Inverse design of metal-organic frameworks for
direct air capture of CO$_2$ via deep reinforcement
learning

Hyunsoo Park$^{1,2}$†, Sauradeep Majumdar$^2$†, Xiaoqi Zhang$^2$, Jihan Kim$^1$*, Berend Smit$^2$*

$^1$Department of Chemical and Biomolecular Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, 34141, Republic of Korea.
$^2$Laboratory of Molecular Simulation, École Polytechnique Fédérale de Lausanne (EPFL), Rue de l’Industrie 17, CH-1951 Sion, Switzerland.

*Corresponding author(s). E-mail(s): jihankim@kaist.ac.kr; berend.smit@epfl.ch;
†These authors contributed equally to this work.

Abstract

The combination of several interesting characteristics makes metal-organic frameworks (MOFs) a highly sought-after class of nanomaterials for a broad range of applications like gas storage and separation, catalysis, drug delivery, and so on. However, the ever-expanding and nearly infinite chemical space of MOFs makes it extremely challenging to identify the most optimal materials for a given application. In this work, we present a novel approach using deep reinforcement learning for the inverse design of MOFs, our motivation being designing promising materials for the important environmental application of direct air capture of CO$_2$ (DAC). We demonstrate that the reinforcement learning framework can successfully design MOFs with critical characteristics important for DAC. Our top-performing structures populate two separate subspaces of the MOF chemical space: the subspace with high CO$_2$ heat of adsorption and the subspace with preferential adsorption of CO$_2$ from humid air, with few structures having both characteristics. Our model can thus serve as an essential tool for the rational design and discovery of materials for different target properties and applications.
1 Introduction

Metal-organic frameworks (MOFs) are a class of crystalline nanomaterials known for their high internal surface areas, tunable chemistries, and wide range of pore sizes[1]. The variety of features makes MOFs promising materials for a wide range of applications like carbon capture[2], methane storage[3], hydrogen storage[4], photocatalysis[5], drug delivery[6], conductivity[7], and so on. To date, there are more than 100,000 experimentally reported MOF structures[8–10] and millions of MOF structures that have been predicted in silico[2, 11–18].

From the perspective of material chemists, the holy grail is to design the most promising materials for a given application[19]. Currently, the most commonly used strategy is to find the best materials through brute force screening of these materials library. However, as we are interested in a small subset of top-performing materials, most of the effort in these brute-force methods is spent on computing properties of those materials that are not interesting. As the number of structures in these databases has grown significantly, several groups have started taking a different approach in searching more efficiently within this infinite chemical space of MOFs, using methods such as diversity-driven searches[16, 17, 20], and active learning based searches[19, 21].

An alternative way to explore this chemical design space is through the inverse design of materials with desired properties[22, 23]. The chemical design space of MOFs involves combining metal clusters and organic linkers with topologies, which gives us a highly complex design space. This complexity is partly why only a limited number of studies focused on the inverse design of MOFs. Yao et al.[24] developed a deep-generative model with variational autoencoders for the inverse design of MOFs with
desired properties by optimizing the latent space. However, optimizing the materials with non-convex objective functions in such a high dimensional latent space can be challenging.\[25\] In addition, evolutionary algorithms have commonly been used to optimize the chemical design space to identify best-performing MOFs across various applications such as gas storage and separation applications\[15, 26–29\].

In this context, deep reinforcement learning has emerged as a promising approach for inverse design with desired properties that maximizes the outcome based on learning from existing knowledge. In particular, this approach has been successfully applied in de novo drug design\[30–32\]. The approach allows learning how to design molecules that maximize a reward function using deep learning models such as recurrent neural networks and generative adversarial networks. Apart from molecule generation, Pan et al.\[33\] demonstrated that the reinforcement learning approach could be extended to inorganic materials such as metal oxide. However, the complex crystal structure of MOFs, which typically contain more than 100 atoms per unit cell, presents a significant challenge for this approach. In this work, we propose a new deep reinforcement learning framework designed for large crystalline materials such as MOFs, which can tackle the complexity of crystalline systems through reinforcement learning.

We apply this reinforcement learning framework to design MOFs for direct air capture (DAC). DAC has been developed to reduce the CO\(_2\) concentration in the atmosphere. As we have a large risk of overshooting these CO\(_2\) levels, DAC has become a very active area of research\[34–36\]. The main challenge of DAC is the low concentration of CO\(_2\) in the air (compared to flue gas), and at such an extremely dilute concentration (\(~400\) ppm), one needs a material with a very high CO\(_2\) enthalpy. Currently, a class of materials known as chemisorbents is used industrially to tackle this challenge. Chemisorbents can bind CO\(_2\) very strongly, which is reflected in its high heat of adsorption (\(>65\) kJ/mol)\[36\]. But such high heat of adsorption also leads to increased difficulty and costs to regenerate the materials\[37, 38\]. Physisorbents, on the
other hand, usually have much lower CO$_2$ heat of adsorption (< 40 kJ/mol), though there have been reports of few structures having CO$_2$ heat of adsorption higher than 50 kJ/mol[39]. So a question that arises here is, could physisorbents be used for DAC? To this end, Findley et al.[38] computationally screened the CoRE MOF database and several classes of zeolites to search for the best physisorbent MOF for DAC. Based on their analysis, they concluded that those structures are not viable for DAC. Hence, whether one can design a library of MOFs specifically for DAC is interesting.

The main objectives of this work are thus: (1) to illustrate the use of a reinforcement learning framework to inverse-design MOFs with desired properties; (2) to highlight the use of these inverse-designed MOFs for an extremely important application like DAC. We first show the working of our reinforcement learning workflow by inverse designing MOFs with CO$_2$ heat of adsorption higher than 40 kJ/mol. To make this design even more challenging, we must consider the effect of H$_2$O. Depending on the climate, there are different levels of water vapor in the atmosphere, and most materials that strongly bind CO$_2$ have a higher affinity for H$_2$O. Hence, we must design MOFs that perform well in humid conditions. And for this reason, our second design criterion is that the material prefers CO$_2$ above H$_2$O (i.e., CO$_2$/H$_2$O selectivity to be higher than 1). The successful inverse design of MOFs for a challenging application like DAC could significantly advance the rational design and discovery of materials for a wide range of applications.

2 Results

2.1 Data representation of MOFs for inverse design

The data representation of MOFs for inverse design is crucial for successfully developing our reinforcement learning framework, as illustrated in Fig. 1a. Each MOF structure in our database is represented by a unique combination of metal clusters, organic linkers, and topologies. We used PORMAKE[15], a Python library
that constructs porous materials from given building blocks with various topologies. Within PORMAKE, there are libraries of building blocks (from the CoRE MOF database[9]) and topologies (from the Reticular Chemistry Structure Resource (RCSR) Database[40]). Given metal clusters in MOFs usually have constraints due to complex chemistry (such as oxidation states), it is difficult to represent them as continuous variables. We thus featurized the metal clusters and topologies (represented in RCSR
using three-letter codes) with categorical variables. On the other hand, organic linkers can be represented with continuous vectors such as SMILES, which is a string representation of molecules ensuring the basic chemical rules. Instead of using SMILES, SELFIES\cite{41} was adopted because it has been shown to outperform SMILES within generative models such as VAE\cite{42}, GAN\cite{43} by generating more diverse and valid molecules. Additional details of the MOF structure representation are provided in the Methods section.

2.2 Reinforcement Learning Framework

The overall schematic of the reinforcement learning framework is depicted in Fig. 1b, which comprises two key components: an agent and an environment. The agent, which serves as the generator, takes action by creating a data representation of a new MOF structure. At the same time, the predictor evaluates the action as the environment by predicting the property of the new MOF representation. Based on the prediction, a reward is returned to update the agent for generating MOF structures to maximize the reward. The objective of reinforcement learning is to find the weights of the agent that maximize the expected return obtained from the environment. The details of the framework are provided in the Methods section.

2.2.1 Generator

Acting as an agent, the generator requires a pre-training stage to learn the essential underlying chemistry of MOFs. This knowledge includes which combinations of topology, metal clusters, and organic linkers can be constructed for MOFs and how to generate organic linkers of MOFs. The generator architecture is based on the Transformer\cite{44} architecture, which consists of an encoder and a decoder, as shown in Fig. 2a. It is important to note that the number of connection points of metal clusters and organic linkers must be appropriately matched within a given topology while constructing MOFs. For instance, the \textit{pcu}, a topology of IRMOF-1, requires a vertex
The architecture of the generator consists of a Transformer encoder and decoder. The encoder takes metal clusters, the number of connection points of organic linkers, and a batch of scaffolds of organic linkers represented by SELFIES as inputs. Scaffolds refer to the core structures of the molecular framework. The decoder selects topologies and metal clusters, which are both categorical variables. The organic linkers are generated based on the scaffolds used as inputs. The schematic illustrates the process of generating MOF representations by the generator and the biased generator (with frozen weights) to balance the trade-off between exploitation and exploration in reinforcement learning. The exploitation-to-exploration ratio is determined by the threshold parameter $\lambda$. The architecture of the predictor is based on the Transformer encoder. The predictor takes the MOF representations from the generator as inputs. A simple dense layer is added to the token at the first position (i.e., class token) to predict the target properties of interest.

comprising six connection points, whereas the tbo, a topology of HKUST-1, necessitates two types of vertices comprising 3 and 4 connection points. To address it, the encoder of the generator receives a metal cluster along with the number of connection points of an organic linker as inputs, and the decoder identifies the suitable topologies based on their connection points. To ensure sufficient diversity of organic linkers effectively, a scaffold-based approach was adopted for organic linkers, which was previously demonstrated as a successful methodology for generating diverse molecules with deep
generative models\cite{45–47}. The decoder of the generator selects topologies and metal clusters sequentially and creates the SELFIES of organic linkers by retaining a batch of scaffolds employed as inputs for the encoder. The details of building scaffolds are explained in the Methods section.

2.2.2 Predictor

The predictor provides the estimated target properties, allowing for the calculation of the rewards. As illustrated in Fig. 2c, the architecture based on the Transformer encoder takes different embedding layers of topologies, metal clusters, and vocabulary of SELFIES. By adding a simple dense layer to the class token at the first position, which is a learnable embedding layer, the predictor can predict the desired target properties. The performance of the predictor was evaluated by measuring the mean absolute errors (MAE) of CO$_2$ heat of adsorption and CO$_2$/H$_2$O selectivity, which were found to be 2.87 kJ/mol and 0.64, respectively (see Supplementary Fig. S1). It should be emphasized that the reward functions of the reinforcement learning framework were assigned based on the predicted targets of the predictor.

2.2.3 Exploration strategy

In reinforcement learning, it is essential to consider the trade-off between exploration and exploitation to achieve optimal performance. Exploitation involves selecting actions that maximize immediate rewards and generate MOFs similar to those created by the agent in the pre-training stage. On the other hand, exploration involves selecting actions that explore the action space for long-term benefits, generating novel MOF structures that have not been previously seen. Achieving this balance is critical to the success of the reinforcement learning algorithm. To address this challenge, we introduced two generators: one for exploration, and one biased towards exploitation (see Fig. 2b). The generator was updated to maximize rewards through the reinforcement learning algorithm, while the weights of the biased generator were frozen to enable
continuous generation of MOFs by the pre-training generator. During training, each token of the MOF representations was selected from either the biased generator or the generator, with the choice determined by a threshold $\lambda$. A higher value of $\lambda$ favored exploitation over exploration, and an empirical observation determined $\lambda$ to be set at 0.5. Moreover, to improve the diversity of topology and metal clusters of the generated MOF structures, 50% of the metal clusters in the inputs were masked during the training stage of reinforcement learning, allowing the decoder of the generator to select topologies and metal clusters without considering the metal clusters used as inputs.

2.3 Results of reinforcement learning

We trained the reinforcement learning framework using the pre-training generator and predictors, with the aim of generating novel MOFs that exhibit high CO$_2$ heat of adsorption and CO$_2$/H$_2$O selectivity. The performance of the optimized generator by the reinforcement learning algorithm was evaluated using three metrics: validity, scaffold, and uniqueness. The validity metric evaluates if the generator can generate MOFs which (1) match connection points of metal clusters and organic linkers for a given topology and (2) produce chemically valid organic linkers. To evaluate (2), we converted the generated SELFIES to canonical SMILES using the RDkit sanitizer. The scaffold metric measures whether the generated organic linkers contain a batch of scaffolds used as encoder inputs. Finally, the uniqueness metric assesses the proportion of distinct organic linkers generated. The performance was tested on a test set of 10,000 data used during the pre-training stage, and the performance was evaluated using metrics and rewards, as summarized in Table 1.

As we have a relatively small number of structures with the desired properties, we carried out the reinforcement learning for three rounds. After each round, we used the top-performing MOFs to retrain the predictors. More details on the training
Fig. 3 Comparison of the distributions of (a) CO$_2$ heat of adsorption and (b) CO$_2$/H$_2$O selectivity for scratch and optimized MOFs based on the predictor’s estimated values. The scratch distributions show the property distributions of the MOFs generated by the pre-training generator before being optimized by the reinforcement learning algorithm. The optimized distributions show the property distributions of the MOFs generated by the generator after being optimized by the reinforcement learning algorithm. After each round, the final top-performing MOFs of that round were selected and added to the next round for retraining the predictor. And, this updated predictor was then used in the training of the reinforcement learning for the next round. The evolution of the generated structures for (c) CO$_2$ heat of adsorption and (d) CO$_2$/H$_2$O selectivity in the chemical space of MOFs. (e) The mapping of the training and generated structures for the two targets. The grey, orange, and green points represent the pre-training structures, the top-performing generated structures for the heat of adsorption, and those for selectivity, respectively.

of the reinforcement learning framework are provided in the Methods section. Supplementary Fig. S1 shows the parity plots of the predictors for each round. In Fig. 3, the property distributions of MOFs generated by the pre-training generator (i.e., scratch) and generators optimized by reinforcement learning for three rounds are illustrated, where the estimated property values are based on the predictors. Notably, it
is observed that the average target values increase as the rounds progress, indicating the success of the reinforcement learning approach in optimizing the generator to produce MOFs with improved properties. Supplementary Fig. S2 reveals that the property distribution improves significantly as training progresses within a particular round. Table 1 shows that the overall scaffold metric decreased after optimization compared to scratch because the optimized generator generated organic linkers without the scaffolds used as input to maximize rewards. After optimization, the metric for uniqueness also exhibits a decrease, suggesting that the generator favors generating organic linkers with desirable properties. Based on the predictors’ estimations, the most frequently observed topologies, metal clusters, and organic linkers for rounds 1, 2, and 3 are presented in Supplementary Fig. 4, 5, and 6, respectively.

<table>
<thead>
<tr>
<th>metric</th>
<th>scratch</th>
<th>CO\textsubscript{2} heat of adsorption</th>
<th>CO\textsubscript{2}/H\textsubscript{2}O selectivity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>validit\texttt{y}</td>
<td>round 1</td>
<td>round 2</td>
</tr>
<tr>
<td>validity</td>
<td>0.69</td>
<td>0.69</td>
<td>0.70</td>
</tr>
<tr>
<td>scaffold</td>
<td>0.96</td>
<td>0.75</td>
<td>0.79</td>
</tr>
<tr>
<td>uniqueness</td>
<td>0.79</td>
<td>0.61</td>
<td>0.64</td>
</tr>
<tr>
<td>target</td>
<td>19.16</td>
<td>26.44</td>
<td>28.52</td>
</tr>
<tr>
<td>reward</td>
<td>0.32</td>
<td>0.44</td>
<td>0.48</td>
</tr>
</tbody>
</table>

Table 1 Performance metrics of the pre-training generator and the reinforcement learning for CO\textsubscript{2} heat of adsorption and CO\textsubscript{2}/H\textsubscript{2}O selectivity

2.4 Chemical space analysis of the MOFs

We performed a chemical space analysis of the training and generated structures for the two targets, with t-distributed stochastic neighbor embedding (t-SNE)[48]. The latent vectors from the transformers of the respective predictors of the two targets of the third round were projected into two dimensions. The grey, orange, and green points represent the structures used to pre-train the predictors, the generated structures for CO\textsubscript{2} heat of adsorption, and those for CO\textsubscript{2}/H\textsubscript{2}O selectivity, respectively. From Fig. 3c, we can visualize the evolution of the generated structures for CO\textsubscript{2} heat of adsorption (for structures having values greater than 30 kJ/mol); the number of structures with
high CO$_2$ heat of adsorption increases over the different rounds (represented by dark orange points). A similar plot for CO$_2$/H$_2$O selectivity (for structures having values greater than 1) is shown in Fig. 3d. These trends are consistent with the distribution shift shown in Fig. 3a and b. Finally, fig. 3e highlights the two separate subspaces occupied by these generated structures for the two targets. This corroborates the fact that structures having high CO$_2$ heat of adsorption have characteristics quite different from structures having high CO$_2$/H$_2$O selectivity. Thus they occupy relatively different subspaces in the MOF chemical space, with a small overlap. This fact is explained in more detail in the Discussion section. It is important to note that we managed to generate a few structures that simultaneously satisfy both targets (represented by the overlapping orange and green points).

2.5 Feasibility tests for the generated top-performing MOFs

To ensure that the structures generated from our reinforcement learning framework are reasonable, we employed different structure feasibility tests to narrow down the ideal MOF candidates for DAC (the workflows for the same are summarized in Supplementary Fig. S2). First, from the 10,000 test set, the valid MOFs, satisfying the validity metric, generated by the optimized generators with CO$_2$ heat of adsorption higher than 30 kJ/mol and CO$_2$/H$_2$O selectivity higher than 1 were selected. Then, to estimate the synthesizability of the organic linkers, we computed the synthetic accessibility (SA) score [49]. The SA score is based on molecular complexity; molecules with a low SA score are less complex and are to be expected to have an easier synthesis route compared to those more complex molecules with a high SA score. The MOFs with an SA score of organic linkers higher than 6 were dismissed. In addition, the number of generated structures was restricted through the topological root mean squared deviation (RMSD) of the atomic positions between the building block node vectors and the target topology node vectors[15]. Lower topological RMSD values indicate that
the strain between the two vectors is low, and the resultant MOF structure is more stable. Given that the MOFs created by the PORMAKE are typically feasible when the topological RMSD is lower than 0.3\cite{15}, the same constraint was adopted. Apart from these, the generated MOFs with more than 3000 atoms and higher than 60 Å cell lengths were also dismissed. Structure optimization and charge generation were then carried out to obtain further reasonable structures. The generated structures that passed all the above tests went through molecular simulations to estimate the respective “true” target values. The details of the molecular simulations are provided in the Methods section. Finally, after all the tests, we obtained 409/497/999 structures with CO$_2$ heat of adsorption higher than 40 kJ/mol and 2215/2304/2426 structures with CO$_2$/H$_2$O selectivity higher than 1 in round 1/round 2/round 3, respectively.

3 Discussion

From a material chemist’s perspective, it is interesting to know what makes a material good for DAC? Or in other words, what are the features or the genes of the top-performing materials for DAC? Our top-performing candidates’ topologies, metal clusters, and organic linkers are shown in Fig. 4. When it comes to CO$_2$ heat of adsorption, we find that the top-performing MOFs mostly contain the metal clusters Mn-based N131 and Eu-based N520 (see Fig. 4; the naming of all the metal clusters can be found in the PORMAKE paper\cite{15}.). The Mn-based N131 cluster has an open metal site that strongly attracts the CO$_2$ molecules towards the metal atom. Many structures built from this cluster have a CO$_2$ heat of adsorption value higher than 40 kJ/mol. One example of such a structure is shown in Fig. 5a. Other metal clusters within the high-performing MOFs include lanthanide metals (i.e. Nd, Sm, Dy) and transition metals (i.e. Ni, Co, Zr).

Apart from the metal cluster, the organic linkers of the top-performing MOFs have more abundant branches (i.e., functional groups) such as F, Cl, Br, NH$_2$ exhibiting
Fig. 4  Frequently occurring topologies, metal clusters, and organic linkers of the top-performing MOFs based on molecular simulations for (a) CO$_2$ heat of adsorption (b) CO$_2$/H$_2$O selectivity in round 3

high electronegativity, which leads to strong binding with CO$_2$. The functionalization of the MOF with fluorine is also a characteristic of one of the few physisorbent MOFs synthesized for DAC by Bhatt et al [39].

The top-performing MOFs with respect to the high selectivity of CO$_2$/H$_2$O, however, have different characteristics from the ones having high CO$_2$ heat of adsorption, as shown in Fig. 4b. The metal clusters include metals frequently used for synthesizing MOFs such as Cu, Zn, Cd. In particular, metal clusters Cu-based N232, and Zn-based N328 appear most frequently. Most metal nodes having open metal sites (like metal node N131) attract H$_2$O more than CO$_2$, and therefore the MOFs consisting of those metal nodes have low CO$_2$/H$_2$O selectivity. Also, given that functional groups with high electronegativity exhibit high H$_2$O affinity and high affinity for CO$_2$, the generated organic linkers of the high-performing MOFs, in this case, rarely include these
branches. This shows that it is indeed a challenge to find a MOF having a good balance between CO$_2$ heat of adsorption and CO$_2$/H$_2$O selectivity[38].

![Fig. 5 Structures of some of the top-performing MOFs designed for DAC. The corresponding metal node, organic linker, topology and some properties are shown/mentioned alongside. The MOF names, as used in this study, are: (a) v0+bcu+N131+225 (b) v1+hex+N262+1184 (c) v1+tfzd+692+2546 (Note: The MOF naming convention followed in this work is in the form: Version+Topology+MetalNode+OrganicLinker. Version 0 refers to structures designed for high CO$_2$ heat of adsorption, and version 1 refers to structures designed for high CO$_2$/N$_2$ selectivity.) C=light brown, H=white, O=red, N=light blue, Cl=green, Mn=dark violet, Cd=pink, Cu=dark blue, and the black atoms in the metal clusters refer to the points where the metal clusters and organic linkers get connected to each other.]

Fig. 5 shows the inverse-designed structures of some of the most promising candidates for DAC application. Structure (a) has a high CO$_2$ heat of adsorption (~62 kJ/mol) but a low CO$_2$/H$_2$O selectivity (<1) whereas structure (b) has a high CO$_2$/H$_2$O selectivity (>1) but a low CO$_2$ heat of adsorption (~29 kJ/mol). And structure (c) has both high CO$_2$/H$_2$O selectivity (>1) and moderately high CO$_2$ heat.
of adsorption (∼40 kJ/mol). It is important to note that these three structures are promising candidates for DAC, depending on the environment in which they are used. For example, if a particular industrial process configuration can have a dehumidifier unit before the adsorption step, then structure (a) would be potentially interesting despite its low CO$_2$/H$_2$O selectivity. This structure would also be promising in regions with very low water vapor content in the atmosphere. The stronger binding of CO$_2$ to the MOF would become the more dominating factor in choosing this material. On the other hand, if there is no dehumidifier unit, then structures (b) and (c) would be the better choice because of their high CO$_2$/H$_2$O selectivity. Structure (c), satisfying both requirements, is likely to be an interesting candidate in both scenarios: with and without humidity. Thus, it would be interesting to evaluate these materials one step further, on a process level, and we are looking into it as future work.

4 Conclusion

In this work, we have developed a deep reinforcement learning framework to inverse design MOFs. We illustrate this approach to design MOFs with important characteristics for direct air capture (DAC) of CO$_2$. We successfully (inverse) designed a set of materials that have a high CO$_2$ heat of adsorption (>40 kJ/mol). In addition, we generated a set of materials that preferentially adsorb CO$_2$ from humid air. Subsequent analysis of the chemical design space shows that the top-performing structures populated two, separate subspaces concerning the two target properties. Yet, few of our structures satisfy both requirements.

The heat of adsorption is a simple proxy for performance in a DAC process; it allows us to eliminate really poor structures and identify some of the promising ones. The next step would be to apply reinforcement learning using a reward function given by a detailed process design.
5 Methods

5.1 Details of MOF structure representation

The building blocks from PORMAKE are composed of vertices (i.e., abstractions with more than two connection points) and edges (i.e., abstractions with two connection points)[15]. In this work, however, they were modified more intuitively into metal clusters and organic linkers depending on the presence of metal atoms in the building blocks. Thereby, 486 vertices containing metal atoms were used for the metal clusters. For organic linkers, there are 103 vertices and 175 edges that do not contain metal atoms. Apart from building blocks, we used 97 topologies extracted from the CoRE-MOF database by MOFkey[50], which is summarized in Supplementary Note S1. It is important to note that data augmentation of the organic linkers is required to train with deep generative models, given the insufficient number of organic linkers. Hence, augmentation was implemented by merging the organic linkers, where the connection points of organic linkers were replaced with the other organic linkers. To this end, 30,642 merged organic linkers were generated, and their validity was examined by the RDkit sanitizer[51]. It is to be noted that these data representations can be used to reconstruct the MOF structures using PORMAKE.

5.2 Computational details for molecular simulations

All structures for pre-training were optimized using UFF with Lammps[52] that can be implemented with multi-core processors, which facilitates calculations with a large number of structures. The number of MOF structures generated via reinforcement learning was much less than the number used for pre-training, and these were optimized using the Universal forcefield (UFF)[53] as implemented in the Forcite Module of Materials Studio 2019[54]. The EQEq (extended charge equilibration) method [55, 56] was used to generate the partial charges of the framework atoms of the MOFs. The lowest common oxidation states of the elements were chosen as their charge
The Henry coefficient and heat of adsorption at infinite dilution, for CO$_2$ and H$_2$O, were computed using Widom’s test particle insertion algorithm\[57\]. All of the calculations were performed at 298.15 K using RASPA \[58\]. We computed the CO$_2$/H$_2$O Henry’s selectivity as the ratio of the respective Henry coefficients:

$$S_{\text{CO}_2/\text{H}_2\text{O}} = \frac{K_{\text{HCO}_2}}{K_{\text{HH}_2\text{O}}} \quad (1)$$

Blocking spheres were calculated using Zeo++\[59\] to block the inaccessible regions of the MOF structures\[60, 61\]. The framework atoms of the MOF structures were described by UFF\[53\]. CO$_2$ and H$_2$O molecules were described by the TraPPE forcefield\[62\] and TIP4P/2005 forcefield respectively\[63\]. The gas-framework interactions were modeled using Lennard Jones potential, truncated at 12 Å, with tail corrections \[64\]. The Lennard Jones interactions between dissimilar atoms were approximated using Lorentz-Berthelot rules \[65\]. The Coulombic electrostatic interactions were computed using Ewald summation.

5.3 Details of Reinforcement Learning Framework

A generator plays the role of the agent and generates the data representations of new MOFs that are defined as a set of state $S$ with length $T$ of MOF representation. The representation consists of a topology, a metal cluster as categorical variables, and an organic linker represented by the SELFIES string. Thus the state $S$ can be expressed as $S(\text{topology, metalcluster, organiclinker})$. Given the maximum length of a SELFIES string, including padding, is 100, the maximum length of $T$ is 102. At each step $t$, \((0 < t < T)\), the generator takes state $s_{1:t}$ in previous steps as inputs and then determines probability distribution $p$ of state $s_t$. At $t = 0$ and $t = 1$, the generator respectively determines topologies and metal clusters that are categorical variables. Then, it generates sequences of SELFIES for organic linker during $1 < t \leq T$.  

18
The predictor takes the data representations created by the generator as inputs and constructs the expected reward $r(s_t)$. A policy gradient method is implemented for the generator to maximize the expected return based on the expected reward of the predictor. The expected return is defined as:

$$J(\theta) = \mathbb{E}[r(s_{1:T})|\theta] = \sum_{t=1}^{T} \log p(s_t|s_{1:t-1})r(s_t)$$ (2)

The objective of reinforcement learning is to find the weights of the agent ($\theta$) that maximize the expected return obtained from the environment ($J(\theta)$).

5.4 Training details for pre-training

In the pre-training step, 646,907 MOFs generated by PORMAKE were used as the dataset for the generator. Each organic linker was decomposed into scaffold fragments with the BRICS algorithm, which allow splitting molecules into chemically semantic fragments. For each organic linker, if the decomposed fragments were a subset of other fragments, the subsets were dismissed. And the maximum number of scaffold fragments was set to four. The smaller fragments were omitted if a molecule was decomposed into more than four scaffold fragments. The scaffold fragments were randomly combined and joined in SELFIES to build a batch of scaffolds. Finally, a dataset of 1,540,889/385,223/10,000 (train/validation/test) was created for the generator. For the predictor, Widom calculations were implemented to compute the CO$_2$ heat of adsorption and CO$_2$/H$_2$O selectivity data for ~30,000 MOFs generated by PORMAKE. They were split in the ratio 8:1:1 (train:validation:test).

The architectures of the generator and predictor were derived from the original Transformer paper[44]. For the generator, the Transformer encoder and decoder consist of 3 layers, 4 heads, and a hidden size of 256. The maximum length of the data representations of MOFs is 128. The model was trained with a batch size of 128 during
50 epochs and AdamW[66] optimizer with a learning rate of $10^{-4}$. The predictor, which is similar to the encoder of the generator, consists of 4 layers of the Transformer encoder. It was trained with a batch size of 128 during 100 epochs. The AdamW optimizer with a learning rate of $10^{-4}$ and weight decay of $10^{-2}$ was used for the predictor. The learning rate was warmed up during the first 5% of the total epoch and then linearly decayed to zero for the remaining epochs.

5.5 Training details for reinforcement learning

The reward functions used in reinforcement learning were assigned by the estimated values provided by the predictor. The reward functions of CO$_2$ heat of adsorption and CO$_2$/H$_2$O selectivity are defined as Eqn. 3 and 4, respectively.

\[
\begin{align*}
\tau & = \frac{Q_{st}}{60} \\
\tau & = \frac{S_{CO2/H2O}}{10}
\end{align*}
\]

The policy gradient algorithm was trained with a batch size of 16 during 20 epochs where each epoch was constructed by randomly selecting 8,000 data in the training dataset of the generator. The optimizer and scheduler were the same as those used in the training process of the generator.

The reinforcement learning was carried out for three rounds. After each round, the top-performing structures of that round were added to retrain the predictor for the next round. And, this updated predictor was then used in the training of the reinforcement learning for the next round.

Acknowledgments. This research was supported by the USorb-DAC Project, which is funded by a grant from The Grantham Foundation for the Protection of the Environment to RMI’s climate tech accelerator program, Third Derivative. H. P. and J. K. acknowledge funding from National Research Foundation of Korea (NRF) under...
Project Number 2021M3A7C208974513. The authors would like to thank Kevin Maik Jablonka for insightful discussions.

Conflicts of interest

There are no conflicts to declare.

Author contributions

H.P. and S.M. contributed equally to this work. H.P. developed the reinforcement learning framework and analyzed the results of reinforcement learning. S.M. conceived the idea of the project, performed molecular simulations, and analyzed the top-performing MOFs. X.Z. performed the chemical space analysis of the MOF structures. J.K and B.S. led the project and provided directions. The manuscript was written through the contributions of all authors. All authors have approved the final version of the manuscript.

Data availability

The cif files and properties of the top-performing MOFs for CO$_2$ heat of adsorption and CO$_2$/$\text{H}_2\text{O}$ selectivity for each of the three rounds, the dataset used for pre-training the predictor, and some additional data are available on the Zenodo platform at https://doi.org/10.5281/zenodo.7853157[67].

Code availability

The reinforcement learning framework is available at https://github.com/hspark1212/ MOFreinforce as open source so that it allows training the reinforcement learning with other target properties. Our GitHub source code will be updated. A consolidated
version of our code used for this work has been exported to Zenodo at https://doi.org/10.5281/zenodo.7851250[68].

References


[29] Han, S., Kim, J.: Design and screening of metal–organic frameworks for ethane/ethylene separation. ACS Omega (2023)

adversarial network for inverse-design chemistry (organic) (2017)


[39] Bhatt, P.M., Belmabkhout, Y., Cadiau, A., Adil, K., Shekah, O., Shkurenko, A.,


H., Engkvist, O.: Smiles-based deep generative scaffold decorator for de-novo drug

(2008)

[49] Ertl, P., Schuffenhauer, A.: Estimation of synthetic accessibility score of drug-
like molecules based on molecular complexity and fragment contributions. J.
Cheminformatics 1(1), 1–11 (2009)

for metal–organic frameworks to enable rapid search and cheminformatics analy-


[52] Thompson, A.P., Aktulga, H.M., Berger, R., Bolintineanu, D.S., Brown, W.M.,
(2022)

a full periodic table force field for molecular mechanics and molecular dynamics


