Inverse Design of Metal-Organic Frameworks for C_2H_4/C_2H_6 Separation

Musen Zhou and Jianzhong Wu*

Department of Chemical and Environmental Engineering, University of California, Riverside, CA, 92521, United States

Abstract

Efficient separation of C_2H_4/C_2H_6 mixtures is of paramount importance in the petrochemical industry. Nanoporous materials, especially metal-organic frameworks (MOFs), may serve the purpose owing to their tailorable structure and pore geometry. In this work, we propose a computational framework for high-throughput screening and inverse design of high-performance MOFs for adsorption and membrane processes. High-throughput screening of the computational-ready, experimental (CoRE 2019) MOF database leads to materials with exceptionally high ethane-selective adsorption selectivity (LUDLAZ: 7.68) and ethene-selective membrane selectivity (EBINUA02: 2167.3). Moreover, the inverse design enables the exploration of a broader chemical space and identification of MOF structures with even higher membrane selectivity and permeability. In addition, a relative membrane performance score (rMPS) has been formulated to evaluate the overall membrane performance relative to the Robeson boundary. The computational framework offers guidelines for the design of MOFs and is generically applicable to materials discovery for gas storage and separation.

* To whom correspondence should be addressed. Email: jwu@engr.ucr.edu

1. Introduction

The efficiency of C_2H_4/C_2H_6 separation is important for the petrochemical industry because high-purity C_2H_4 is used as the primary feedstock for the synthesis of diverse chemical products including plastics, polyesters and rubber materials.[1,2] Conventional processes for C_2H_4/C_2H_6 separation are mostly based on high-pressure cryogenic distillation, which requires extensive energy input while suffers from low separation efficiency. To reduce the energy cost and increase the selectivity, it is desirable to develop alternative approaches such as adsorption or permeation processes based on nanoporous materials.[3-5]

Metal-organic frameworks (MOFs) are ideal candidates for efficient separation of C_2H_4/C_2H_6 because they have good mechanic stability, large specific surface area, and tailorable pore structure and geometry.[6-8] In particular, such materials show promising performance for separating molecules with similar size and interaction energy, such as the mixtures of H_2/D_2 isotopes, of noble gases (Ar/Kr/Xe), and of xylene isomers.[9-12] For C_2H_4/C_2H_6 separation, promising MOF candidates have been identified by experiments.[3,4,13,14] Whereas the possible variations of MOFs are virtually unlimited and the separation efficiency is sensitive to the atomic details, it is practically infeasible to explore the design space only through experimentation. Previously, computational methods have been used to identify best material candidates for separation process through high-throughput screening.[9,15,16] While the adsorption isotherms predicted by the computational methods are found in good agreement with experimental measurement,[17] membrane processes are often considered more efficient to separate C_2H_4 from C_2H_6 leveraging the difference in both adsorption affinity and gas diffusivity.[18] To the best of our knowledge, previous research on the computational screening of MOF database is mostly concerned with the separation of C_2H_4/C_2H_6 by adsorption.[17,19-22] From the computational perspective, the assessment of MOF

materials for membrane separation is much more demanding because the evaluation of gas diffusivity in confined geometry is typically more time-consuming than that for gas adsorption. In particular, the strong confinement makes it computationally prohibitive to predict the diffusion coefficients of gas molecules in a large library of nanoporous materials using conventional methods such as molecular simulation (MD) simulation.[23,24]

Although computational methods (*e.g.*, MD, grand canonical Monte Carlo simulation, and classical density function theory) have been well established for accurate prediction of gas adsorption and diffusivity,[25-28] the inverse design of nanoporous materials for separation processes remains a theoretical challenge from both computational and practical perspectives. While generative adversarial network (GAN) shows early success in the inverse design of zeolites for methane storage,[29] its computational complexity increases significantly with the number of elements in the crystal structure. Besides, GAN easily breaks down and fails to converge for complicated crystalline materials such as MOFs because a large number of atomic types need to be considered. In contrast, variational autoencoder (VAE) can well accommodate the complex topology and molecular structure of the secondary building blocks (SBUs) by compressing the MOF structure into a text string and projecting it into the latent space.[30] However, VAE requires accurate projection (*viz*., encode and decoder) between the crystal structure and a latent space, and the VAE training would become infeasible when a vast number of SBUs are considered for the MOF design. Alternatively, evolutionary algorithms, such as the genetic algorithm, are promising for the inverse design of MOFs because they can accommodate not only a large number of SBUs for MOF design, but also find the solution in a nonlinear space consisted of the material topology and SBUs.[31,32]

In previous work,[25,26] we developed a computational procedure with GPU-acceleration that offers fast and accurate evaluation of sorption and diffusion properties of gas molecules in nanoporous materials. Empowered by the developed computational capability, here we perform high-throughput screening of the computational-ready experimental (CoRE 2019) MOF database (over 10k MOFs) for the separation of C_2H_4/C_2H_6 with adsorption and membrane processes. Compared with the state-of-art materials from the literature,[4,33] the best MOFs identified in this work have significantly higher separation selectivity. The highest ethane-selective adsorption selectivity in LUDLAZ is up to 7.68, and the highest ethene-selective membrane selectivity in EBINUA02 can reach 2167.3. Leveraging on the high-throughput capability, a genetic algorithm (GA) is incorporated into our computational workflow to achieve the inverse design of MOF membranes with both high membrane selectivity and permeability. The inverse design allows us to explore a broader chemical space in comparison with high-throughput screening and identify MOFs with even higher membrane selectivity and permeability. The structural analyses of MOFs with best separation performance offer useful guidelines for the experimental design of MOFs for adsorption and membrane separation.

2. Results and Discussion

2.1 Screening CoRE MOF 2019 Database

We first perform the high-throughput screening of the computational-ready experimental (CoRE) MOF 2019 database for separation of C_2H_4/C_2H_6 through adsorption and membrane processes. While high-throughput screening has been commonly used to find the best material candidates for gas adsorption, the procedure is more challenging for membrane separation because of the steep computational cost in evaluating the diffusion coefficients. Figure 1 shows the separation selectivity versus capacity for both adsorption and membrane separations. In Figure 1(a)

and Supplementary Figure 1, MOFs with Henry's constants smaller than 1×10^{-5} mol·m^{-3}·Pa^{-1} and the largest cavity diameters smaller than the hard-sphere diameter of ethene have been removed. Those structures can hardly accommodate ethane/ethene molecules thus are not further considered in this work. The hard-sphere diameters of ethane and ethene are calculated from the Barker-Henderson theory[34] based on their Lennard-Jones (LJ) parameters and more details are provided in Supporting Information. As shown in Figure 1(a), the maximum selectivity is less than 30 for ethene-selective MOFs suitable for adsorption separation. It decreases exponentially with the increase of the separation capacity (*viz*. adsorption amount) because highly confined pores are needed in order to achieve high ethene-selectivity. The structural properties of top ethene-selective and ethane-selective MOFs are shown in Supplementary Table 2 and Supplementary Table 3, respectively. Such materials offer little pore volume to achieve high adsorption capacity. Table 1 lists the properties of top 5 ethene-selective MOFs for the adsorption separation of C_2H_4/C_2H_6 at room temperature (300 K).

Figure 1. Selectivity vs. capacity for CoRE MOFs used in C_2H_4/C_2H_6 separation. (a) Adsorption selectivity. The green line is fitted with $\alpha^{IM}=0.2516e^{-0.7411\log(K)}$. (b) Membrane selectivity. The red line denotes the Robeson boundary, and the color stands for the percentile of relative membrane performance score (rMPS): the red, gray and blue represent the highest, intermediate, and the lowest rMPS, respectively. The dashed line marks the membrane selectivity of 1.

Table 1. Henry's constants (K_h), ideal selectivity (α^{IM}) and self-diffusivity (D_0) of top ethene-selective MOFs for adsorption separation of C_2H_4/C_2H_6 at 300 K. The diffusion coefficients are shown only if they are larger than 1×10^{-20} m^2·s^{-1}.

Although ethene-selective materials yield high selectivity in the separation of C_2H_4/C_2H_6 by gas adsorption, industrial applications desire ethane-selective processes because they can significantly reduce the energy cost. Supplementary Figure 1 shows that, consistent with the literature,[4,5,20] the highest selectivity of ethane-selective MOFs identified in this work is much

smaller than that of ethene-selective MOFs because the stronger adsorption of smaller molecules (*e.g.*, ethene) yields a larger adsorption selectivity of C_2H_4/C_2H_6 via the ultra-small pores of promising MOFs. It is worth noting that, different from ethene-selective MOF for the separation of C_2H_4/C_2H_6 via adsorption, the selectivity of ethane-selective MOFs increases with the capacity and none of the CoRE MOFs have the ethane-selective adsorption selectivity above 8.

Table 2. Top ethane-selective MOFs for the separation of C_2H_4/C_2H_6 at 300 K by gas adsorption.

Table 2 lists the top 5 ethane-selective MOFs with the highest adsorption selectivity of C_2H_6/C_2H_4 at 300 K. Although MOF candidates with high adsorption selectivity of C_2H_4/C_2H_6 have been reported before,[17,19-22] the materials identified in this work yield much higher selectivity for both ethene-selective and ethane-selective separations. Previously, the computational screening was carried either on a smaller structural database or a subset of large structural library (*e.g.*, CoRE MOF 2019) that was restricted by certain structural and chemical criteria. While the application of these criteria would speed up the computation by reducing the number of materials to be evaluated, they ignore promising candidates due to the complex topology and structure. For the adsorption separation of C_2H_4/C_2H_6, the selectivity declines with the increase of loading amount. The reduction in adsorption selectivity can be attributed to the smaller difference between the adsorbate-adsorbate interactions in comparison to that between adsorbate and adsorbent interactions.

According to the structural analysis of the promising materials (results shown in Supplementary Figure 2 and Supplementary Figure 3), the ethane-selective MOFs have less confined geometry in terms of the pore limit diameter (PLD), the largest cavity diameter (LCD) and pore size distribution in comparison with the ethene-selective MOFs. The increase in pore size leads to a much higher adsorption capacity for those MOFs with higher ethane-selectivity. It is

140 worth mentioning that the highest adsorption selectivity of ethane-selective MOF [LUDLAZ:
141 $\alpha^{IM}(C_2H_6/C_2H_4)$=7.68] identified in this work is about 60% higher than the best nanoporous
142 material found by the previous computational screening[22], a hypothetical zeolite structure
143 [$\alpha^{IM}(C_2H_6/C_2H_4)$=4.86]. The adsorption selectivity is more than 70% higher than that of the state-
144 of-art nanoporous material reported in the experimental literature[4], $Fe_2(O_2)$(dobdc) with
145 C_2H_6/C_2H_4 adsorption selectivity of 4.4. Besides, LUDLAZ has a much larger Henry's constant
146 [$K_h(C_2H_6)$=4.6056 $cm^3 \cdot g^{-1} \cdot Pa^{-1}$] than $Fe_2(O_2)$(dobdc) [$K_h(C_2H_6)$=0.0147 $cm^3 \cdot g^{-1} \cdot Pa^{-1}$], meaning
147 much higher gravimetric adsorption capacity. We note in passing that LUDLAZ was originally
148 synthesized by McKellar and coworkers to examine how ligand exchange affects the stability and
149 compressibility of MOF materials.[35] For the top 5 ethane-selective MOFs, their pore size
150 distributions characterized by N_2 adsorption are shown in Supplementary Figure 3(b). These
151 materials have similar micropores between 5 Å and 7 Å, rendering higher selectivity of ethane
152 over ethene in adsorption separation.

153 Compared with adsorption, gas separation *via* permeation through MOF membranes may
154 achieve not only higher selectivity but also larger separation capacity. In addition, a membrane
155 splits the feed stream into two purified sub-streams (*viz*., retentate stream and permeate stream)
156 such that it does not require a recovery process even for ethene-selective operations. Therefore,
157 the membrane process is often much less energy-intensive in comparison with adsorption. Figure
158 1(b) shows the membrane selectivity versus membrane permeability in the units of barrer. The red
159 line in Figure 1(b) denotes the Robeson boundary, a semi-empirical upper limit summarized by
160 Rungta et al.[2] based on the state-of-art polymer membranes for specifically separating ethane and
161 ethene. We see that many MOFs in CoRE MOF 2019 database surpass the Robeson boundary,
162 indicating their superior performance compared with the polymer membranes. The highest

membrane selectivity of C_2H_4/ C_2H_6 is 2167.3 in MOF – EBINUA02, which was synthesized by Tian et al.[36] with 1D rhombic channel in the 3D diamond topology network. The selectivity is several orders of magnitude larger than the state-of-art membrane materials discovered by previous computational screening and experimental synthesis.[19,33] Table 3 lists the top 5 MOFs with the highest membrane selectivity of C_2H_4/C_2H_6.

Table 3. Top MOFs for C_2H_4/C_2H_6 separation with the highest membrane selectivity (k^{IM}) at 300 K.

In our previous work,[25,26] we proposed the membrane performance score (MPS) to evaluate the overall performance of nanoporous materials by combining the membrane selectivity and permeability. Although MPS offers a direct comparison of nanoporous materials with different permeability and selectivity, it does not evaluate the membrane performance relative to the upper limit of the state-of-art polymer membranes (*viz*., the Robeson boundary). Here, we propose a modification of MPS, which is originally defined as $\mathrm{MPS} = k_{\mathrm{fast/slow}} \times P_{\mathrm{fast}}$. The relative membrane performance score (rMPS) is defined as

$$\mathrm{rMPS} = \left(k_{\mathrm{fast/slow}} - k_{\mathrm{fast/slow}}^{\mathrm{Robeson}}\right) \times P_{fast} \quad (1)$$

where k represents the membrane selectivity, P stands for the gas permeability, and the fast component refers to the one with higher permeability in the binary mixture. Since rMPS evaluates the overall performance of nanoporous materials relative to the Robeson boundary, MOFs with a selectivity below the Robeson boundary would have a smaller and negative rMPS compared with MPS if the membrane selectivity is larger than 1. When the membrane selectivity is smaller than 1, MOFs with membrane selectivity below the Robeson boundary would have a higher rMPS because the selectivity of fast component over slow component is larger than that on the Robeson boundary with the same value of permeability. Because in rMPS the membrane selectivity is

187 defined in terms of the fast component over the slow component and the identity of the fast
188 component might change for different materials, rMPS reflects the relative separation efficiency
189 of the fast component in a MOF membrane compared to that in the state-of-art polymer membrane
190 at the same permeability. For the separation of C_2H_4/C_2H_6, Figure 1(b) shows that a high rMPS
191 value favors MOFs with high permeability but intermediate selectivity (bottom right) instead of
192 intermediate permeability and high selectivity (top center). Because ethane and ethene have similar
193 molecular size and interaction energy, the increase of diffusion selectivity from an intermediate
194 value requires a larger energy barrier along the minimum energy path (MEP), which leads to a
195 significant reduction of the diffusion coefficient and rMPS. Conversely, the increase of diffusion
196 coefficient results in the reduction of diffusion selectivity due to the absence of a large energy
197 barrier for molecular sieving. As shown in Supplementary Figure 1, only the ideal adsorption
198 selectivity of C_2H_6/C_2H_4 increases with the capacity. Therefore, we conclude that MOFs with high
199 rMPS are mostly ethane-selective, and that the membrane selectivity is mostly attributed to their
200 difference in the adsorption amount. Table 4 lists the top 5 MOFs with the highest rMPS.

201 **Table 4.** Top MOFs with the highest relative membrane performance score (rMPS) for the
202 separation of C_2H_4/C_2H_6 at 300 K.
203

204 ***2.2 Structural Features of Promising MOF Membranes***

205 As discussed above, the selectivity of MOF membranes is less compromised (and much higher)
206 at high separation capacity in comparison with MOF adsorbents. To explore the synergetic effects
207 between adsorption and diffusion, we have further examined the structural features of top MOFs
208 with the highest membrane selectivity and rMPS.
209

Figure 2. Distributions of the pore limit diameter (a) and the largest cavity diameter (b) for all CoRE MOFs and MOFs with top 5% ideal membrane selectivity and relative performance score (rMPS) for C_2H_4/C_2H_6 separation.

Figure 2 shows the distributions of pore limit diameter (PLD) and the largest cavity diameter (LCD) for all MOFs in CoRE MOF 2019 database and MOFs with top 5% membrane selectivity and rMPS. Compared with the distributions of PLD and LCD for all CoRE MOFs, it is clear that MOFs with top 5% membrane selectivity (and rMPS) have significantly different structural features. The PLD and LCD distributions suggest that MOFs with top 5% membrane selectivity have much smaller pores than those with top 5% rMPS. However, their void fractions are rather similar, both in the range from 0.4 to 0.7 (shown in Supplementary Figure 4). For MOFs with top 5% membrane selectivity, the PLD mostly distributes between 2.75 Å and 3.5 Å, where the narrow end is even slightly smaller than the LJ diameter of the methylene group in ethene. It is worth mentioning that nanoporous materials do not prohibit gas adsorption even when the PLD is slightly smaller than the LJ diameter of the gas molecules.[37] As a matter of fact, the selectivity is maximized for MOFs with the PLD slightly smaller than the LJ diameter because the extremely narrow pore aperture magnifies the difference in the potential energy between C_2H_4 and C_2H_6 at the transition state. By contrast, MOFs with top 5% rMPS have a PLD distribution spanning from 3.75 Å to 4.75 Å. Compared with MOFs with top 5% membrane selectively, the slightly larger PLD for MOFs with top 5% rMPS leads to a smaller difference in the potential energy between ethane and ethene at the transition state, and therefore, much higher permeability with intermediate membrane selectivity. Similar to the PLD distribution, the LCD distribution for the MOFs with top 5% membrane selectivity is mostly localized at the smaller pore size than those with top 5% rMPS. For MOFs with top 5% membrane selectivity, the ultra-narrow pore apertures contribute to a larger membrane selectivity but a smaller diffusion coefficient and lower permeability.

Figure 3. The position and orientation of an ethene molecule along the minimum energy path (MEP) in EBINUA02 (a) and LUDLAZ (b). Here the detailed molecular structures are only for illustration. The energy landscape along the MEP for ethane and ethene in EBINUA02 (c) and LUDLAZ (d) at 300 K.

Figure 3 shows the minimum energy path (MEP), molecular orientation along the MEP, and the energy landscape for gas diffusion in MOFs with the highest membrane selectivity [EBINUA02: $k^{IM}(C_2H_4/C_2H_6)$=2167.3 and rMPS=1012.4 barrer] and the highest rMPS [LUDLAZ: $k^{IM}(C_2H_6/C_2H_4)$=11.1]. Supplementary Figure 5 presents the energy barrier along the MEP for EBINUA02 and LUDLAZ. For both EBINUA02 and LUDLAZ, only one direction along the lattice vector can accommodate the diffusion of ethane or ethene molecule. Figure 3(a) and (b) show that, despite the significant difference between EBINUA02 and LUDLAZ in the energy landscape along the MEP, their MEPs inside MOFs are almost identical. Both EBINUA02 and LUDLAZ yield near straight trajectories for the molecular center of mass on the MEP with a minimal change of the molecular orientation, suggesting that the high membrane selectivity is attributed to extremely narrow pores. It is worth noting that the local chemical environments are very similar along the MEP in EBINUA02 and LUDLAZ, and their different pore structures result in the distinct energy landscape along the MEP.

Although EBINUA02 and LUDLAZ have a similar void fraction, their pore structures (*e.g.*, PLD and LCD) are very different thus result in different separation mechanisms. According to the solution-diffusion theory,[18] the membrane selectivity can be improved by increasing the difference in adsorption, diffusion, or a combination of both quantities. EBINUA02 has a much smaller pore aperture (PLD: 2.91 Å and LCD: 3.96 Å) than LUDLAZ (PLD: 4.18 Å and LCD: 5.96 Å), which leads to the preferential diffusion and adsorption (*viz*. solubility in solution-diffusion theory) of C_2H_4 and the extremely high membrane selectivity of C_2H_4 over C_2H_6. Whereas in LUDLAZ, its

interaction with the gas molecules is attractive even at the transition state and the difference in energy barrier between C_2H_4 and C_2H_6 is almost negligible (Fig. 3D). The relatively spacious pore structure in LUDLAZ results in a slightly faster diffusion of C_2H_6 than C_2H_4 because ethane experiences a stronger van der Waals attraction. As a result, the membrane selectivity of LUDLAZ is mostly contributed by the difference in adsorption (*viz*., Henry's constant) between C_2H_4 and C_2H_6. According to the above analysis of the top MOFs with high membrane selectivity and rMPS, a large energy barrier (*viz*., extremely narrow pore aperture) is not preferred for the design of ideal MOF membrane (with both high selectivity and permeability). The trend is intuitively understandable because a narrow pore aperture significantly reduces the gas diffusion coefficient and permeability, For the rational design of ideal MOF membranes for C_2H_4/C_2H_6 separation, the selectivity and permeability need to be harnessed by enhancing the difference in adsorption and diffusion, respectively.

2.3 Inverse Design of MOF Membranes

In comparison with adsorption, the membrane process has major advantages in terms of both separation selectivity and capacity (viz. permeability). As a result, our inverse design is concerned only with MOF membranes. To find nanoporous materials ideal for C_2H_4/C_2H_6 separation (*viz*., high separation selectivity and capacity), we use a genetic algorithm (GA) with the fitness score of

$$F^{\text{total}} = 0.5F^{k^{\text{IM}}} + 0.5F^{P}. \quad (2)$$

In Equation (2), the total fitness score, F^{total}, is evenly weighted according to the member selectivity and permeability. The member selectivity fitness score is formulated as

$$F^{k^{\text{IM}}} = \begin{cases} \left(k^{\text{IM}} - 5\right)^2 + 2 & k^{\text{IM}} \geq 5 \\ e^{k^{\text{IM}} - 5} + 1 & k^{\text{IM}} < 5 \end{cases}$$

where k^{IM} stands for the (ethene-selective) membrane selectivity. Conversely, the permeability fitness score is defined by

$$F^{P} = \begin{cases} \log(P) & P \geq 100 \\ e^{P-100} + 1 & P < 100 \end{cases}.$$

In this work, the membrane selectivity of 5 and permeability of 100 barrers are used as threshold values because most CoRE MOFs have a permeability larger than 100 barrer but few have a membrane selectivity larger than 5. For materials with the permeability or selectivity lower than the corresponding threshold value, the exponential form allows for a smooth variation of the fitness score. Such a smooth form is important especially for materials with an extremely small value of permeability. Compared with improving permeability, it is much more difficult to improve the membrane selectivity. Therefore, for permeability and selectivity larger than the threshold values, the fitness score of selectivity is constructed in the quadratic form while the natural logarithm is used for the permeability. According to this formulation of the fitness score, a relatively small increase in selectivity would lead to a large increase in the fitness thus allow GA to generate MOFs with both high membrane selectivity and permeability instead of MOFs with solely high permeability.

Figure 4. Schematic illustration for chromosome representation of MOF and genetic algorithm used in this work. (a) Chromosome representation of MOFs investigated in this work where topology, node and edge are treated as genes in the chromosome. (b) Workflow of the genetic algorithm (GA) for the inverse MOF design. Here square boxes represent the secondary building blocks (SBUs) used for MOF construction. In analogy to the genes in the chromosome, the choice of SBUs directly determines the physiochemical properties and separation performance. Black, brown and green boxes represent the topology, node and edge, respectively.

Figure 4 shows how each MOF is represented as a 'chromosome', i.e., in terms of the MOF topology, node and edge. The schematic flowchart elucidates the computational steps in the inverse

design of MOF membranes. Compared with CoRE MOF database or other existing MOF structural databases, a much larger chemical space can be explored for the inverse design. In this work, the secondary building block (SB) database consists of 1,687 topologies, 648 nodes and 219 edges. As each MOF is defined by a topology, and up to two types of metal nodes and three types edges, enormous combinations are possible for the MOF construction. Although the number 7.4×10^{15} does not necessarily reflect the actual size of the design space because many combinations of the edges, nodes and topology cannot lead to the construction of synthesizable MOFs, it shows the extensiveness of the design space in enumeration of all possible combinations of the topology and building blocks. Not all chromosomes would lead to a successful MOF design owing to the unmatched coordination number and bonding distance in the topologies, nodes and edges. Invalid chromosomes were identified during the construction of MOF structures with PORMAKE based on the compatibility criteria for the building blocks and topology.[37] The percentage of invalid chromosomes is about 70% on average among all generations of chromosomes sampled by GA. In general, the successful rate increases with the GA evolution. For valid chromosomes, we obtain the MOF structure and cell parameters from PORMAKE that is able to accommodate the nodes and linkers in arbitrary topologies while avoid clashing atoms and distorting bonds. The structures generated by PORMAKE were optimized during the MOF construction. No further geometry optimization was attempted in this work.

Figure 5. Membrane selectivity vs. capacity for CoRE MOFs and inverse designed MOFs used in C_2H_4/C_2H_6 separation. (a) Membrane separation selectivity *vs*. permeability for CoRE MOFs (filled dots) and inverse designed MOFs (open symbols). The distribution of membrane selectivity (b) and permeability (c) for inversed designed MOFs. The red line denotes the Robeson boundary, and color spectrum stands for the percentile of relative membrane performance score (rMPS): the red, white and blue represent the highest, intermediate and the lowest rMPS, respectively. Brown box, purple circle, green triangle and gold star stand for inverse designed MOFs in generation 0, 1, 2, 3, respectively.

Figure 5 shows the membrane selectivity and permeability of the designed MOFs in comparison with CoRE MOFs. Here, we consider only materials in the region where the scale of permeability is similar to the experiment results. Supplementary Figure 6 shows the same figure with the full ranges of permeability and selectivity. Figure 5(a) compares the properties of MOF structures generated by GA with those from the CoRE MOF library in terms of the membrane selectivity and permeability. Clearly, GA is able to identify MOF structures with the targeted properties, i.e., high selectivity and high capacity as shown in the area above the Robeson boundary. Because no constraint was imposed in sampling the design space, the GA sampling does not always lead to the successful construction of MOF structures. For those chromosomes not generating valid MOFs, their fitness scores would be assigned to the lowest value.

Figure 5(b) and (c) present the percentages MOF structures in different generations of GA sampling. Similar plots are given in Supplementary Figure 6(b) and (c) but for the entire ranges of membrane selectivity and permeability. As the area under each curve represents the percentage of valid MOF structures in each generation, the successful evolution of GA is evident not only because it generates more valid MOF structures after each round of evolution but because the designed materials show noticeable improvement in both permeability and membrane selectivity. Although many MOF structures do not surpass the Robeson boundary even in the final generation (gold stars), GA is able to create successful candidates with high membrane selectivity and permeability. Importantly, GA is computationally much more efficient than conventional approaches of material discovery such as high-throughput screening because it avoids enumeration of the entire design space.

The best MOF identified by the inverse design (yfk-N379) significantly improves the overall membrane separation performance in terms of both membrane selectivity and permeability. Table

5 lists the detailed properties of yfk-N379. It should be noted that the best materials identified by GA may vary with the initial generation. In this work, we tested the efficiency of the inverse design with a different initial generation (shown in Supplementary Figure 8). Whereas different sets of materials were sampled during the evolution, it seems that GA sampling is robust to meet the goal of the inverse design independent of the initial condition. Although the final MOF structures are not identical, they have similar performance in terms of both membrane selectivity and permeability.

Table 5. Properties of designed MOFs with both high membrane selectivity and permeability for the separation of C_2H_4/C_2H_6 at 300 K.

Figure 6 shows molecular orientation and energy landscape for the diffusion of gas molecules along the MEP. In yfk-N379, the metal node, $(CO_2\text{-}\kappa^2O)TbO_2(\mu\text{-}CO_2\text{-}\kappa^2O)_4TbO_2(CO_2\text{-}\kappa^2O)$, is connected with the yfk topology network to form a 1D channel for the diffusion of C_2H_4 and C_2H_6 molecules. The energy landscape along the MEP in yfk-N379 is similar to that in LUDLAZ where the intermolecular interaction along the MPE is all attractive. Also, like that in LUDLAZ, the difference in the energy barrier between C_2H_4 and C_2H_6 is relatively small. As discussed above, in order to design an ideal MOF membrane with both high membrane selectivity and permeability, the separation selectivity shall be harnessed by the difference in the adsorption (*viz*., solubility), and the high permeability should be obtained by fast diffusion. As shown in Figure 6(b), the 1D channel in yfk-N379 offers strong attraction and relatively moderate energy barrier along the MEP, which results in the extremely fast diffusion of gas molecules. The distinct difference of adsorption properties (*viz*., Henry's constant) between C_2H_4 and C_2H_6 in yfk-N379 leads to an exceptionally high membrane selectivity compared to the CoRE MOFs. Although the membrane selectivity yfk-N379 is not much larger than those corresponding to the top 5 CoRE MOFs, its permeability (8420

384 barrer) is at least three orders of magnitudes higher. Conversely, the membrane selectivity of yfk-
385 N379 is 30% higher than those of CoRE MOFs with similar permeability. The synergetic effects
386 from adsorption and diffusion help yfk-N379 achieve both high membrane selectivity and
387 permeability simultaneously, making it promising for industry applications.
388
389 **Figure 6.** Minimum energy path (MEP) of C_2H_6 and C_2H_4 in yfk-N379. (a) The position and
390 orientation of an ethene molecule along the MEP (yellow line) in yfk-N379. (b) Energy landscape
391 along the MEP for ethane and ethene in yfk-N379 at 300 K.
392
393 We have performed molecular dynamics (MD) simulation to further validate the diffusion
394 coefficients of ethane and ethene in top 5 MOFs with the highest rMPS and in the best MOF
395 candidate identified from the inverse design (yfk-N379). As shown in Supplementary Figure 7,
396 the diffusion coefficients predicted by the transition-state theory agree well with the MD results.
397 In comparison with MOF-5 tested in our earlier work, these MOFs have more complicated pore
398 structure and topology, indicating the accuracy of MEPs calculated by the simplified string method.
399 Compared with high-throughput screening, the inverse design via GA is computationally much
400 more efficient. For example, the ideal MOF candidate (both high membrane selectivity and
401 permeability) for the membrane separation of C_2H_4/C_2H_6 can be found with only 8,000 attempts
402 in a much larger chemical space. The computational workflow thus demonstrates that, with the
403 efficient theoretical tools for high-throughput evaluation of materials performance, how the
404 inverse design can significantly accelerate the material discovery, especially for the construction
405 of reticular materials (*e.g.*, MOFs and COFs) for the gas storage and separation.
406 In this work, we have integrated high-throughput screening and inverse design to find the best
407 metal-organic frameworks (MOFs) for C_2H_4/C_2H_6 separation. Both adsorption and membrane
408 processes have been considered in the high-throughput screening of computational-ready

experimental (CoRE 2019) MOF database. For the adsorption separation, the separation selectivity of ethene-selective MOF decreases with the increase of separation capacity because highly ethene-selective materials have extremely small pores with low adsorption capacity. While the selectivity of ethane-selective MOF increases with the adsorption capacity, the highest adsorption selectivity of ethane-selective MOF (LUDLAZ) [$\alpha^{IM}(C_2H_6/C_2H_4)$=7.68] is smaller than that of ethene-selective MOF (PIRYOF) [$\alpha^{IM}(C_2H_4/C_2H_6)$=27.01]. Nevertheless, LUDLAZ is more than 70% higher than the state-of-art ethane-selective MOF identified by previous work.

Compared with that in the adsorption process, the selectivity of membrane process is less compromised by the increase of the separation capacity. Through high-throughput screening, we find that EBINUA02 yields the highest membrane selectivity [$k^{IM}(C_2H_4/C_2H_6)$=2167.3]. To evaluate the overall membrane performance, we introduced a relative membrane performance score (rMPS) in terms of the selectivity and permeability with respect to the Robeson boundary. For the separation of C_2H_4/C_2H_6, high rMPS favors MOFs with high permeability and intermediate membrane selectivity because high membrane selectivity requires large energy barrier along the minimum energy path (MEP) and leads to the slow diffusion. According to the structural analysis, MOFs with top 5% membrane selectivity have a much more confined diffusion path in terms of pore limit diameter (PLD) and largest cavity diameter (LCD) than those with top 5% rMPS, despite their similarity in the distribution of void fraction. The separation mechanism is quite different between EBINUA02 and the MOF (LUDLAZ) with the highest rMPS. While a small pore aperture (PLD: 2.91 Å and LCD: 3.96 Å) in EBINUA02 results in faster diffusion and stronger adsorption of C_2H_4 over C_2H_6, the less confined diffusion path in LUDLAZ (PLD: 4.18 Å and LCD: 5.96 Å) leads to negligible difference in the diffusion. In that case, the membrane selectivity is mostly contributed by its ethane-selective solubility.

The computational efficiency of the theoretical tools for predicting the sorption and diffusion properties of nanoporous materials enables the design MOF membranes with both high membrane selectivity and permeability using the genetic algorithm (GA). Compared with high-throughput screening, not only can GA explore the material design space with targeted properties, but it takes less attempts to identify the most promising candidates as well. The best MOF discovered by GA consists metal node - $(CO_2\text{-}\kappa^2O)TbO_2(\mu\text{-}CO_2\text{-}\kappa^2O)_4TbO_2(CO_2\text{-}\kappa^2O)$ with the yfk topology. The designed material has both permeability and membrane selectivity significantly larger than the threshold values set in the fitness function. Besides, its overall membrane separation performance is better than all existing experimental MOF candidates. The computational workflow used in the work thus demonstrates the capability of inverse design to accelerate the discovery of nanoporous materials, especially reticular materials (such as MOFs and COFs) for gas storage and separation.

4. Methods

4.1 Molecular Models

In this work, ethane (C_2H_6) and ethene (C_2H_4) molecules are modeled as two united-atom groups separated by a fixed bond length.[38] The detailed force field parameters can be found in Supplementary Table 1. These force-field parameters are able to reproduce the adsorption isotherms of ethane and ethene in nanoporous materials.[39] They also predict reasonable diffusion coefficients in comparison with limited experimental data. For example, the diffusivity predicted in this work agrees well with that from experiment for C_2H_6 in MOF-5 [2.42×10^{-8} m^2/s vs 1.8~2.1 $\times10^{-8}$ m^2/s (exp)]. Metal-organic frameworks (MOFs) are considered to be rigid with the universal force field (UFF) for all nonbonded interactions.[40] While the flexibility of MOFs may play an important role in determining the efficiency of gas separation, a reliable description of such effects is computationally prohibitive for high-throughput screening because it would require input from

high-level quantum-mechanical calculations.[41] Recently, a systematic examination of the impacts of MOF flexibility on molecular diffusivities indicates that the rigid-structure assumption yields reasonable diffusion coefficients for rigid molecules such as ethane and ethene investigated in the present work.[42] As mentioned above, the diffusivity of C_2H_6 in MOF-5 predicted in this work agrees well with that from experiment.

The Lennard-Jones (LJ) 12-6 potential is truncated and shifted to zero at 12.9 Å, and the Lorentz-Berthelot mixing rule is used for the energy and size parameters between different kinds of atoms. The periodic boundary condition is applied to all cell axes. The unit cell is duplicated such that the length along each lattice axis is at least two times the cutoff distance. The structural properties, such as the pore limit diameter, the largest cavity diameter and the void fraction, are calculated with Zeo++.[43]

4.2 Adsorption Separation

Nanoporous materials have been widely used as adsorbent in industrial applications. In the low pressure region, the adsorption selectivity for an equimolar mixture of two gas species can be measured with the ratio of Henry's constants[34]

$$\alpha^{\mathrm{IM}} = \frac{K_{\mathrm{h},2}}{K_{\mathrm{h},1}} \tag{3}$$

where $K_{\mathrm{h},i}$ represents the Henry's constant of component i. For a gas molecule with a rigid conformation, the Henry's constant can be calculated via the integration of the external potential due to its interaction with the nanoporous material[44]

$$K_{\mathrm{h}} = \frac{1}{8\pi^2 k_B T V} \int d\omega \int d\mathbf{r} \exp\left[-\beta \phi^{\mathrm{ext}}(\mathbf{r}, \omega)\right] \tag{4}$$

where $\beta = 1/(k_B T)$, k_B stands for the Boltzmann constant, T is the absolute temperature, V represents the system volume, ϕ^{ext} is the external potential, *i.e.*, the potential energy due to the

interaction of a gas molecule with all atoms from the porous material, **r** represents the cartesian coordinates for the center of mass of the gas molecule, and ω stands for its Euler angles. For each MOF, the Henry's constants for C_2H_6 and C_2H_4 are numerically evaluated via midpoint rule with the step size of 1 Å and 45° for spatial and rotational variables, respectively.

4.3 Membrane Separation

According to the solution-diffusion model, the membrane permeability is defined as the Henry's constant multiplied by the gas diffusion coefficient at infinite dilution. The membrane selectivity can thus be calculated from[44]

$$k^{\mathrm{IM}} = \frac{K_{\mathrm{h},2}}{K_{\mathrm{h},1}} \frac{D_{0,2}}{D_{0,1}} = \frac{P_2}{P_1} \tag{5}$$

where $D_{0,i}$ stands for the diffusion coefficient of component i at infinite dilution , and P_i represent the permeability. In evaluating the membrane selectivity, we use the average of diffusion coefficients along the three lattice vectors (*viz*., *x*-, *y*-, *z*-axis in cartesian coordinate if the lattice vectors are mutually orthogonal)

$$D_0 = \frac{D_{0,a} + D_{0,b} + D_{0,c}}{3}. \tag{6}$$

Along each direction, the diffusivity can be calculated independently according to the transition-state theory (TST)

$$D_{0,\alpha} = \frac{1}{2} k a_\alpha^2 . \tag{7}$$

where α stands for the direction of lattice vector, k represents the transmission rate (*viz*., the hopping rate of the gas molecule), and a_α stands for the hopping distance between neighboring unit cells along the direction of lattice vector α. The hopping rate can be obtained from the minimum energy path (MEP) via the Bennett-Chandler formula[45,46]

$$k = \sqrt{\frac{k_B T}{2\pi m}} \frac{\exp\left[-\beta\phi^{\text{ext}}\left(s^*\right)\right]}{\int_0^1 \exp\left[-\beta\phi^{\text{ext}}\left(s\right)\right] ds}. \quad (8)$$

where s is a normalized dimensionless variable along the MEP, and s^* represents the transition state. The diffusion coefficient predicted by Equation (7) depends heavily on the accuracy of the MEP.

The mathematical details and the accuracy of MEP calculations for predicting the diffusivity of rigid molecules, such as C_2H_4 and C_2H_6, can be found in our previous work.[25] Here, we recapitulate only the key steps to identify the MEP via the simplified string method. First, we locate the starting point of the string (i.e., the diffusion path or MEP) by searching the position and orientation of a gas molecule that minimize the external potential at the entrance plane. Because of the periodic structure of MOFs, the starting and end points of the string have identical relative configurations within their own unit cells. Next, we construct an initial string by evenly placing a certain number of points (viz., images) between the starting and end points. Finally, MEP is obtained by iterative evolution of the images along the string according to the gradient of the full external potential. After each round of iteration, the string is renormalized to avoid images collectively falling into any local energy minima.

All diffusion coefficients reported in this work are predicted from the MEP obtained from the GPU-accelerated simplified string method. We have also validated the diffusion coefficients calculated with our method by carrying out molecular dynamics (MD) simulation for top 5 MOFs with highest rMPS and for the best inverse designed MOF. For all MD simulations, we used the LAMMPS package with the Nose-Hoover thermostat for controlling the temperature. The detailed settings follow the protocol suggested in the literature for infinite dilution.[47]

4.4 Genetic Algorithm

For inverse design, we use the genetic algorithm (GA) to construct MOFs with desired properties. Because MOFs can be decomposed into the secondary building blocks (SBUs), each material may be considered as a 'chromosome' of different topologies, metal nodes and organic linkers. Because the supermajority of MOF topologies can accommodate no more than two types of metal nodes and three types of organic linkers, each chromosome consists of 6 genes, and each gene is represented by an integer which corresponds to a specific topology, node or linker. The chromosome representation allows for the efficient sampling of all possible combinations of topology and SBUs with GA. The population is set as 2,000 which makes the initial generation to have diverse combinations of topologies, nodes and linkers. Three evolutions are carried after the initial population and a total of 8,000 combinations of topologies, nodes and linkers is explored to find the optimal MOF structure with desired properties. Compared with the MOF structural database used in high-throughput screening, less MOF structures are used in the GA to benchmark its computational performance for the inverse design despite a much larger chemical space is considered. In the initial population, the 2,000 chromosomes are generated by the random selection of topologies, nodes and edges in the SBU database.[37] In each generation, 10 MOFs are used to generate the offspring via single point crossover. The next generation of MOFs are selected by stochastic universal selection to avoid bias towards the SBUs with low fitness values while 30% MOFs would have a random mutation on their genes. In this work, PyGAD library is used for the genetic algorithm.[48] PORMAKE is used to construct MOFs when the chromosome values are assigned.[37]

Data availability

The authors declare that the data supporting this study are available within the paper and its Supplementary Information files. Additional data that support the findings of this study are available from the corresponding author upon reasonable request.

Code availability

The code used in this work can be accessed from the associated GitHub repository (https://github.com/MusenZhou/inverse-design-MOF-membrane).

Acknowledgement

This work is financially supported by the National Science Foundation's Harnessing the Data Revolution (HDR) Big Ideas Program under Grant No. NSF 1940118. We acknowledge the NVIDIA corporation for donation of NVIDIA A100 GPU through NVIDIA Academic Hardware Grant Program.

Author Contributions

M. Z. and J. W. conceptualized the project and designed the methodology. M. Z. wrote the code and performed the computation including collection of literature data. J.W. procured funding and supervised the project. Both authors were involved with the writing of the paper.

Competing Interests

The Authors declare no Competing Financial or Non-Financial Interests.

Reference

1 Sholl, D. S. & Lively, R. P. Seven chemical separations to change the world. *Nature* **532**, 435-437 (2016). https://doi.org:10.1038/532435a

2 Rungta, M., Zhang, C., Koros, W. J. & Xu, L. Membrane ‐ based ethylene/ethane separation: The upper bound and beyond. *AIChE J.* **59**, 3475-3489 (2013).

3 Ding, Q. *et al.* Exploiting equilibrium-kinetic synergetic effect for separation of ethylene and ethane in a microporous metal-organic framework. *Sci. Adv.* **6**, eaaz4322 (2020). https://doi.org:10.1126/sciadv.aaz4322

4 Li, L. *et al.* Ethane/ethylene separation in a metal-organic framework with iron-peroxo sites. *Science* **362**, 443-446 (2018). https://doi.org:10.1126/science.aat0586

5 Lv, D. *et al.* Recent advances in adsorptive separation of ethane and ethylene by C2H6-selective MOFs and other adsorbents. *Chem. Eng. J.*, 133208 (2021).

6 Rangnekar, N., Mittal, N., Elyassi, B., Caro, J. & Tsapatsis, M. Zeolite membranes - a review and comparison with MOFs. *Chem. Soc. Rev.* **44**, 7128-7154 (2015). https://doi.org:10.1039/c5cs00292c

7 Li, J. R., Sculley, J. & Zhou, H. C. Metal-organic frameworks for separations. *Chem. Rev.* **112**, 869-932 (2012). https://doi.org:10.1021/cr200190s

8 Tranchemontagne, D. J., Mendoza-Cortes, J. L., O'Keeffe, M. & Yaghi, O. M. Secondary building units, nets and bonding in the chemistry of metal-organic frameworks. *Chem. Soc. Rev.* **38**, 1257-1283 (2009). https://doi.org:10.1039/b817735j

9 Zhou, M., Vassallo, A. & Wu, J. Toward the inverse design of MOF membranes for efficient D2/H2 separation by combination of physics-based and data-driven modeling. *J. Membr. Sci.* **598** (2020). https://doi.org:10.1016/j.memsci.2019.117675

10 Wang, J., Zhou, M., Lu, D., Fei, W. & Wu, J. Virtual Screening of Nanoporous Materials for Noble Gas Separation. *ACS Appl. Nano Mater.* **5**, 3701-3711 (2022). https://doi.org:10.1021/acsanm.1c03907

11 Bárcia, P. S. *et al.* Reverse shape selectivity in the adsorption of hexane and xylene isomers in MOF UiO-66. *Microporous Mesoporous Mater.* **139**, 67-73 (2011).

12 Oh, H. *et al.* A cryogenically flexible covalent organic framework for efficient hydrogen isotope separation by quantum sieving. *Angew. Chem. Int. Ed.* **52**, 13219-13222 (2013). https://doi.org:10.1002/anie.201307443

13 Chen, G. *et al.* M-gallate MOF/6FDA-polyimide mixed-matrix membranes for C2H4/C2H6 separation. *J. Membr. Sci.* **620**, 118852 (2021).

14 Bao, Z. *et al.* Molecular Sieving of Ethane from Ethylene through the Molecular Cross‐Section Size Differentiation in Gallate‐based Metal‐Organic Frameworks. *Angew. Chem. Int. Ed.* **57**, 16020-16025 (2018).

15 Zhou, M., Tian, Y., Fei, W. & Wu, J. Fractionation of Isotopic Methanes with Metal–Organic Frameworks. *J. Phys. Chem. C* **123**, 7397-7407 (2019). https://doi.org:10.1021/acs.jpcc.8b11393

16 Wang, J., Zhou, M., Lu, D., Fei, W. & Wu, J. Computational screening and design of nanoporous membranes for efficient carbon isotope separation. *Green Energy Environ.* **5**, 364-373 (2020). https://doi.org:10.1016/j.gee.2020.07.025

17 Kang, M. *et al.* High-Throughput Discovery of Ni(IN)(2) for Ethane/Ethylene Separation. *Adv. Sci.* **8**, e2004940 (2021). https://doi.org:10.1002/advs.202004940

18 Wijmans, J. G. & Baker, R. W. The solution-diffusion model: a review. *J. Membr. Sci.* **107**, 1-21 (1995).

19 Altintas, C. & Keskin, S. Computational screening of MOFs for C2H6/C2H4 and C2H6/CH4 separations. *Chem. Eng. Sci.* **139**, 49-60 (2016).

20 Tang, H. & Jiang, J. In silico screening and design strategies of ethane‐selective metal‐organic frameworks for ethane/ethylene separation. *AIChE J.* **67**, e17025 (2021).

21 Halder, P. & Singh, J. K. High-throughput screening of metal–organic frameworks for ethane–ethylene separation using the machine learning technique. *Energy Fuels* **34**, 14591-14597 (2020).

22 Kim, J. *et al.* Large-scale computational screening of zeolites for ethane/ethene separation. *Langmuir* **28**, 11914-11919 (2012). https://doi.org:10.1021/la302230z

23 Kärger, J., Ruthven, D. M. & Theodorou, D. N. *Diffusion in Nanoporous Materials*. (WILEY-VCH, 2012).

24 Verploegh, R. J., Nair, S. & Sholl, D. S. Temperature and loading-dependent diffusion of light hydrocarbons in ZIF-8 as predicted through fully flexible molecular simulations. *J. Am. Chem. Soc.* **137**, 15760-15771 (2015).

25 Zhou, M. & Wu, J. Massively Parallel GPU-Accelerated String Method for Fast and Accurate Prediction of Molecular Diffusivity in Nanoporous Materials. *ACS Appl. Nano Mater.* **4**, 5394-5403 (2021). https://doi.org:10.1021/acsanm.1c00727

26 Zhou, M. & Wu, J. A GPU implementation of classical density functional theory for rapid prediction of gas adsorption in nanoporous materials. *J. Chem. Phys.* **153**, 074101 (2020). https://doi.org:10.1063/5.0020797

27 Glaser, J. *et al.* Strong scaling of general-purpose molecular dynamics simulations on GPUs. *Comput. Phys. Commun.* **192**, 97-107 (2015).

28 Kutzner, C. *et al.* More bang for your buck: Improved use of GPU nodes for GROMACS 2018. *J. Comput. Chem.* **40**, 2418-2431 (2019). https://doi.org:10.1002/jcc.26011

29 Kim, B., Lee, S. & Kim, J. Inverse design of porous materials using artificial neural networks. *Sci. Adv.* **6**, eaax9324 (2020). https://doi.org:10.1126/sciadv.aax9324

30 Yao, Z. *et al.* Inverse design of nanoporous crystalline reticular materials with deep generative models. *Nat. Mach. Intell.* **3**, 76-86 (2021). https://doi.org:10.1038/s42256-020-00271-1

31 Gustafson, J. A. & Wilmer, C. E. Intelligent Selection of Metal-Organic Framework Arrays for Methane Sensing via Genetic Algorithms. *ACS Sens.* **4**, 1586-1593 (2019). https://doi.org:10.1021/acssensors.9b00268

32 Chung, Y. G. *et al.* In silico discovery of metal-organic frameworks for precombustion CO(2) capture using a genetic algorithm. *Sci. Adv.* **2**, e1600909 (2016). https://doi.org:10.1126/sciadv.1600909

33 Dou, H. *et al.* Boron Nitride Membranes with a Distinct Nanoconfinement Effect for Efficient Ethylene/Ethane Separation. *Angew. Chem. Int. Ed.* **58**, 13969-13975 (2019). https://doi.org:10.1002/anie.201907773

34 Cotterman, R. L., Schwarz, B. J. & Prausnitz, J. M. Molecular thermodynamics for fluids at low and high densities. Part I: Pure fluids containing small or large molecules. *AIChE J.* **32**, 1787-1798 (1986). https://doi.org:10.1002/aic.690321104

35 McKellar, S. C. *et al.* The effect of pressure on the post-synthetic modification of a nanoporous metal-organic framework. *Nanoscale* **6**, 4163-4173 (2014). https://doi.org:10.1039/c3nr04161a

36 Tian, L., Chen, Z., Yu, A., Song, H.-B. & Cheng, P. Novel flexible bis-triazole bridged copper (II) coordination polymers varying from one-to three-dimensionality. *CrystEngComm* **14**, 2032-2039 (2012).

37 Lee, S. *et al.* Computational Screening of Trillions of Metal-Organic Frameworks for High-Performance Methane Storage. *ACS Appl. Mater. Interfaces* **13**, 23647-23654 (2021). https://doi.org:10.1021/acsami.1c02471

38 Wu, Y., Chen, H., Liu, D., Qian, Y. & Xi, H. Adsorption and separation of ethane/ethylene on ZIFs with various topologies: Combining GCMC simulation with the ideal adsorbed solution theory (IAST). *Chem. Eng. Sci.* **124**, 144-153 (2015). https://doi.org:10.1016/j.ces.2014.07.019

39 Bux, H., Chmelik, C., Krishna, R. & Caro, J. Ethene/ethane separation by the MOF membrane ZIF-8: molecular correlation of permeation, adsorption, diffusion. *J. Membr. Sci.* **369**, 284-289 (2011).
40 Rappe, A. K., Casewit, C. J., Colwell, K. S., Goddard, W. A. & Skiff, W. M. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *J. Am. Chem. Soc.* **114**, 10024-10035 (2002). https://doi.org:10.1021/ja00051a040
41 Heinen, J. & Dubbeldam, D. On flexible force fields for metal-organic frameworks: Recent developments and future prospects. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **8**, e1363 (2018). https://doi.org:10.1002/wcms.1363
42 Yang, Y. & Sholl, D. S. A systematic examination of the impacts of MOF flexibility on intracrystalline molecular diffusivities. *J. Mater. Chem. A* **10**, 4242-4253 (2022).
43 Willems, T. F., Rycroft, C. H., Kazi, M., Meza, J. C. & Haranczyk, M. Algorithms and tools for high-throughput geometry-based analysis of crystalline porous materials. *Microporous Mesoporous Mater.* **149**, 134-141 (2012). https://doi.org:10.1016/j.micromeso.2011.08.020
44 Haldoupis, E., Nair, S. & Sholl, D. S. Efficient calculation of diffusion limitations in metal organic framework materials: a tool for identifying materials for kinetic separations. *J. Am. Chem. Soc.* **132**, 7528-7539 (2010). https://doi.org:10.1021/ja1023699
45 Bennett, C. Molecular Dynamics and Transition State Theory: The Simulation of Infrequent Events in Algorithms for Chemical Computations *ACS Symposium Series* Vol. 46 Ch. 4, 63-97 (American Chemical Society, 1977).
46 Chandler, D. Statistical mechanics of isomerization dynamics in liquids and the transition state approximation. *J. Chem. Phys.* **68**, 2959-2970 (1978).
47 Xu, H., Cabriolu, R. & Smit, B. Effects of Degrees of Freedom on Calculating Diffusion Properties in Nanoporous Materials. *J. Chem. Theory Comput.* **18**, 2826-2835 (2022). https://doi.org:10.1021/acs.jctc.2c00094
48 Gad, A. F. Pygad: An intuitive genetic algorithm python library. Preprint at https://arxiv.org/abs/2106.06158 (2021).

Figure captions

Figure 1. Selectivity vs. capacity for CoRE MOFs used in C_2H_4/C_2H_6 separation. (a) Adsorption selectivity. The green line is fitted with $\alpha^{IM}=0.2516e^{-0.7411\log(K)}$. (b) Membrane selectivity. The red line denotes the Robeson boundary, and the color stands for the percentile of relative membrane performance score (rMPS): the red, gray and blue represent the highest, intermediate, and the lowest rMPS, respectively. The dashed line marks the membrane selectivity of 1.

Figure 2. Distributions of the pore limit diameter (a) and the largest cavity diameter (b) for all CoRE MOFs and MOFs with top 5% ideal membrane selectivity and relative performance score (rMPS) for C_2H_4/C_2H_6 separation.

Figure 3. The position and orientation of an ethene molecule along the minimum energy path (MEP) in EBINUA02 (a) and LUDLAZ (b). Here the detailed molecular structures are only for

illustration. The energy landscape along the MEP for ethane and ethene in EBINUA02 (c) and LUDLAZ (d) at 300 K.

Figure 4. Schematic illustration for chromosome representation of MOF and genetic algorithm used in this work. (a) Chromosome representation of MOFs investigated in this work where topology, node and edge are treated as genes in the chromosome. (b) Workflow of the genetic algorithm (GA) for the inverse MOF design. Here square boxes represent the secondary building blocks (SBUs) used for MOF construction. In analogy to the genes in the chromosome, the choice of SBUs directly determines the physiochemical properties and separation performance. Black, brown and green boxes represent the topology, node and edge, respectively.

Figure 5. Membrane selectivity vs. capacity for CoRE MOFs and inverse designed MOFs used in C_2H_4/C_2H_6 separation. (a) Membrane separation selectivity *vs*. permeability for CoRE MOFs (filled dots) and inverse designed MOFs (open symbols). The distribution of membrane selectivity (b) and permeability (c) for inversed designed MOFs. The red line denotes the Robeson boundary, and color spectrum stands for the percentile of relative membrane performance score (rMPS): the red, white and blue represent the highest, intermediate and the lowest rMPS, respectively. Brown box, purple circle, green triangle and gold star stand for inverse designed MOFs in generation 0, 1, 2, 3, respectively.

Figure 6. Minimum energy path (MEP) of C_2H_6 and C_2H_4 in yfk-N379. (a) The position and orientation of an ethene molecule along the MEP (yellow line) in yfk-N379. (b) Energy landscape along the MEP for ethane and ethene in yfk-N379 at 300 K.

Table 1. Henry's constants (K_h), ideal selectivity (α^{IM}) and self-diffusivity (D_0) of top ethene-selective MOFs for adsorption separation of C_2H_4/C_2H_6 at 300 K (The diffusion coefficients are shown only if they are larger than 1×10^{-20} $m^2\cdot s^{-1}$)

MOF	$K_h(C_2H_4)$ ($mol\cdot m^{-3}\cdot Pa^{-1}$)	$K_h(C_2H_6)$ ($mol\cdot m^{-3}\cdot Pa^{-1}$)	$\alpha^{IM}(C_2H_4/C_2H_6)$	$D_0(C_2H_4)$ ($m^2\cdot s^{-1}$)	$D_0(C_2H_6)$ ($m^2\cdot s^{-1}$)
PIRYOF	5.474×10^{-2}	2.027×10^{-3}	27.01		
BADHIA	3.454×10^{-4}	1.325×10^{-5}	26.07		
BADHOG	4.205×10^{-4}	1.766×10^{-5}	23.81		
EBINUA02	1.682×10^{-2}	9.829×10^{-4}	17.11	9.342×10^{-15}	7.377×10^{-17}
FEDKAB	9.997×10^{-2}	6.036×10^{-3}	16.56		

Table 2. Top ethane-selective MOFs for the separation of C_2H_4/C_2H_6 at 300 K by gas adsorption

MOF	$K_h(C_2H_4)$ ($mol \cdot m^{-3} \cdot Pa^{-1}$)	$K_h(C_2H_6)$ ($mol \cdot m^{-3} \cdot Pa^{-1}$)	$\alpha^{IM}(C_2H_6/C_2H_4)$	$D_0(C_2H_4)$ ($m^2 \cdot s^{-1}$)	$D_0(C_2H_6)$ ($m^2 \cdot s^{-1}$)
LUDLAZ	4.374×10^1	3.360×10^2	7.68	4.350×10^{-9}	6.313×10^{-9}
EFILUA	1.863×10^2	1.301×10^3	6.98	1.956×10^{-9}	1.188×10^{-9}
XUJSAY	3.360×10^1	2.107×10^2	6.27	6.291×10^{-9}	2.145×10^{-9}
ZAZNUL	2.735×10^1	1.697×10^2	6.20	6.604×10^{-9}	2.134×10^{-9}
KAXQIL	5.689×10^1	3.479×10^2	6.12	1.362×10^{-9}	1.088×10^{-9}

Table 3. Top MOFs for C_2H_4/C_2H_6 separation with the highest membrane selectivity (k^{IM}) at 300 K

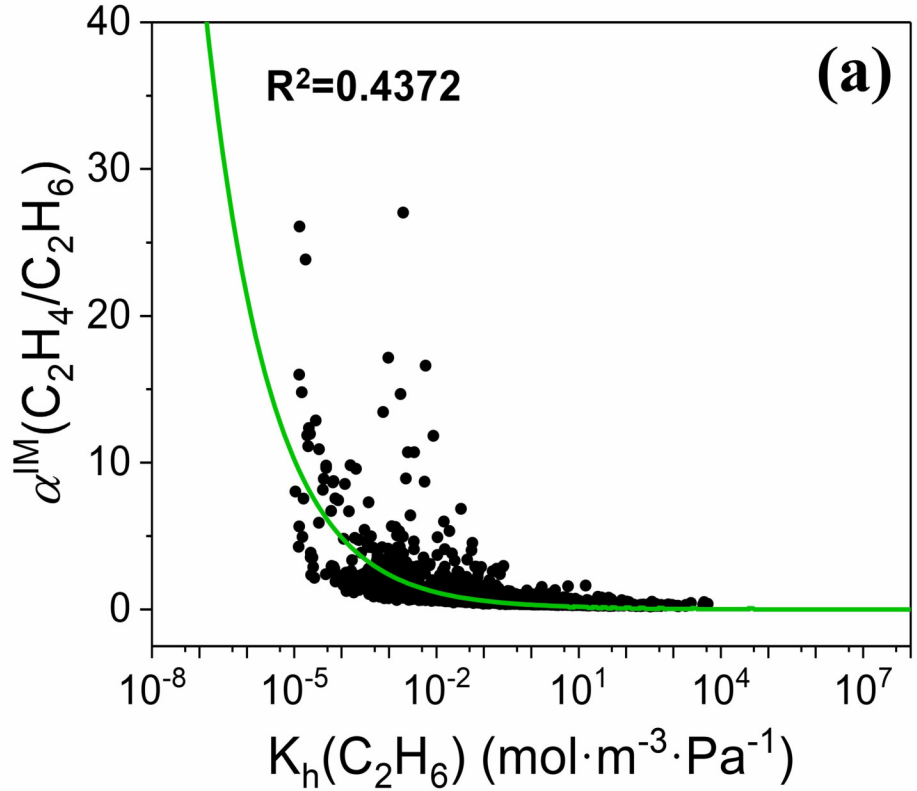
MOF	$K_h(C_2H_4)$ ($mol \cdot m^{-3} \cdot Pa^{-1}$)	$K_h(C_2H_6)$ ($mol \cdot m^{-3} \cdot Pa^{-1}$)	$D_0(C_2H_4)$ ($m^2 \cdot s^{-1}$)	$D_0(C_2H_6)$ ($m^2 \cdot s^{-1}$)	$k^{IM}(C_2H_4/C_2H_6)$
EBINUA02	1.682×10^{-2}	9.828×10^{-4}	9.342×10^{-15}	7.377×10^{-17}	2167.3
HAZGOF	9.344×10^{-4}	1.939×10^{-4}	1.438×10^{-12}	4.202×10^{-15}	1649.1
ALOLES	2.770×10^{-2}	6.686×10^{-2}	8.276×10^{-15}	3.271×10^{-18}	1048.2
EBINUA	1.030×10^{-2}	7.678×10^{-4}	4.007×10^{-14}	5.599×10^{-16}	960.1
EBINUA01	2.056×10^{-2}	2.320×10^{-3}	3.033×10^{-13}	6.128×10^{-15}	438.2

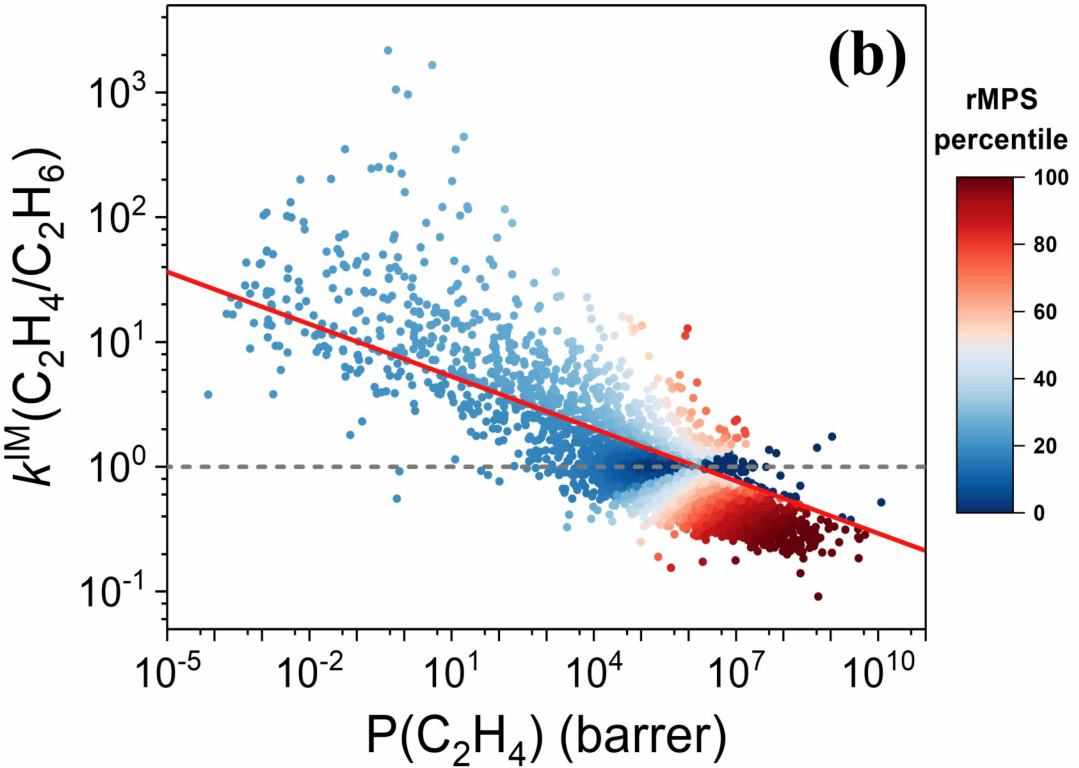
Table 4. Top MOFs with the highest relative membrane performance score (rMPS) for the separation of C_2H_4/C_2H_6 at 300 K

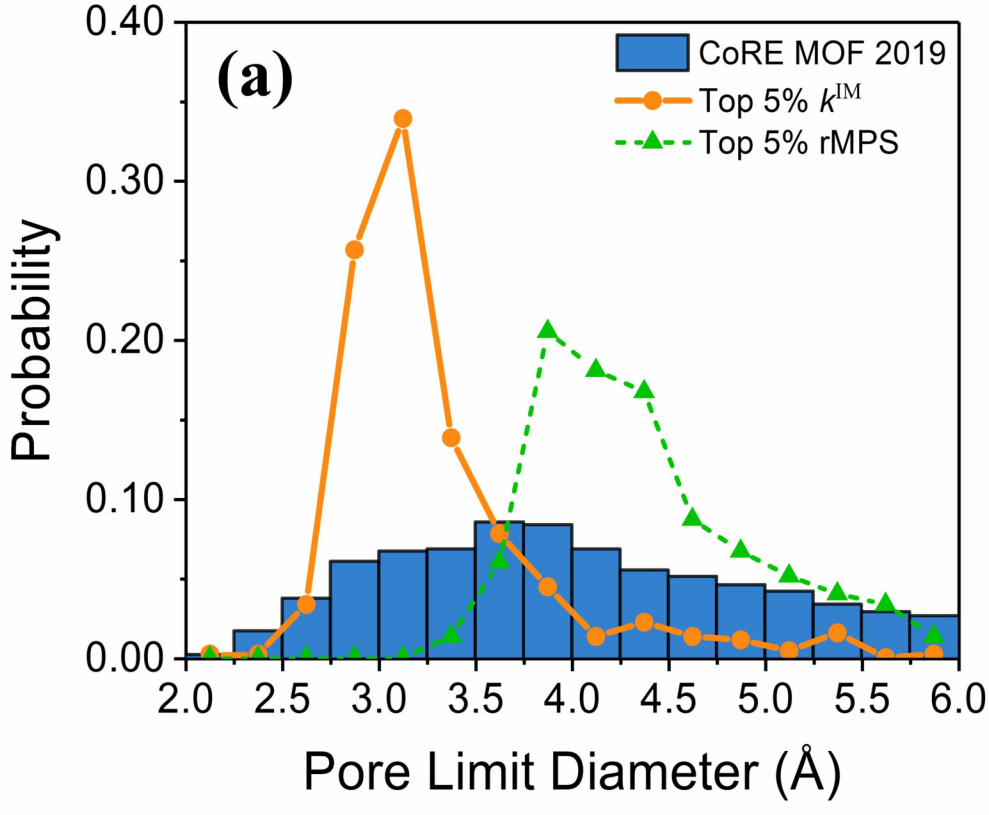
MOF	$K_h(C_2H_4)$ ($mol \cdot m^{-3} \cdot Pa^{-1}$)	$K_h(C_2H_6)$ ($mol \cdot m^{-3} \cdot Pa^{-1}$)	$D_0(C_2H_4)$ ($m^2 \cdot s^{-1}$)	$D_0(C_2H_6)$ ($m^2 \cdot s^{-1}$)	$k^{IM}(C_2H_6/C_2H_4)$	rMPS (barrer) $\times 10^{10}$
LUDLAZ	4.374×10^{1}	3.360×10^{2}	4.350×10^{-9}	6.313×10^{-9}	11.1	5.597
PARMIG	9.666×10^{1}	5.599×10^{2}	1.404×10^{-8}	1.333×10^{-8}	5.5	5.484
BEKSAM	4.511×10^{2}	1.892×10^{3}	9.052×10^{-9}	4.231×10^{-9}	2.0	1.339
MIMVEJ	2.869×10^{1}	1.443×10^{2}	1.291×10^{-8}	1.276×10^{-8}	5.0	1.268
MORZID	1.842×10^{2}	8.200×10^{2}	7.613×10^{-9}	6.568×10^{-9}	3.8	1.046

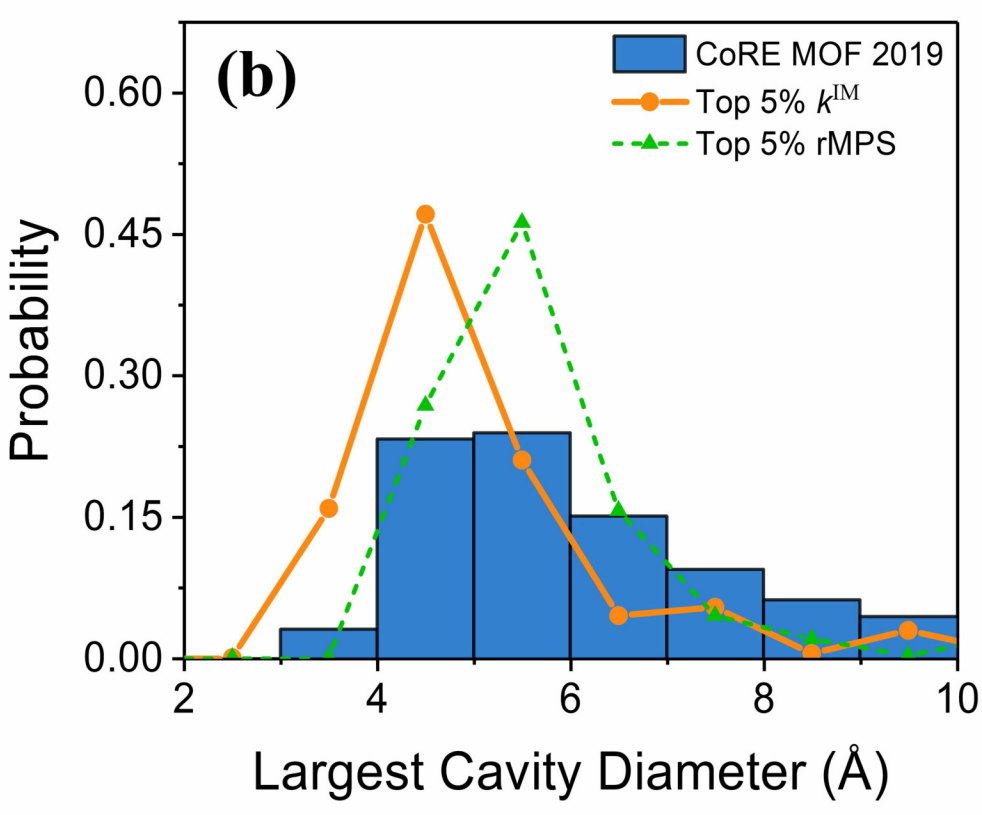
Table 5. Properties of designed MOFs with both high membrane selectivity and permeability for the separation of C_2H_4/C_2H_6 at 300 K

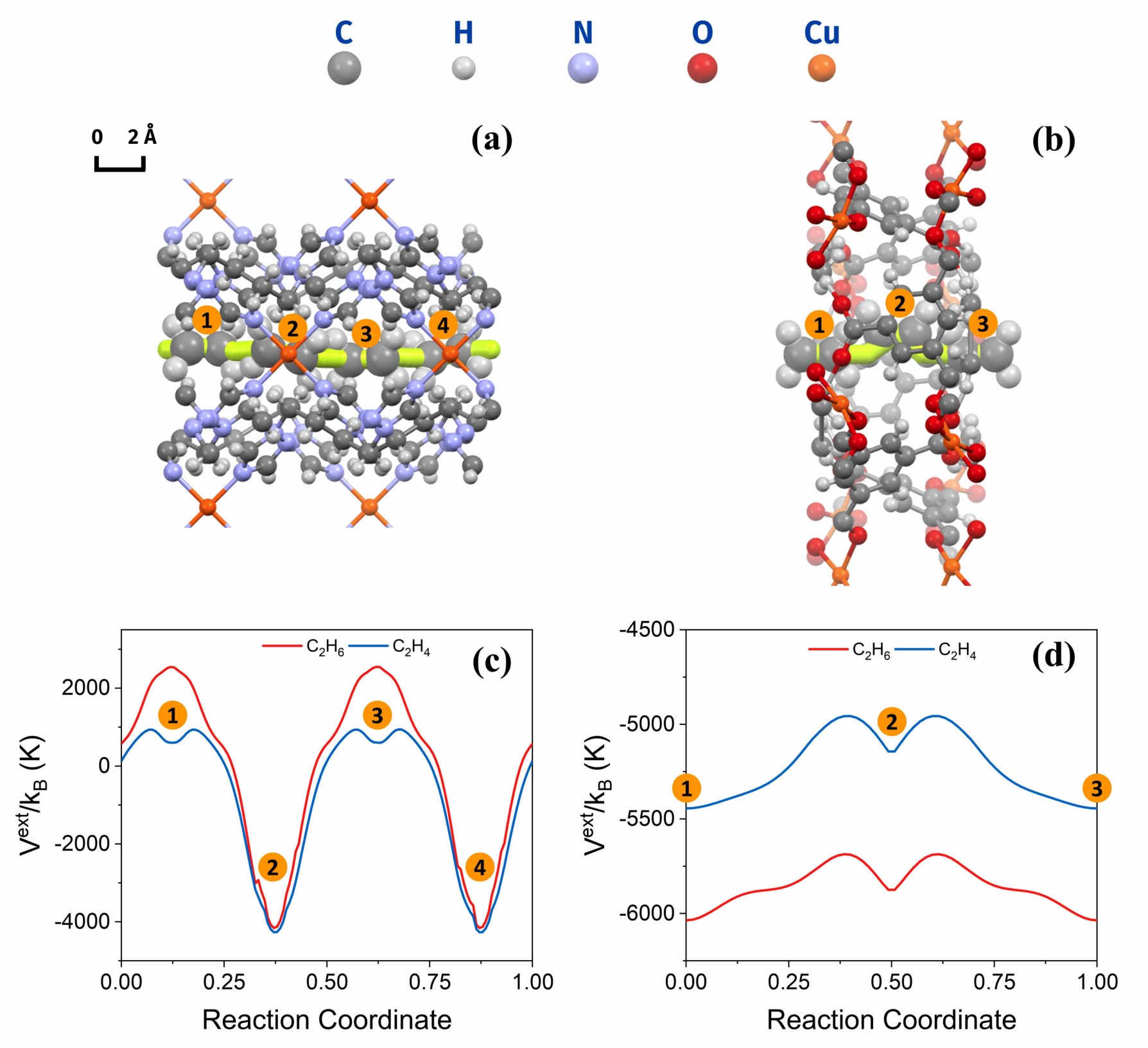
MOF	$K_h(C_2H_4)$ ($mol \cdot m^{-3} \cdot Pa^{-1}$)	$K_h(C_2H_6)$ ($mol \cdot m^{-3} \cdot Pa^{-1}$)	$D_0(C_2H_4)$ ($m^2 \cdot s^{-1}$)	$D_0(C_2H_6)$ ($m^2 \cdot s^{-1}$)	$k^{IM}(C_2H_4/C_2H_6)$
yfk-N379	1.682×10^{-2}	9.828×10^{-4}	1.677×10^{-9}	1.467×10^{-9}	20.8

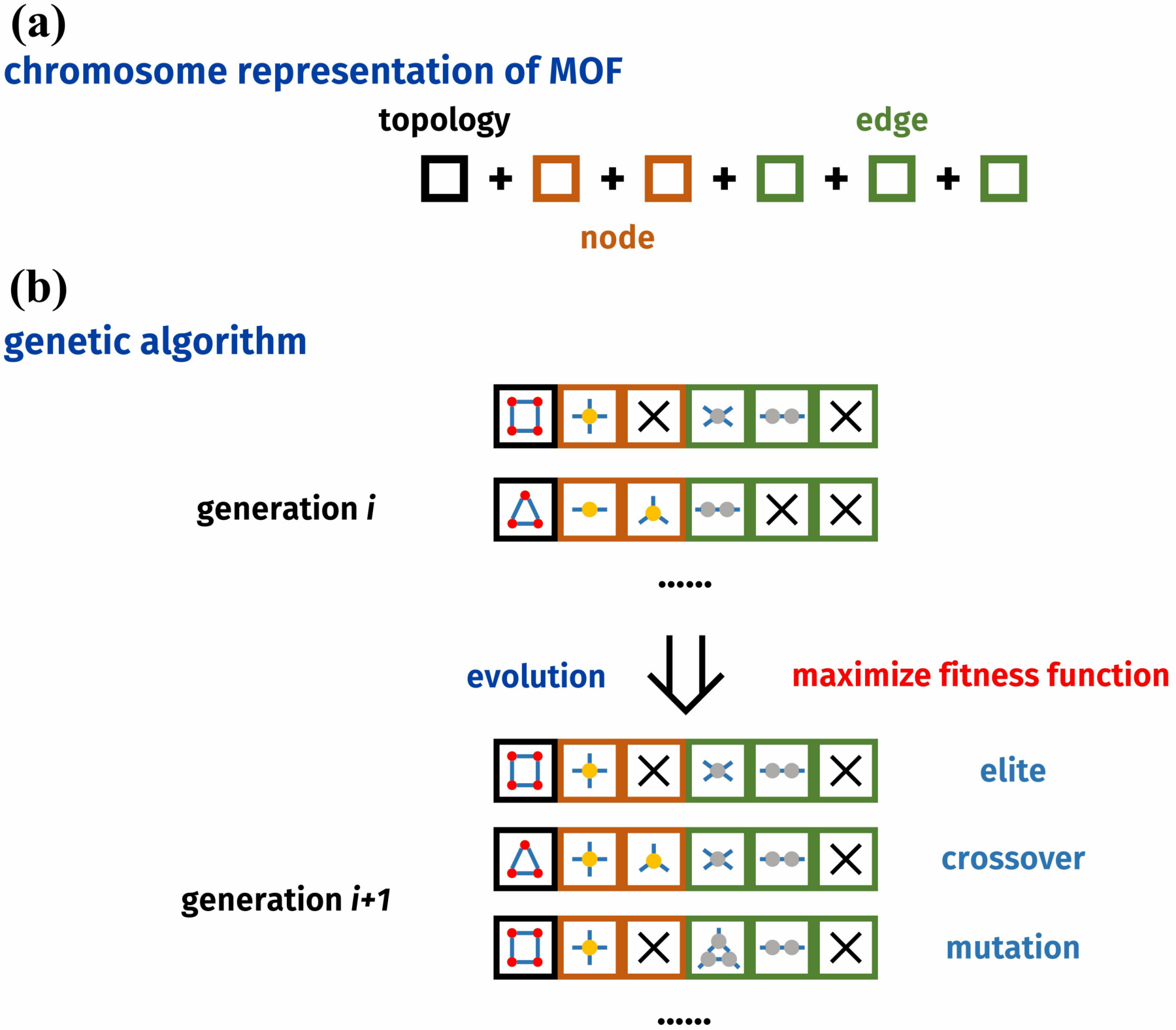

(a)
R²=0.4372
40
30
20
10
0
$\alpha^{IM}(C_2H_4/C_2H_6)$
10^{-8}
10^{-5}
10^{-2}
10^{1}
10^{4}
10^{7}
$K_h(C_2H_6)$ (mol·m^{-3}·Pa^{-1})

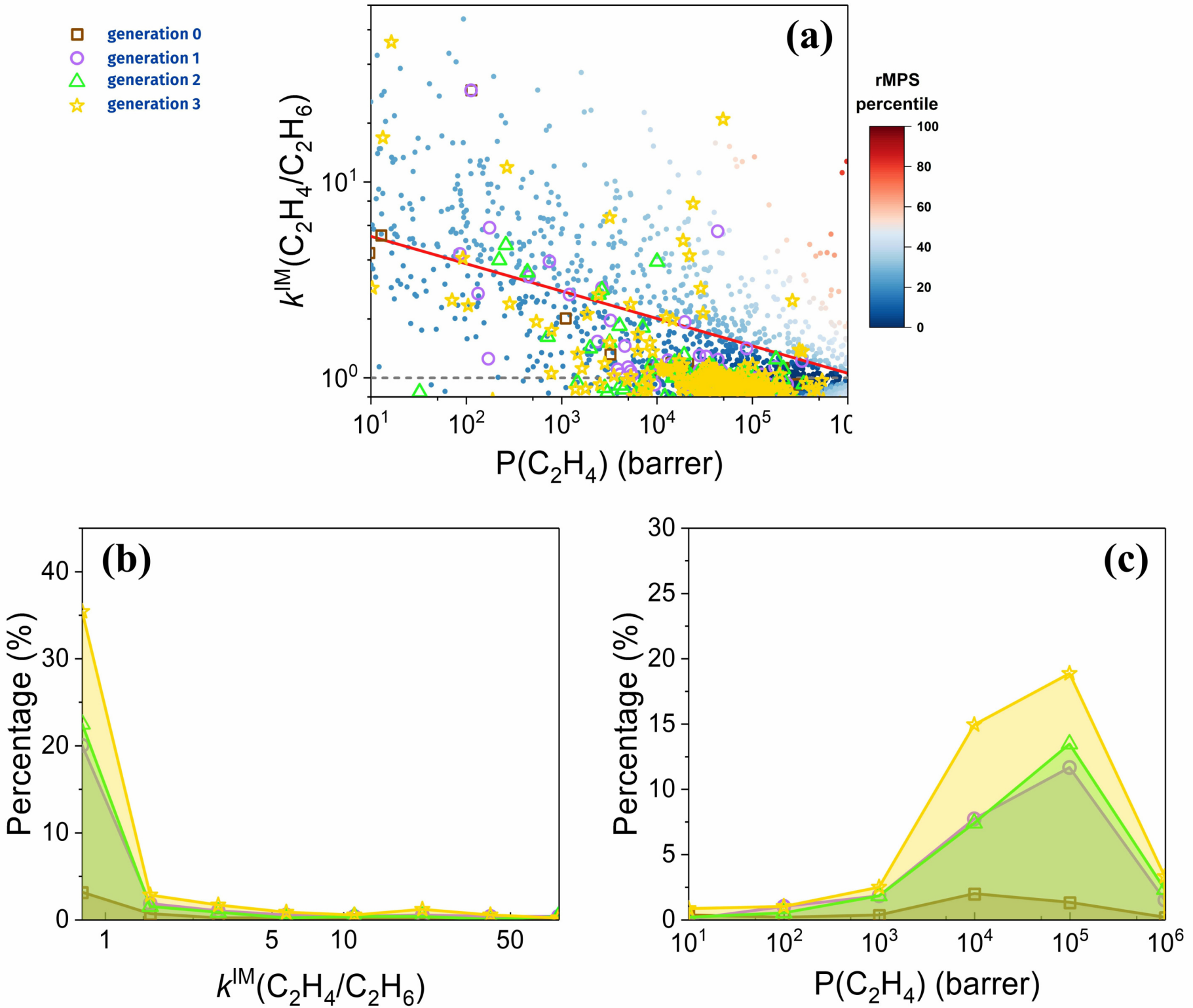

(b)
rMPS percentile
100
80
60
40
20
0
$k^{IM}(C_2H_4/C_2H_6)$
10^{3}
10^{2}
10^{1}
10^{0}
10^{-1}
10^{-5}
10^{-2}
10^{1}
10^{4}
10^{7}
10^{10}
$P(C_2H_4)$ (barrer)

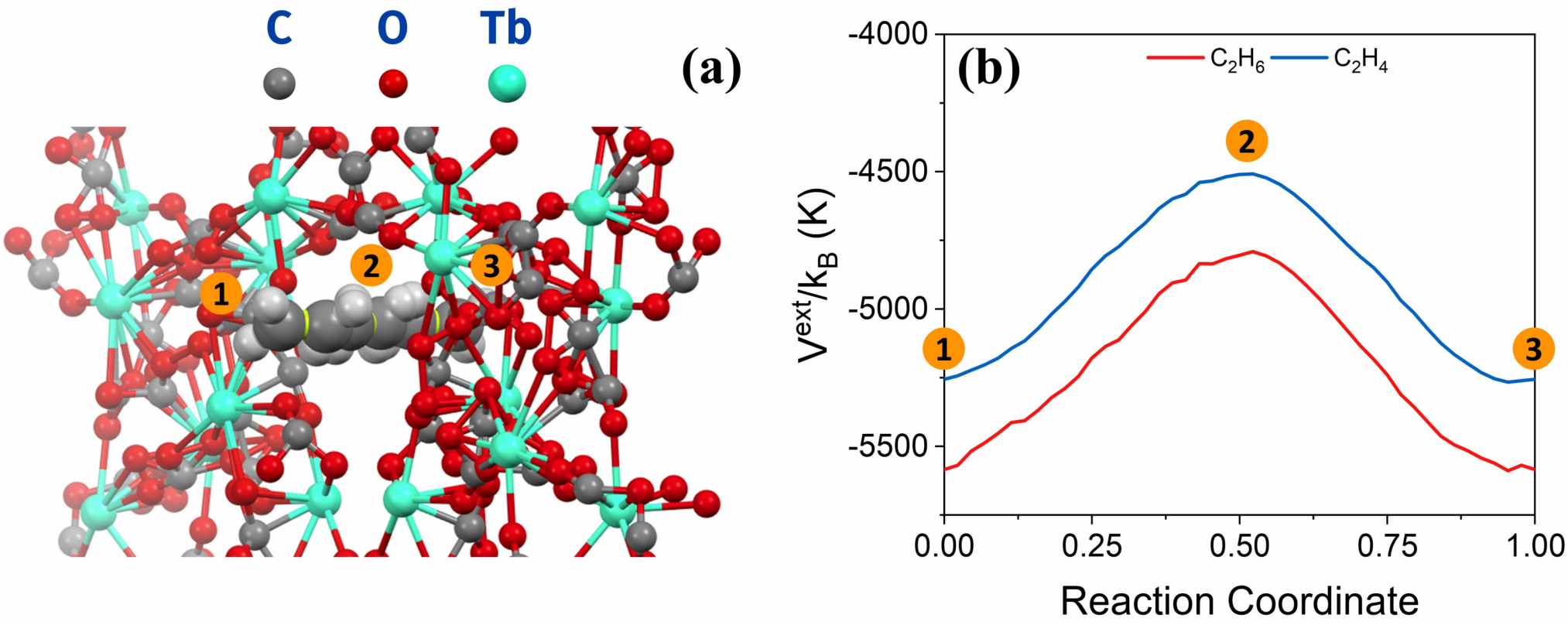

(a)
CoRE MOF 2019
Top 5% k^{IM}
Top 5% rMPS
Probability
0.40
0.30
0.20
0.10
0.00
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
Pore Limit Diameter (Å)


(b)
CoRE MOF 2019
Top 5% k^{IM}
Top 5% rMPS
Probability
0.60
0.45
0.30
0.15
0.00
2
4
6
8
10
Largest Cavity Diameter (Å)


C
H
N
O
Cu
0 2 Å
(a)
(b)
1
2
3
4
(c)
(d)
C_2H_6
C_2H_4
V^{ext}/k_B (K)
Reaction Coordinate
2000
0
-2000
-4000
-4500
-5000
-5500
-6000
0.00
0.25
0.50
0.75
1.00


(a)
chromosome representation of MOF
topology
edge
node
(b)
genetic algorithm
generation i
.....
evolution
maximize fitness function
generation i+1
elite
crossover
mutation
.....


generation 0
generation 1
generation 2
generation 3
(a)
rMPS percentile
100
80
60
40
20
0
$k^{IM}(C_2H_4/C_2H_6)$
$P(C_2H_4)$ (barrer)
(b)
Percentage (%)
$k^{IM}(C_2H_4/C_2H_6)$
(c)
Percentage (%)
$P(C_2H_4)$ (barrer)


C
O
Tb
(a)
1
2
3
(b)
C_2H_6
C_2H_4
-4000
-4500
-5000
-5500
V^{ext}/k_B (K)
0.00
0.25
0.50
0.75
1.00
Reaction Coordinate