation were smaller than those obtained for a single set of descriptors and the other methods. Notably, even in the case of significantly decreased search performance for many methods in Case 4 (Figure S5), the outlier descriptors did not have a significant effect on the search performance. Moreover, the method of selecting descriptor sets did not significantly improve search performance, regardless of the number of selected descriptor sets (single or multiple). This could be explained by the absence of a high correlation between the five indices used to select descriptors and the search performance in BO. In addition, the method with high diversity did not improve the search performance owing to the absence of a relationship between the diversity of the dataset and the search performance of BO. Although the method involving all of the descriptor sets showed better search performance than the method involving some descriptors, it was not as effective as the method involving the average of all descriptors. Figures 4 and S8−S12 show the cumulative relative frequency of the number of rounds required to reach the target yield normalized by the cumulative relative frequency of the averaging method (ave). The horizontal axis represents the number of rounds required to reach the target yield in each trial, and the vertical axis represents the normalized cumulative relative frequency.

https://doi.org/10.1021/acsomega.3c04891
ACS Omega 2023, 8, 33032−33038

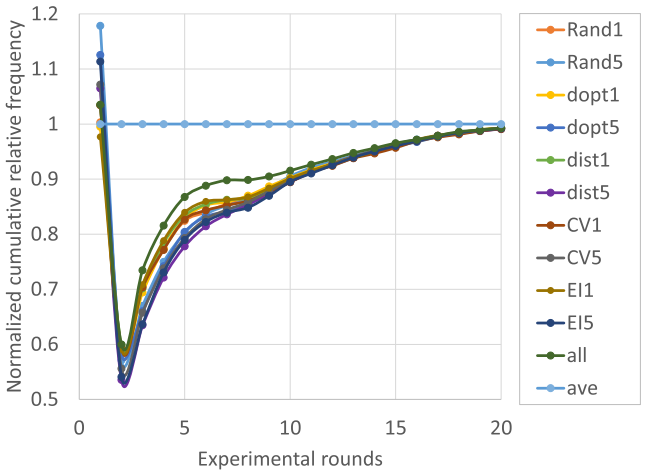


Figure 4. Experimental rounds vs normalized cumulative relative frequency for Case 1. Case descriptions are summarized in Table 1.

Although the results of the first round slightly varied owing to the effect of random numbers, the search performance of the averaging method (ave) was higher than that of the other methods in most cases from the second round onward.

In many cases, the dataset creation method involving the averaged descriptor sets improved the search performance compared to the method with a single descriptor dataset. The search performance in Case 1 was investigated for all combinations of the descriptors, where two, three, four, five, eight, and ten were averaged (Figure 5 and Tables S13 and S14). The mean and standard deviation of the number of rounds required to reach the target yield were smaller in the averaged descriptor sets than those obtained for the single descriptor dataset. In addition, as the number of descriptors used for averaging increased, the percentage improvement in the search performance also increased (Table 3).

The performance of the BO search varied depending on the type of basis sets and functionals selected for the DFT calculations, which could be attributed to the errors in the structure-optimized molecular structures and calculated descriptors. The convergence state of the molecular structure was also affected by other calculation conditions. For example, the target molecule would not be represented correctly if the initial structure is poor and the local optimum is reached, or if the convergence is not complete because of poor setting parameters. In general, the data obtained from analytical instruments is highly dependent on the operator's technique, inter-device differences, and setting parameters, which may lead to values with errors. Similarly, DFT calculations are highly dependent on the initial structure of a given molecule, convergence conditions, basis sets, functions, and other setting parameters, which could lead to values with errors. In both cases, the true value is unknown, and obtaining values with errors is common. To overcome this issue in the case of analytical instruments and obtain reliable values, measures such as increasing the number of measurements or using the average value of multiple observations obtained from different instruments are adopted. Similarly, considering DFT calculations as an analytical instrument, the average value of the structural information and descriptors obtained from multiple

Table 3. Percentage Improvement of the Average Number of Rounds and Standard Deviation Required to Reach the Target Using the Proposed Method Compared to Those Obtained Using a Single Descriptor Set (A: Multi Datasets, B: Single Dataset)

number of descriptor sets	number of combinations	average number of rounds of A compared to that of B [%]	standard deviation of A compared to that of B [%]
2	45	86.7	86.7
3	120	90.8	95.8
4	210	98.1	99.5
5	252	98.6	99.5
8	45	100	100
10	1	100	100

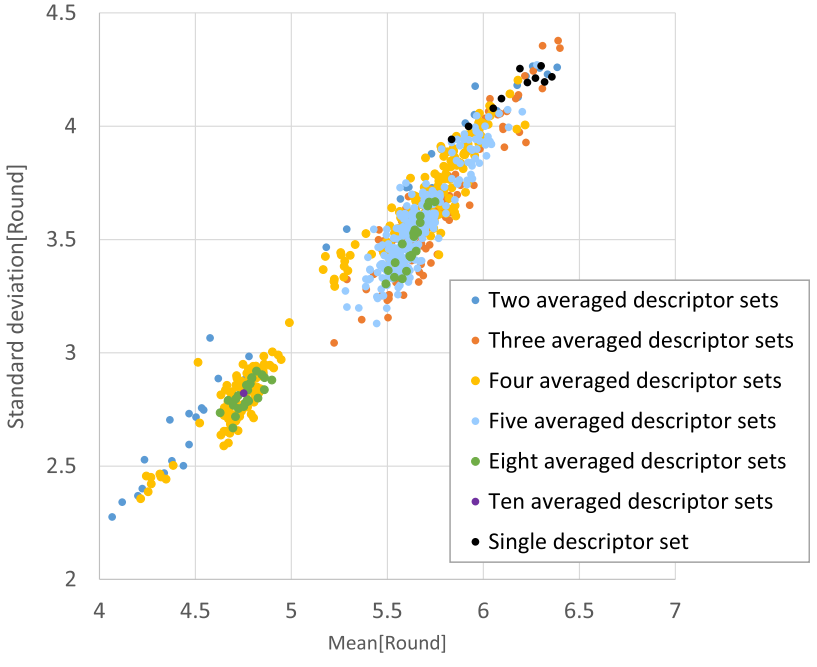


Figure 5. Mean and standard deviation of the number of experimental rounds with averaged descriptor sets.

https://doi.org/10.1021/acsomega.3c04891
ACS Omega 2023, 8, 33032−33038

DFT calculations should be closer to the true value. Hence, we believe that the search performance in BO may have improved by using the proposed method. However, even if the averaging approaches the true state, the space represented by the descriptors must represent the behavior of the objective variable that should be optimized. Assuming that the descriptors used in the reactions in this study conform to the abovementioned conditions, we believe that the effect may have been achieved. Although averaging the descriptors is simple and not novel, it achieved a considerable effect. In addition, the computational load was small (Figures S1−S2), and researchers with little domain knowledge can easily use it. The validity of the proposed method was confirmed for several reactions and combinations of descriptors, and we believe that this method may be applicable to other synthetic reactions. However, this strategy should be verified in other reactions under optimized reaction conditions and should be reproducible using different combinations of basis sets and functionals.

CONCLUSIONS

Descriptors of chemical compounds calculated from their molecular structure information can be used as explanatory variables in Bayesian optimization (BO). In this study, we created descriptor sets for direct palladium-catalyzed arylation and Suzuki−Miyaura coupling reactions using several combinations of basis sets and functionals and verified the search performance in BO using newly created descriptor datasets. The dataset created by averaging multiple sets of descriptors exhibited better BO search performance than that obtained for the dataset created from a single set of descriptors. In addition, the percentage of improved search performance increased as the number of descriptor sets used for averaging increased. Furthermore, this strategy of averaging multiple descriptor sets has a low computational load and can be easily used by researchers with limited knowledge of quantum chemical computations. We intend to confirm whether the proposed method can be used for other synthetic reactions using different combinations of basis sets and functionals in a future study.

ASSOCIATED CONTENT

Data Availability Statement

The data (molecular information and calculation results) underlying this study are available in the published article and its Supporting Information.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.3c04891.

Reaction space, compound list, and experimental results (Tables S1−S14 and Figures S1 and S2) (PDF)

AUTHOR INFORMATION

Corresponding Author

Hiromasa Kaneko − *Department of Applied Chemistry, School of Science and Technology, Meiji University, Kawasaki, Kanagawa 214-8571, Japan;* orcid.org/0000-0001-8367-6476; Email: hkaneko@meiji.ac.jp

Author

Toshiharu Morishita − *Department of Applied Chemistry, School of Science and Technology, Meiji University, Kawasaki, Kanagawa 214-8571, Japan;* orcid.org/0000-0003-4150-2675

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.3c04891

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This study was supported by a Grant-in-Aid for Scientific Research (KAKENHI) (Grant Number 19K15352) from the Japan Society for the Promotion of Science.

REFERENCES

(1) Greenhill, S.; Rana, S.; Gupta, S.; Vellanki, P.; Venkatesh, S. Bayesian Optimization for Adaptive Experimental Design: A Review. *IEEE Access* **2020**, *8*, 13937−13948.

(2) Shahriari, B.; Swersky, K.; Wang, Z.; Adams, R. P.; De Freitas, N. Taking the Human Out of the Loop: A Review of Bayesian Optimization. *Proc. IEEE* **2016**, *104*, 148−175.

(3) Honarmandi, P.; Attari, V.; Arroyave, R. Accelerated Materials Design Using Batch Bayesian Optimization: A Case Study for Solving the Inverse Problem from Materials Microstructure to Process Specification. *Comput. Mater. Sci.* **2022**, *210*, No. 111417.

(4) Hanaoka, K. Bayesian Optimization for Goal-Oriented Multi-Objective Inverse Material Design. *iScience* **2021**, *24*, No. 102781.

(5) Park, S.; Na, J.; Kim, M.; Lee, J. M. Multi-objective Bayesian Optimization of Chemical Reactor Design Using Computational Fluid Dynamics. *Comput. Chem. Eng.* **2018**, *119*, 25−37.

(6) Griffiths, R. R.; Hernández-Lobato, J. M. Constrained Bayesian Optimization for Automatic Chemical Design Using Variational Autoencoders. *Chem. Sci.* **2020**, *11*, 577−586.

(7) Fang, L.; Makkonen, E.; Todorović, M.; Rinke, P.; Chen, X. Efficient Amino Acid Conformer Search with Bayesian Optimization. *J. Chem. Theory Comput.* **2021**, *17*, 1955−1966.

(8) Shields, B. J.; Stevens, J.; Li, J.; Parasram, M.; Damani, F.; Alvarado, J. I. M.; Janey, J. M.; Adams, R. P.; Doyle, A. G. Bayesian Reaction Optimization as a Tool for Chemical Synthesis. *Nature* **2021**, *590*, 89−96.

(9) Experimental Design via Bayesian Optimization. https://github.com/b-shields/edbo (accessed June 20, 2023).

(10) A Mordred Descriptor Calculator. https://github.com/mordred-descriptor/mordred (accessed June 20, 2023).

(11) Moriwaki, H.; Tian, Y. S.; Kawashita, N.; Takagi, T. Mordred: A Molecular Descriptor Calculator. *J. Cheminform.* **2018**, *10*, 1−14.

(12) Descriptor List. https://mordred-descriptor.github.io/documentation/master/descriptors.html (accessed June 20, 2023).

(13) Bayesian Reaction Optimization. https://b-shields.github.io/edbo/index.html (accessed June 20, 2023).

(14) Yang, K.; Palar, P. S.; Emmerich, M.; Shimoyama, K.; Bäck, T. et al. In *A Multi-point Mechanism of Expected Hypervolume Improvement for Parallel Multi-objective Bayesian Global Optimization*, Proceedings of the Genetic and Evolutionary Computation Conference, Association for Computing Machinery, Inc., 2019; pp 656−663.

(15) Gaussian Home Page. https://gaussian.com/ (accessed June 20, 2023).

(16) Codessa Pro QSPR/QSAR Software Home Page. http://www.codessa-pro.com (accessed June 20, 2023).

(17) Collins, E. M.; Raghavachari, K. Effective Molecular Descriptors for Chemical Accuracy at DFT Cost: Fragmentation, Error-Cancellation, and Machine Learning. *J. Chem. Theory Comput.* **2020**, *16*, 4938−4950.