# Failure-experiment-supported optimization of poorly reproducible synthetic conditions for novel lanthanide metal-organic frameworks

Yu Kitamura,[a] Emi Terado,[a] Zechen Zhang,[b] Hirofumi Yoshikawa,[b] Tomoko Inose,[c,d] Hiroshi Uji-i,[c,e] Masaharu Tanimizu,[f] Akihiro Inokuchi,[g] Yoshinobu Kamakura,[a] and Daisuke Tanaka*[a,h]

[a] Y. Kitamura, E. Terado, Y. Kamakura, Prof. D. Tanaka
Department of Chemistry, School of Science and Technology,
Kwansei Gakuin University
2-1 Gakuen, Sanda, Hyogo 669-1337 (Japan)
E-mail: dtanaka@kwansei.ac.jp

[b] Z. Zhang, Prof. H. Yoshikawa
Department of Nanotechnology for Sustainable Energy, School of Science and Technology,
Kwansei Gakuin University
2-1 Gakuen, Sanda, Hyogo 669-1337 (Japan)

[c] Dr. T. Inose, Prof. H. Uji-i
Division of Photonics and Optical Science, Research Institute for Electronic Science (RIES),
Hokkaido University
North 20 West 10, Kita ward, Sapporo, Hokkaido 001-0020 (Japan)

[d] Dr. T. Inose
Institute for Integrated Cell-Material Sciences (WPI-iCeMS),
Kyoto University
Yoshida, Sakyo-ku, Kyoto 606-8501 (Japan)

[e] Prof. H. Uji-i
Department of Chemistry,
Katholieke Universiteit Leuven, Celestijnenlaan 200F, Heverlee, 3001 (Belgium)

[f] Prof. M. Tanimizu
Department of Applied Chemistry for Environment, School of Science and Technology,
Kwansei Gakuin University
2-1 Gakuen, Sanda, Hyogo 669-1337 (Japan)

[g] Prof. A. Inokuchi
Department of Informatics, School of Science and Technology,
Kwansei Gakuin University
2-1 Gakuen, Sanda, Hyogo 669-1337 (Japan)

[h] Prof. D. Tanaka
JST PRESTO
2-1 Gakuen, Sanda, Hyogo 669-1337 (Japan)

Supporting information for this article is given via a link at the end of the document.

**Abstract:** A series of novel metal-organic frameworks with lanthanide double-layer-based inorganic subnetworks (KGF-3) was synthesized assisted by machine learning. Pure KGF-3 was difficult to isolate in the initial screening experiments. The synthetic conditions were successfully optimized by extracting the dominant factors for the synthesis of KGF-3 using two machine-learning techniques. Cluster analysis was used to classify the obtained PXRD patterns of the products and to decide automatically whether the

experiments were successful or had failed. Decision tree analysis was used to visualize the experimental results, with the factors that mainly affected the synthetic reproducibility being extracted. The water adsorption isotherm revealed that KGF-3 possesses unique hydrophilic pores, and impedance measurements demonstrated good proton conductivities ($\sigma = 5.2 \times 10^{-4}$ S cm$^{-1}$ for KGF-3(Y)) at a high temperature (363 K) and high relative humidity (95%).

## Introduction

Crystalline nanoporous materials, such as metal–organic frameworks (MOFs), covalent organic frameworks (COFs), and zeolites, have been widely studied as promising environmental energy materials.[1-6] The crystallization of these materials has been mainly achieved by solvothermal or hydrothermal synthesis in sealed reaction vessels.[1,4,7] However, the discovery of new materials by such synthetic methods unfortunately requires patient trial and error rather than precise reaction design. It is therefore desirable to understand the crystallization mechanism of nanoporous materials under solvothermal conditions and thereby to develop a rational approach to the selective crystallization of target compounds.[8-11] Although many in situ measurements have been carried out to understand the crystallization process in a sealed reaction vessel, it remains a challenge to determine the relationship between the reaction parameters and the outcome of a reaction. Facile and reliable techniques to understand and control the target reaction, which can be applicable during the screening of a synthesis process, are therefore necessary to efficiently discover novel crystalline nanoporous materials.

Recently, machine learning has attracted attention as an efficient exploration tool, especially in the area of materials synthesis.[12-14] In particular, several attempts to use machine learning to search for the crystallization conditions of materials have been reported, and it is beginning to be regarded as a powerful method.[15-22] While the application of machine learning techniques to predict crystallization condition for nanoporous materials seems promising, its application in the exploration of novel materials remains limited, and no studies have introduced machine learning as a tool for the preparation of unknown porous MOFs. This is partially due to a lack of training data; open databases for the exploration of unknown MOFs are limited and the generation of such training data is expensive.[21] In addition, during the

exploration of novel MOFs by screening synthesis, the compounds obtained in failure experiments are generally present as complex mixtures. Therefore, facile techniques to describe such crude mixtures in a format suitable for analysis by machine learning is necessary.

In this work, we focus on lanthanide-based MOFs (Ln-MOFs) in which lanthanide ions or clusters are linked by organic linkers. Ln-MOFs are promising materials for a wide range of applications, such as luminescent materials, proton conductors, and magnetic materials, as well as porous materials.[23-27] Furthermore, lanthanides have been reported to form giant clusters under solvothermal synthetic conditions,[28-32] and so multiple or novel properties are expected to arise due to synergism when such large clusters or more highly dimensional infinite structures are incorporated into the skeletons of MOFs.[4,33,34] However, since the formation of lanthanide clusters is easily affected by the reaction conditions, the rational design and syntheses of MOFs containing Ln-cluster-based subnetworks remains a challenge. In addition, Ln-MOFs are also known to provide many crystal polymorphisms due to their flexible coordination nature. It is therefore difficult to selectively synthesize the crystal polymorphs, and the preparation of Ln-MOFs frequently suffers from a poor reproducibility. A facile and intuitive means to evaluate the factors that dominantly influence the reaction outcome is therefore required.

We herein report the synthesis of a novel Ln-MOF containing lanthanide-double-layer-based subnetworks (KGF-3) assisted by machine learning. This is the first example of applying machine learning to the solvothermal synthesis for the exploration of unknown MOFs. The synthetic results are evaluated using both cluster analysis and decision tree analysis. These analyses will enable us to determine the optimal conditions for the reproducible synthesis of KGF-3. Figure 1 shows the flow process used to optimize the synthetic conditions using machine-learning techniques. Water adsorption experiments and impedance measurements are also used to analyze the pores and the proton conductivity of the prepared KGF-3.

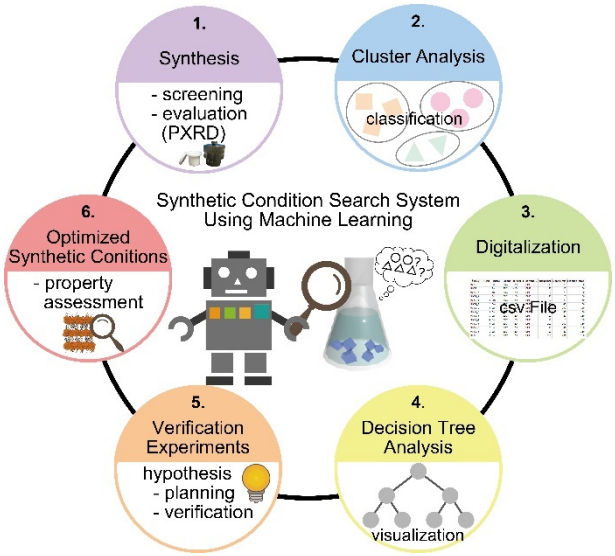


**Fig. 1** Schematic representation of the research flow process used to optimize the synthetic conditions by incorporating machine-learning methods.

## Results and Discussion

### Screening synthesis

Solvothermal synthesis was performed using nitrate hexahydrate salts of lanthanide metals ($Ln^{3+}$=Sm, Eu, Gd, and Tb) in the presence of terephthalic acid (BDC) in $H_2O$/DMF (DMF=*N,N*-dimethylformamide). A total of 108 experiments were carried out, with variables including the lanthanide metal, the concentration of the metal and/or ligand solution, the reaction temperature and time, the cooling time, and the type of reaction vessel (see Synthetic Conditions 1 in Methods). Solid powders were obtained under all conditions, and were characterized by powder X-ray diffractometry (PXRD), where, in some cases, novel phases were observed.

The novel $[Ln_{10}(BDC)_3(HCOO)_4(\mu_3\text{-}OH)_{12}(\mu_5\text{-}CO_3)_4(H_2O)_2]$ phase, which we refer to as "KGF-3", was obtained in the presence of various lanthanide ions, although single crystals suitable for crystal structure analysis were obtained only when $Gd^{3+}$ was used as the metal source. Based on the single-crystal X-ray data, KGF-3(Gd) was found to contain five types of non-equivalent $Gd^{3+}$ ions with coordination numbers of eight or nine. Four complexes share ridges with each other to form a chain, which is then

connected by another lanthanide ion to form a porous layer (Figure 2a). The two layers are cross-linked by carbonate ions (the origin of the carbonates will be discussed later) to form a double layer, and BDC bridges the layers as pillar molecules, resulting in a three-dimensional pillar-layered structure (Figures 2b–2c). The $\mu_5$-$CO_3^{2-}$ coordination mode is common in giant cluster synthesis but is unusual in MOFs (Figure 2d). Moreover, we found that the formic acid generated by the decomposition of DMF was also coordinated. Many $\mu$-OH groups are aligned on the KGF-3 pore surfaces, and disordered guest molecules (most likely water) are incorporated in the pores, suggesting that the pores are highly hydrophilic in nature.

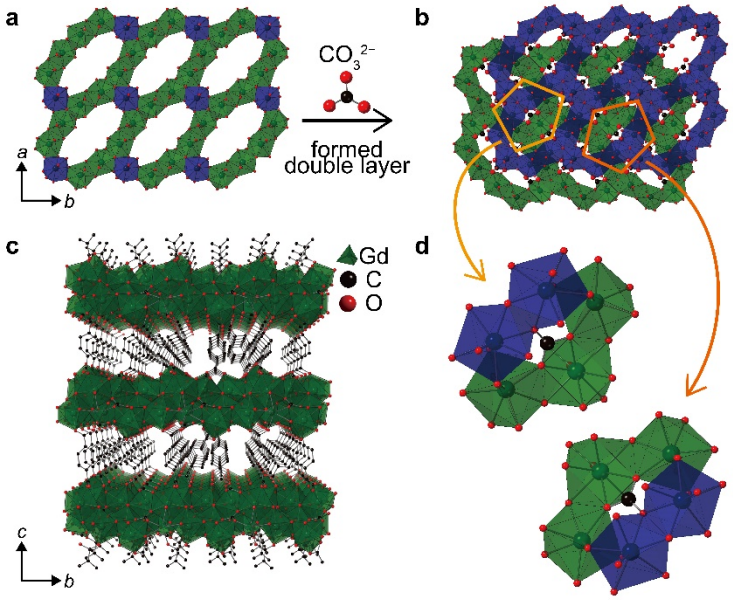


**Fig. 2** (a) Monolayer with 10 spread-out lanthanide clusters. (b) Double layer connected by carbonate ions. (c) KGF-3(Gd) viewed along the a-axis. H atoms are omitted for clarity. (d) A pentagonal pocket connected by a carbonate ion.

**Analysis by machine learning techniques**

KGF-3 was difficult to isolate, and its preparation suffered from a poor reproducibility, with different phases frequently being obtained even under the same synthetic conditions. In addition, a pure KGF-3 phase was not obtained after 108 experimental trials. To optimize the synthetic conditions, we extracted the dominant factors of the reaction using machine-learning techniques. To predict the conditions under which KGF-3 can be obtained more reproducibly, it was necessary to determine whether or not the

reaction was successful using PXRD. However, the products are complex mixtures of different phases in many cases; hence, assigning each PXRD pattern to an appropriate phase is challenging. A simple method that excluded arbitrariness was therefore required; hence, we classified the obtained patterns using cluster analysis.[35] All diffraction patterns were automatically analyzed and successfully classified into six categories, in which the main products were KGF-3 (cluster 1), four reported phases (clusters 2,[36] 3,[37], 4,[38] and 5[39]), and another unknown phase (cluster 6), as shown in Figures 3 and S1. By examining these results, we revealed that automated classification was consistent with "researcher intuition," with the exception of only two data points in 108 patterns (see the Cluster Analysis section in the SI).

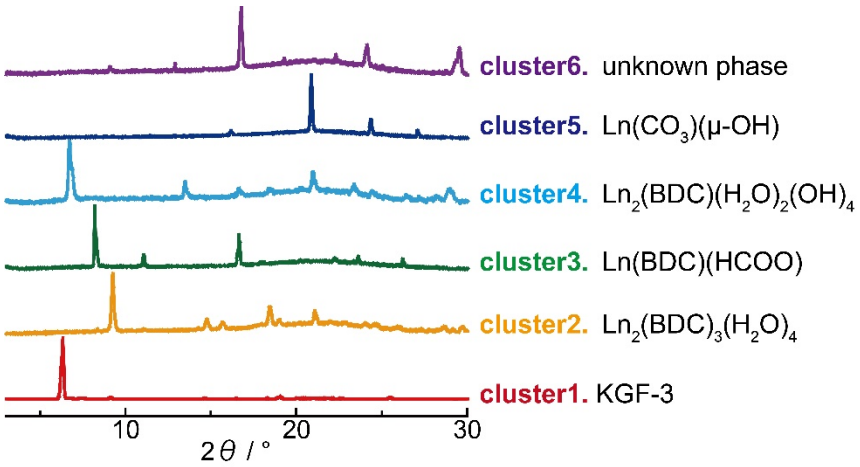


**Fig. 3** Using cluster analysis to classify the acquired PXRD patterns.

We next turned our attention to extracting the dominant factors responsible for the poor reproducibility using decision tree analysis, which is considered to be one of the most interpretable machine-learning techniques.[15,20,22] Initially, the experimental data and cluster analysis results were linked together in a text file, after which the data file was analyzed using the decision tree technique, where the objective variables were the crystal phases assigned by cluster analysis of the PXRD patterns, and the explanatory variables were the synthetic parameters (see the Decision Tree Analysis section in the SI). The results presented in Figure 4a suggest that the most suitable synthetic conditions for the preparation of KGF-3 are as follows: Ligand solution concentration, 18–22 mM; cooling time, >12 h; and metal salt source, company A. The parameters appearing at the branches in the decision tree were also suggested to be important based on random forest analysis (Figure 4b). Thus, visualization of the experimental data by decision tree analysis

allowed us to understand the synthetic conditions at a glance. The information extracted from the decision tree is summarized as follows. Firstly, it is likely that when the ligand concentration is <18 mM, a product is formed with bridging carbonate ligands (cluster 5) (Figure 4a-(1)), suggesting that carbonate is generated by the decomposition of DMF or through capture from the air.[40,41] Secondly, the success or failure of the KGF-3 synthesis was determined by the reagent company employed, with the nitrate provided by company A being most suitable (Figure 4a-(2)). Finally, as shown in Figure 4a-(3), the success rate was 17% on the left branch (Eu) and 45% on the right branch (Gd and Tb), suggesting that the metal ion affects the synthetic process. To determine the synthetic conditions that maximize reproducibility, the extracted dominant factors were verified as follows.

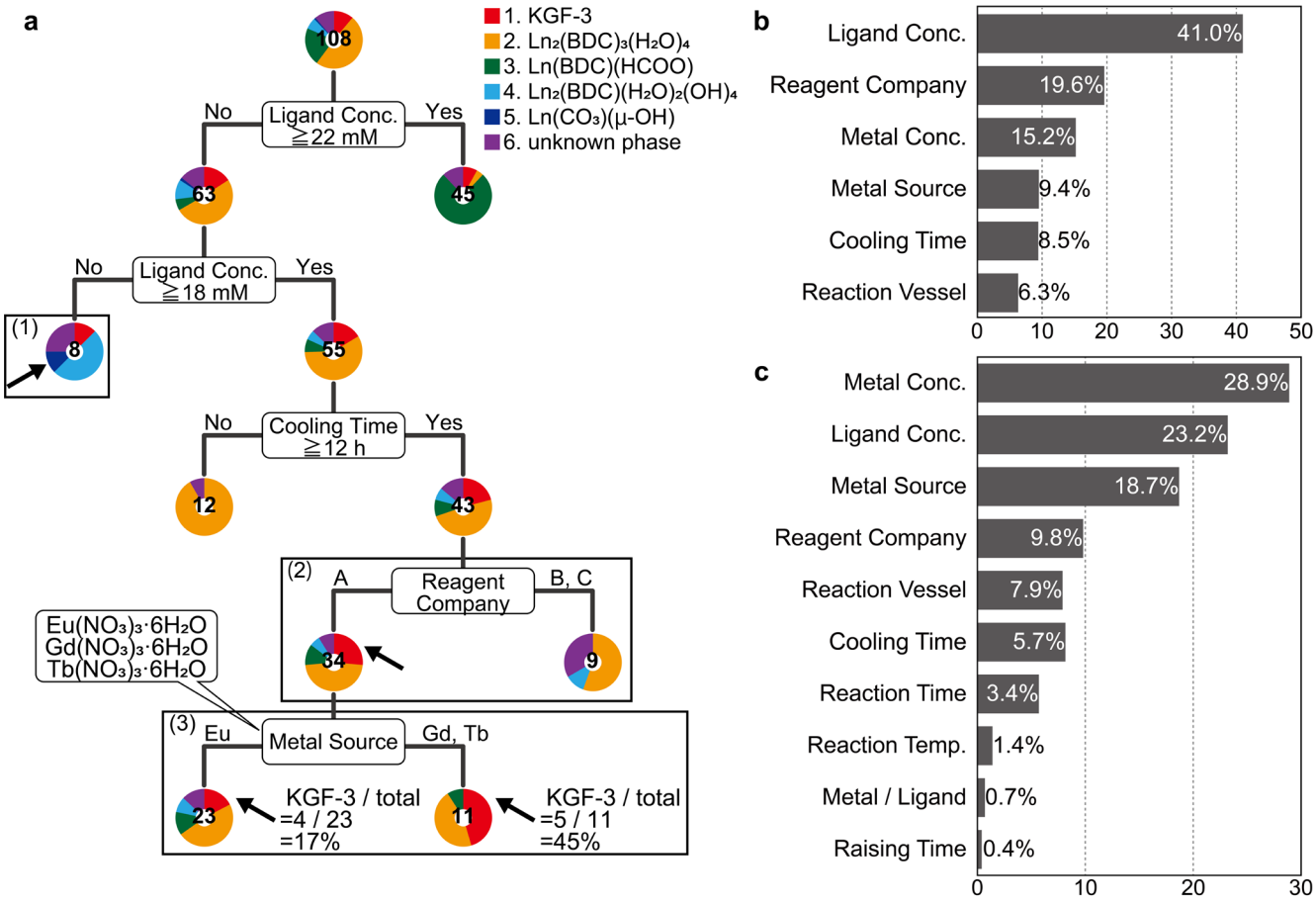


**Fig. 4 a** Visualizing the relationships between the experimental conditions and the products, based on decision tree analysis. The ovals show the decision nodes, the pie charts show the product ratios, and the number of experiments is shown in the center of each pie chart. (1), (2), and (3) are branches related to

the concentration conditions, the reagent company, and the lanthanide metal, respectively. Branches that are not discussed in the main text have been omitted. The complete version of the decision tree is given in Figure S2. The percentage of each parameter that contributes to the branch is shown. **b** Decision tree analysis. **c** Random forest analysis.

**Exploration of the optimized synthetic conditions**

In terms of the supply company for the lanthanide nitrate, the reagents purchased from company A were superior (Figure S3), exhibiting a higher purity (99.95%) compared to those obtained from companies B (99.9%) and C (99.5%). The purities guaranteed by the reagent companies were evaluated based only on the metal ion concentration. To estimate the influence of the purity of the metal source, $Tb(NO_3)_3 \cdot 6H_2O$ with the highest purity (99.999%), i.e., from company D, was also used, which gave an improved success rate compared to that obtained using $Tb(NO_3)_3 \cdot 6H_2O$ from company A; the success rate was 20% (7 out of 34 trials) using the Tb source from company A while the success rate increased to 95% using the Tb source from company D (19 out of 20 trials), as shown in Figure S4. The different purities were evaluated by inductively coupled plasma-mass spectrometry, which revealed that the $Tb(NO)_3 \cdot 6H_2O$ obtained from company A contained slightly higher levels of Eu than that from company D (Table S2). Therefore, the preparation of KGF-3 was carried out using lanthanide salts purchased from company D in all of the following experiments.

The results from decision tree analysis also showed that under low concentration conditions, carbonate ions are captured within the structure (cluster 5). Generally, carbonate ions play important roles as anion templates during the formation of polynuclear lanthanide clusters.[32,42-44] In many cases, carbonate ions are generated by the decomposition of the precursor and/or the uptake of carbon dioxide from the air. Therefore, in the crystal structure of KGF-3, it is likely that the molecules coordinating to five metal ions in a pentagonal pocket (Figure 2d) are carbonate ions. To estimate the effect of these carbonate ions, KGF-3 was synthesized with the addition of sodium carbonate (Synthetic Conditions 2). For Gd and Tb, the formation of the $Ln_2(BDC)_3(H_2O)_4$ impurity (cluster 2) was suppressed with the addition of carbonate

ions. In the case of Eu, $Ln_2(BDC)_3(H_2O)_4$ (cluster 2) was preferentially synthesized, regardless of whether carbonate ions were added or not (Figure 5). These results suggest that the effect imparted by the carbonate ions depends on the lanthanide metal ion.

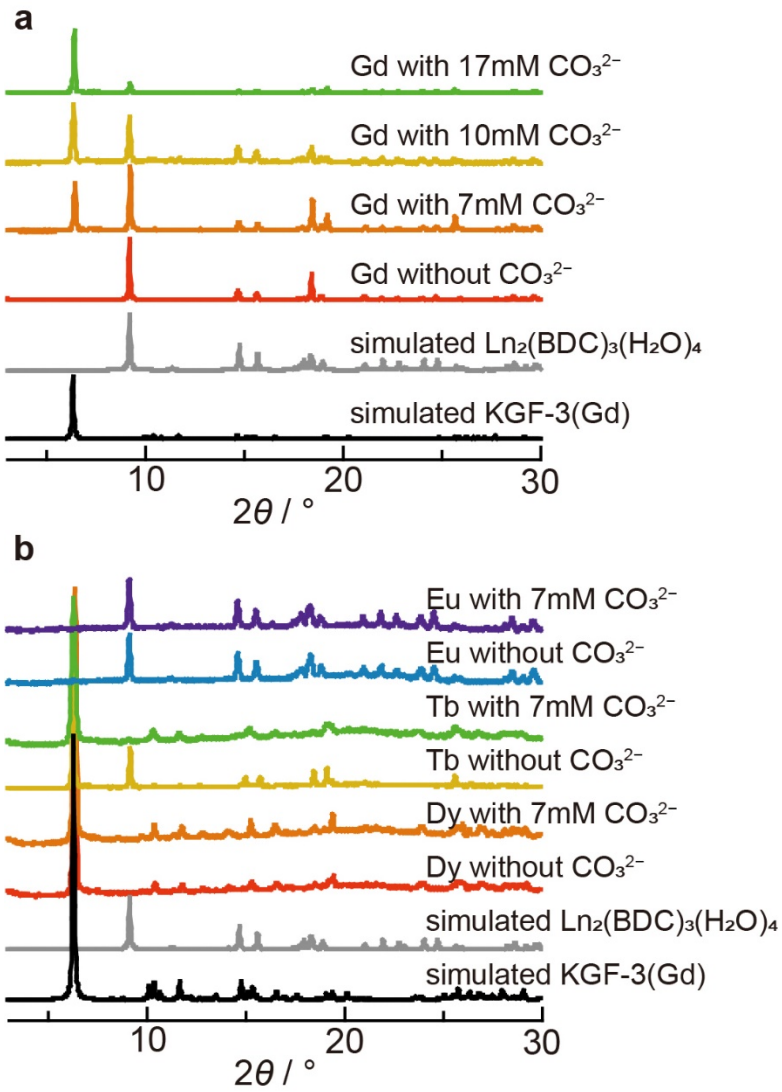


**Fig. 5** PXRD patterns obtained in the presence of sodium carbonate. **a** The results of adding various concentrations of sodium carbonate to Gd. **b** comparisons with and without the addition of sodium carbonate for Eu, Tb, and Dy.

The decision tree analysis (Figure 4a-(3)) and the response to the addition of carbonate ions strongly suggest that the metal ion affects the probability of successfully synthesizing KGF-3. We therefore synthesized KGF-3(Ln) with various lanthanide metal ions (La–Lu, excluding Pm) under optimized Synthetic Conditions 3, 4, 5 and 6 (Figures 6 and S5). Based on PXRD and elemental analysis (EA) data,

it was found that only Dy and Ho provided pure isolated KGF-3, and only the EA data for Dy (ionic radius $r$ = 1.03 Å), and Ho ($r$ =1.02 Å) matched the calculated values, which suggests that the ionic radius is important for the selective synthesis of pure KGF-3, and that the optimal ionic radius is 1.02–1.03 Å. The EA results show that no nitrogen atom exists in KGF-3 (Table S4), thereby confirming that all three directional ligands in the crystal structure are carbonate ions (see the Elemental Analysis section in the SI). The observation that the presence of trace amounts of Eu in the purchased reagent render the experiments more likely to fail coincides with the fact that other crystal phases form with high probabilities using Eu. In many cases, the ionic radius of the lanthanide ion strongly influences the construction of the obtained Ln complex.[45,46] In fact, KGF-3(Y) was successfully isolated when $Y^{3+}$ ($r$ =1.02 Å), whose ionic radius is close to those of $Dy^{3+}$ and $Ho^{3+}$ (Figure 6), was used, which strongly suggests that the ionic radius of the lanthanide metal is a crucial and dominant factor for KGF-3 formation.

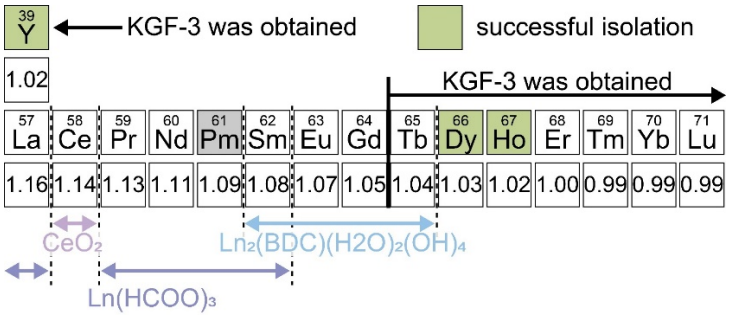


**Fig. 6** Summary of the synthetic results for the KGF-3 under Synthetic Conditions 3. See Figure S5 for details.

**Adsorption properties and proton conductivity measurements of KGF-3**

To evaluate the permanent porosity of KGF-3(Dy, Ho, and Y), $N_2$ and $H_2O$ adsorption isotherms were acquired (Figures 7a and S8), which suggests that KGF-3 adsorbs $H_2O$ molecules into its pores, whereas $N_2$ is not adsorbed; hence KGF-3 is likely to possess narrow hydrophilic channels whose diameters are too small for nitrogen diffusion at 77 K. This result is consistent with the obtained crystal structure of KGF-3(Gd). Although the water molecules are disordered within the pores, they are expected to form a pathway for proton conduction, with the water assembly being stabilized by the hydrophilic pore surface (Figure S10). The hydrophilic nature of KGF-3 therefore prompted us to evaluate its proton conductivity.

Thus, the alternating current impedances of KGF-3(Dy, Ho, and Y) were measured at 313–363 K and at 95% relative humidity (Figures 7b and S11). KGF-3(Y) showed the highest proton conductivity among the three MOFs, i.e., $5.2 \times 10^{-4}$ S $cm^{-1}$ at 363 K, and KGF-3 retained its crystalline nature after the impedance experiments or even after soaking in water (Figures S9 and S12). The conductivity of KGF-3(Y) was observed to increase with increasing temperature due to thermal activation of the water molecules. The activation energy for proton conduction was calculated to be 0.65 eV at low temperatures (40–343 K), whereas at high temperatures (343–363 K), it was 0.14 eV (Figure 7c), suggesting that changes in the transport mechanism occurs at ~343 K.[47,48]

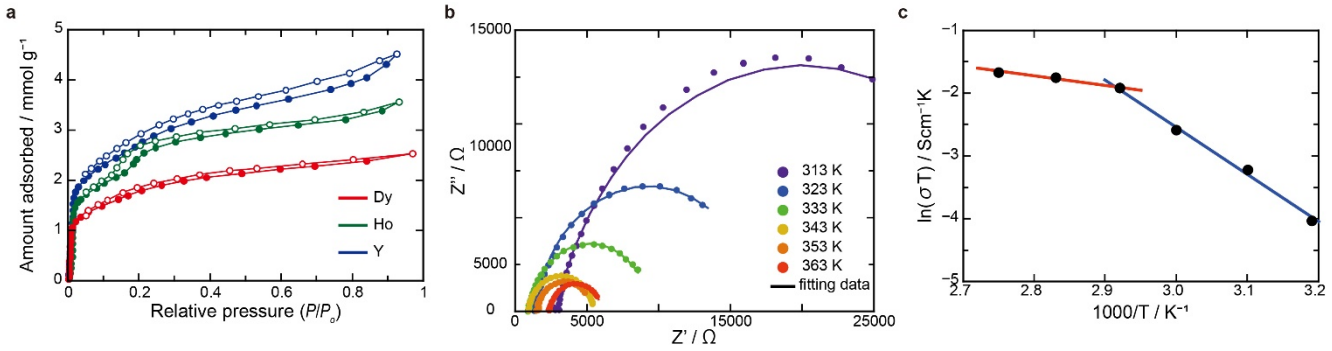


**Fig. 7 a** Adsorption and desorption isotherms for $H_2O$. The solid and open symbols correspond to adsorption and desorption, respectively. **b** Nyquist plots at various temperatures and at 95% RH for a pellet sample of KGF-3(Y). **c** Arrhenius plot of the ion conductivity at 95% RH of KGF-3(Y).

## Conclusion

In summary, we successfully synthesized a series of novel pillar-layered Ln-MOFs (metal-organic frameworks) containing lanthanide double-layer-based inorganic subnetworks that we refer to as "KGF-3." Although it was difficult to isolate KGF-3 in our initial screening experiments due to the poor synthetic reproducibility, we successfully extracted the dominant factors for KGF-3 synthesis by evaluating both failure and success using machine-learning techniques. The extracted chemical insight suggests that the lanthanide ion affects the synthetic results, and systematic synthesis experiments demonstrated the effect of the ionic radius of the metal ion. This method is a useful tool for preparing new MOFs and related

compounds, such as coordination polymers and covalent organic frameworks that suffer from poor synthetic reproducibilities. Through the application of this method, the exploration of novel MOFs and coordination polymers, for which it is challenging to obtain highly crystalline samples, is currently underway, and the results will be presented in due course.

## Methods

### General characterization and analytical methods

See the Supplementary Information for further details.

### Synthetic conditions 1 (exploratory experiments)

$Ln(NO_3)_3 \cdot 6H_2O$ (Ln = Sm, Eu, Gd, Tb, 0.008–0.8 mmol) was dissolved in distilled water (2–200 mM), and $H_2BDC$ (0.008–0.8 mmol) was separately dissolved in DMF (2–200 mM). These two solutions were mixed in a Teflon-lined stainless-steel container (4, 8, 16, 30, or 100 mL) and heated at 145 or 150 °C for 24–80 h. At the end of the heating process, the container was cooled to 30 °C. The heating time was either 5 or 12 h, and the cooling time was either rapid cooling or cooling over 12–80 h.

### Synthetic conditions 2 (carbonate ion-added synthesis)

$Ln(NO_3)_3 \cdot 6H_2O$ (Ln = Eu, Gd, Tb) was dissolved in distilled water (0.08 mmol, 2.4 mL) and $H_2BDC$ was separately dissolved in DMF (0.08 mmol, 1.6 mL), while $Na_2CO_3$ was dissolved in distilled water (0–0.036 mmol, 2.0 mL). These three solutions were mixed in an 8 mL Teflon-lined stainless-steel container and heated at 150 °C for 48 h. At the end of the heating process, the container was cooled to 30 °C. The heating time was 5 h, and the cooling time was 12 h.

### Synthetic conditions 3 (optimized synthetic conditions A: stoichiometry synthesis)

$Ln(NO_3)_3 \cdot xH_2O$ (Ln = La, Ce, Pr, Nd, Sm, Eu, Gd, Tb, Dy, Ho, Er, Tm, Yb, Lu, or Y) was dissolved in distilled water (0.16 mmol, 4.8 mL) and $H_2BDC$ was separately dissolved in DMF (0.048 mmol, 3.2 mL).

These two solutions were mixed in a 16 mL Teflon-lined stainless-steel container and heated at 150 °C for 48 h. At the end of the heating process, the container was cooled to 30 °C. The heating time was 5 h, and the cooling time was 12 h. The obtained residue was washed with DMF and MeOH (×3 for each solvent).

**Synthetic conditions 4 (optimized synthetic conditions B; preparation using excess BDC)**

$Ln(NO_3)_3 \cdot xH_2O$ (Ln = La, Ce, Pr, Nd, Sm, Eu, Gd, Tb, Dy, Ho, Er, Tm, Yb, Lu, or Y) was dissolved in distilled water (0.16 mmol, 4.8 mL) and $H_2BDC$ was separately dissolved in DMF (0.16 mmol, 3.2 mL). These two solutions were mixed and separated by centrifugation, and the supernatant was placed in a 16 mL Teflon-lined stainless-steel container and heated at 150 °C for 48 h. At the end of the heating process, the container was cooled to 30 °C. The heating time was 5 h, and the cooling time was 12 h. The obtained residue was washed with DMF and MeOH (×3 for each solvent).

**Synthetic conditions 5 (optimized synthetic conditions for KGF-3(Eu, Gd)**

$Ln(NO_3)_3 \cdot 6H_2O$ (Ln = Eu, Gd) was dissolved in distilled water (0.08 mmol, 2.4 mL) and $H_2BDC$ was separately dissolved in DMF (0.024 mmol, 1.6 mL), while $Na_2CO_3$ was dissolved in distilled water (0.04 mmol, 2.0 mL). These three solutions were mixed in an 8 mL Teflon-lined stainless-steel container and heated at 150 °C for 6 h. At the end of the heating process, the container was cooled to 30 °C. The heating time was 5 h, and the cooling time was 12 h. The residue was washed with DMF and MeOH (×3 for each solvent).

**Synthetic conditions 6 (optimized synthetic conditions for KGF-3(Tb)**

$Tb(NO_3)_3 \cdot 6H_2O$ was dissolved in distilled water (0.08 mmol, 2.4 mL) and $H_2BDC$ was separately dissolved in DMF (0.024 mmol, 1.6 mL). These two solutions were mixed in an 8 mL Teflon-lined stainless-steel container and heated at 150 °C for 48 h. At the end of the heating process, the container

was cooled to 30 °C. The heating time was 5 h, and the cooling time was 12 h. The obtained residue was washed with DMF and MeOH (×3 for each solvent).

**Preparation of a single crystal of KGF-3 (Gd)**

$Gd(NO_3)_3 \cdot 6H_2O$ was dissolved in distilled water (0.08 mmol, 2.4 mL) and $H_2BDC$ was separately dissolved in DMF (0.024 mmol, 1.6 mL), while $Na_2CO_3$ was dissolved in distilled water (0.04 mmol, 2.0 mL). These three solutions were mixed in an 8 mL Teflon-lined stainless-steel container and heated at 150 °C for 6 h. At the end of the heating process, the container was cooled to 30 °C. The heating time was 5 h, and the cooling time was 12 h.

**Machine-learning analysis**

Custer analysis was carried out using PDXL 2.8 (Rigaku). Decision tree analysis (Partition) was carried out using JMP® Pro14.3, and random forest analysis (Bootstrap Forest) was carried out using JMP® Pro 15.2.1 (SAS Institute Inc., Cary, NC, USA).

**Acknowledgements**

This work was supported by grants from JST PRESTO (Grant No. JPMJPR17NA) and JSPS KAKENHI (Grant Nos. 20H02577, 17K00320, 20H04680, 20H04646). We thank Dr. Satoki Okabayashi, Prof. Koichi Chiba and Dr. Akane Ito (KwanseiGakuin University) for ICP−OES support. This work was performed under the Cooperative Research Program of "Network Joint Research Center for Materials and Devices." Prof. Masaharu Tanimizu was supported by a Joint Research Grant for the ICP−MS experiments from the Environmental Isotope Study of the Research Institute for Humanity and Nature.

**Author contributions**

Y. K. and E. T. carried out the synthesis. D. T. conceived the experiments and supervised the project. Z. Z. and H. Y. measured the proton conductivities and H. Y. carried out the $H_2O$ adsorption experiments.

T. I. and H. U. acquired the Raman data and M. T. acquired the ICP-MS data. Y. Kamakura performed the single crystal X-ray diffraction analysis. A. I. discussed the machine-learning method.

**Keywords:** lanthanide • machine learning • metal-organic framework • proton conductivity • solvothermal synthesis

**References**

1 Cundy, C. S. & Cox, P. A. The hydrothermal synthesis of zeolites: History and development from the earliest days to the present time. *Chem. Soc. Rev.,* **103**, 663-701 (2003).
2 Hooper, D. *et al.* An in situ study of crystallisation gradients during the hydrothermal/autoclave synthesis of zeolites. *Phys. Chem. Chem. Phys.* **5**, 4946-4950 (2003).
3 Kitagawa, S., Kitaura, R. & Noro, S. Functional porous coordination polymers. *Angew. Chem. Int. Ed.* **43**, 2334-2375 (2004).
4 Ferey, G. Hybrid porous solids: past, present, future. *Chem. Soc. Rev.* **37**, 191-214 (2008).
5 Das, S., Heasman, P., Ben, T. & Qiu, S. Porous organic materials: strategic design and structure-function correlation. *Chem. Rev.* **117**, 1515-1563 (2017).
6 Kumar, K. V., Preuss, K., Titirici, M. M. & Rodriguez-Reinoso, F. Nanoporous materials for the onboard storage of natural gas. *Chem. Rev.* **117**, 1796-1825 (2017).
7 Kumar, M. *et al.* Lanthanide contraction in action: structural variations in 13 lanthanide(III) thiophene-2,5-dicarboxylate coordination polymers (Ln = La–Lu, except Pm and Tm) featuring magnetocaloric effect, slow magnetic relaxation, and luminescence-lifetime-based thermometry. *Cryst. Growth Des.* **20**, 6430-6452 (2020).
8 Goesten, M. G. *et al.* Evidence for a chemical clock in oscillatory formation of UiO-66. *Nat. Commun.* **7**, 11832 (2016).
9 Saha, S. *et al.* Insight into fast nucleation and growth of zeolitic Imidazolate framework-71 by in situ time-resolved light and X-ray scattering experiments. *Cryst. Growth Des.* **16**, 2002-2010 (2016).
10 Yeung, H. H. *et al.* In situ observation of successive crystallizations and metastable intermediates in the formation of metal-organic frameworks. *Angew. Chem. Int. Ed.* **55**, 2012-2016 (2016).
11 Van Vleet, M. J., Weng, T., Li, X. & Schmidt, J. R. In situ, time-resolved, and mechanistic studies of metal-organic framework nucleation and growth. *Chem. Rev.* **118**, 3681-3721 (2018).
12 Yang, X., Wang, Y., Byrne, R., Schneider, G. & Yang, S. Concepts of artificial intelligence for computer-assisted drug discovery. *Chem. Rev.* **119**, 10520-10594 (2019).
13 Coley, C. W., Eyke, N. S. & Jensen, K. F. Autonomous discovery in the chemical sciences Part I: Progress. *Angew. Chem. Int. Ed.* **59**, 22858-22893 (2020).
14 Moosavi, S. M., Jablonka, K. M. & Smit, B. The role of machine learning in the understanding and design of materials. *J. Am. Chem. Soc.* **142**, 20273−20287 (2020).
15 Raccuglia, P. *et al.* Machine-learning-assisted materials discovery using failed experiments. *Nature* **533**, 73-76 (2016).
16 Duros, V. *et al.* Human versus robots in the discovery and crystallization of gigantic polyoxometalates. *Angew. Chem. Int. Ed.* **56**, 10815-10820 (2017).
17 Greenaway, R. L. *et al.* High-throughput discovery of organic cages and catenanes using computational screening fused with robotic synthesis. *Nat. Commun.* **9**, 2849 (2018).
18 Duros, V. *et al.* Intuition-enabled machine learning beats the competition when joint human-robot teams perform inorganic chemical experiments. *J. Chem. Inf. Model.* **59**, 2664-2671 (2019).
19 Moosavi, S. M. *et al.* Capturing chemical intuition in synthesis of metal-organic frameworks. *Nat. Commun.* **10**, 539 (2019).

20 Muraoka, K., Sada, Y., Miyazaki, D., Chaikittisilp, W. & Okubo, T. Linking synthesis and structure descriptors from a large collection of synthetic records of zeolite materials. *Nat. Commun.* **10**, 4459 (2019).
21 Jablonka, K. M., Ongari, D., Moosavi, S. M. & Smit, B. Big-data science in porous materials: materials genomics and machine learning. *Chem. Rev.* **120**, 8066-8129 (2020).
22 Xie, Y. *et al.* Machine learning assisted synthesis of metal-organic nanocapsules. *J. Am. Chem. Soc.* **142**, 1475-1481 (2020).
23 Rocha, J., Carlos, L. D., Paz, F. A. & Ananias, D. Luminescent multifunctional lanthanides-based metal-organic frameworks. *Chem. Soc. Rev.* **40**, 926-940 (2011).
24 Roy, S., Chakraborty, A. & Maji, T. K. Lanthanide-organic frameworks for gas storage and as magneto-luminescent materials. *Coord. Chem. Rev.* **273**, 139-164 (2014).
25 Lustig, W. P. *et al.* Metal-organic frameworks: functional luminescent and photonic materials for sensing applications. *Chem. Soc. Rev.* **46**, 3242-3285 (2017).
26 Meng, X., Wang, H. N., Song, S. Y. & Zhang, H. J. Proton-conducting crystalline porous materials. *Chem. Soc. Rev.* **46**, 464-480 (2017).
27 Saraci, F., Quezada-Novoa, V., Donnarumma, P. R. & Howarth, A. J. Rare-earth metal-organic frameworks: from structure to applications. *Chem. Soc. Rev.* **49**, 7949-7977 (2020).
28 Peng, J. B. *et al.* Beauty, symmetry, and magnetocaloric effect-four-shell keplerates with 104 lanthanide atoms. *J. Am. Chem. Soc.* **136**, 17938-17941 (2014).
29 Dong, J., Cui, P., Shi, P. F., Cheng, P. & Zhao, B. Ultrastrong alkali-resisting lanthanide-zeolites assembled by [Ln60] nanocages. *J. Am. Chem. Soc.* **137**, 15988-15991 (2015).
30 Wang, K. *et al.* A single-stranded {Gd18} nanowheel with a symmetric polydentate diacylhydrazone ligand. *Chem. Commun.* **52**, 8297-8300 (2016).
31 Wang, J., Feng, M., Akhtar, M. N. & Tong, M. L. Recent advance in heterometallic nanomagnets based on TM(x)Ln(4-x) cubane subunits. *Coord. Chem. Rev.* **387**, 129-153 (2019).
32 Zheng, X.-Y., Xie, J., Kong, X.-J., Long, L.-S. & Zheng, L.-S. Recent advances in the assembly of high-nuclearity lanthanide clusters. *Coord. Chem. Rev.* **378**, 222-236 (2019).
33 Devic, T. & Serre, C. High valence 3p and transition metal based MOFs. *Chem. Soc. Rev.* **43**, 6097-6115 (2014).
34 Miras, H. N., Vila-Nadal, L. & Cronin, L. Polyoxometalate based open-frameworks (POM-OFs). *Chem. Soc. Rev.* **43**, 5679-5699 (2014).
35 Butler, B. M. *et al.* Pre-treatment of soil X-ray powder diffraction data for cluster analysis. *Geoderma* **337**, 413-424 (2019).
36 Reineke, T. M., Eddaoudi, M., Fehr, M., Kelley, D. & Yaghi, O. M. From condensed lanthanide coordination solids to microporous frameworks having accessible metal sites. *J. Am Chem. Soc.* **121**, 1651-1657 (1999).
37 Huang, G., Yang, P., Wang, N., Wu, J.-Z. & Yu, Y. First lanthanide coordination polymers with N,N-dimethylformamide hydrolysis induced formate ligands. *Inorg. Chim. Acta* **384**, 333-339 (2012).
38 Serre, C., Millange, F., Marrot, J. & Ferey, G. Hydrothermal synthesis, structure determination, and thermal behavior of new three-dimensional europium terephthalates: MIL-51(LT,HT) and MIL-52 or EU2n(OH)(x)(H2O)(y)(O2C-C6H4-CO2)(z) (n = III, III, II; x = 4, 0, 0; y = 2, 0, 0; z = 1, 1, 2). *Chem. Mater.* **14**, 2409-2415 (2002).
39 Tian, H. *et al.* Quadruple-CO3(2-) bridged octanuclear dysprosium(III) compound showing single-molecule magnet behaviour. *Chem. Commun.* **48**, 708-710 (2012).
40 Zhang, B. *et al.* Efficient fixation of atmospheric CO2 as carbonate by lanthanide-based complex via synergistic effect of zinc ion. *Dalton Trans.* **42**, 8571-8574 (2013).
41 Chen, H.-M. *et al.* Luminescence and magnetocaloric effect of Ln4 clusters (Ln = Eu, Gd, Tb, Er) bridged by CO32− deriving from the spontaneous fixation of carbon dioxide in the atmosphere. *Inorg. Chem. Front.* **5**, 394-402 (2018).
42 Zheng, X. Y. *et al.* A gigantic molecular wheel of {Gd140}: a new member of the molecular wheel family. *J. Am. Chem. Soc.* **139**, 18178-18181 (2017).

43 Zhou, Y. *et al.* Three giant lanthanide clusters Ln37 (Ln = Gd, Tb, and Eu) featuring a double-cage structure. *Inorg. Chem.* **56**, 2037-2041 (2017).
44 Li, X. Y. *et al.* A giant Dy76 Ccuster: a fused Bi-nanopillar structural model for lanthanide clusters. *Angew. Chem. Int. Ed.* **58**, 10184-10188 (2019).
45 Xiao, H. P., Zhou, J., Zhao, R. Q., Zhang, W. B. & Huang, Y. A series of lanthanoid selenidoantimonates(V): rare examples of lanthanoid selenidoantimonates based on dinuclear lanthanide complexes. *Dalton Trans.* **44**, 6032-6039 (2015).
46 Ellart, M., Blanchard, F., Rivenet, M. & Abraham, F. Structural variations of 2D and 3D lanthanide oxalate frameworks hydrothermally dynthesized in the presence of hydrazinium Ions. *Inorg. Chem.* **59**, 491-504 (2020).
47 Ramaswamy, P., Wong, N. E. & Shimizu, G. K. MOFs as proton conductors--challenges and opportunities. *Chem. Soc. Rev.* **43**, 5913-5932 (2014).
48 Lim, D. W. & Kitagawa, H. Proton transport in metal-organic frameworks. *Chem. Rev.* **120**, 8416-8467 (2020).