Toward the Inverse Design of MOF Membranes for Efficient D_2/H_2 Separation by Combination of Physics-Based and Data-Driven Modeling

Musen Zhou, Anthony Vassallo and Jianzhong Wu*

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

ABSTRACT

Hydrogen isotopes are useful for scientific research, energy generation and medical treatment. However, their industrial production is expensive because conventional processes for separation of hydrogen isotopologues are mostly based on energy-intensive macroscopic procedures with extremely low separation efficiency. Metal-organic frameworks (MOFs) provide a promising route to D_2/H_2 separation by leveraging their well-defined chemistry and nanoporous structures. In this work, we report high-throughput screening of 12,723 experimentally synthesizable MOF membranes for D_2/H_2 separation by predicting gas adsorption and transport properties underpinning the separation efficiency. A membrane performance score is introduced to identify top ranked MOFs with the best selectivity and capacity. The extensive data generated from the physics-based modeling enables application of machine learning methods to predict desirable features of novel nanoporous materials for more efficient separation of hydrogen isotopes.

Keywords: Metal-organic frameworks, isotope separation, computational materials design, machine learning

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

1. Introduction

Isotopologues are chemical species with the same molecular structure but a different number of neutrons in certain atom types. Because of their unique properties, hydrogen isotopologues have been extensively used in scientific research, energy production and medical treatment, ranging from neutron scattering[1, 2], isotopic tracing[3-5] and nuclear fusion reaction[6], to medical imaging and cancer therapy[7]. As the only difference is the number of neutrons, hydrogen isotopologues have virtually identical chemical properties, making their separation at industrial scale exceedingly difficult. Conventional processes like cryogenic distillation and thermal diffusion suffer from low separation efficiency (e.g., the selectivity of D_2/H_2 is only about 1.5 at 24 K) and intensive energy consumption[8]. Whereas newly proposed technologies (e.g., atomic vapor laser isotope separation[9] and magnetically activated and guided isotope separation[10]) can achieve higher separation selectivity and consume less energy, their capacity for D_2/H_2 separation is severely limited thus hampering industrial applications.

Recently, metal-organic frameworks (MOFs) have been proposed for more efficient separation of hydrogen isotopes by utilizing quantum sieving effects. The mechanism was initially discovered by Beenakker and coworkers in 1995 and can be attributed to the disparity in the zero-point energies of isotopologues [11]. For gas separation with porous materials, the quantum sieving effect is most pronounced when the difference between the pore size and the diameter of gas molecules becomes comparable with the molecular de Broglie length. When the number of neutrons increases, hydrogen isotopologue has a smaller effective size and a larger binding energy. As a result, a porous material often adsorbs a heavier isotopologue more favorably than a lighter isotopologue and a smaller size makes the latter diffuse faster. By tuning the difference in the zero-point energies of hydrogen isotopologues, FitzGerald et al. demonstrated that quantum sieving

could lead to D_2/H_2 selectivity of 1.5 at a temperature as high as 150 K[12]. Because the good performance is mainly attributed to the difference between the binding energies of hydrogen isotopes, the procedure is also known as chemical-affinity quantum sieving (CAQS). A further improvement of the separation performance could be achieved by a combination different quantum sieving mechanisms[13].

MOFs have been recognized as one of the most promising nanoporous structures for gas storage and separation. Compared with alternative nanoporous materials like activated carbons or zeolites, MOFs have the advantages of tunable pore size, geometry, and local chemical composition[14-17]. Besides, MOFs are particularly promising for D_2/H_2 separation not only because of ultrahigh porosity and large specific surface area but also for the wide varieties of metal clusters and organic linkers that can be finely tuned to amplify the quantum sieving effects[18]. Previous experiments and molecular dynamics (MD) simulation have shown that MOF structures can be utilized to separate D_2/H_2 with a selectivity up to 41.4 at 20 K[13, 19-21]. In addition, a flexible MOF (MIL-53) has been investigated for separation of D_2/H_2 at 40 K; a selectivity of 13.6 was achieved by controlling the "breathing effect"[22]. Impressive selectivity for D_2/H_2 separation was also predicted for other sub-nanometer structures such as carbon nanotubes (CNTs)[23, 24], albeit they are less promising for industrial applications due to difficulties in materials synthesis[18].

Whereas quantum sieving effects are sensitive to the microscopic details of gas-pore interactions, such effects have not been systematically investigated and their influence on D_2/H_2 separation remains largely unknown. Previously[25, 26], we studied separation of isotopic methanes by high-throughput screening of both hypothetical and experimentally attainable MOF databases based on classical density functional theory (cDFT) calculations. In this work, we

demonstrate that similar procedures can be generalized to account for the quantum sieving effects by using the Feynman-Hibbs (FH) method[27]. Through the theoretical prediction of Henry's constants and self-diffusivity coefficients, the physics-based models allow us to evaluate 12,723 experimentally synthesizable MOF structures for D_2/H_2 separation and rank these materials by using the membrane performance score. The extensive properties data generated from physic-based modeling provide a sound basis for application of machine-leaning methods to identify important features of nanoporous materials that may achieve both high selectivity and separation capacity.

2. Methods

2.1 Molecular Model

We consider D_2/H_2 separation with various MOFs at 77 K, a cryogenic temperature commonly used for characterization of porous materials by nitrogen adsorption. The latest computation-ready, experimentally synthesizable MOF database (CoRE MOF 2019) is used for high-throughput screening and data generation[28]. For computational efficiency, all MOF structures are assumed to be rigid. While MOFs typically have good mechanical strength and structural rigidity, the structure flexibility may have significant effect on gas diffusion coefficient and thus separation efficiency. The same as the CoRE MOF 2014 library, the CoRE MOF 2019 structures do not contain solvent molecules or ionic species in the pore, which have been removed mimicking the experimental "activation" procedure. Different from those in hypothetical MOF databases, all MOF structures in the CoRE database have been experimentally synthesized, thus paving the way for experimental verification of theoretical predictions.

At the cryogenic condition, hydrogen molecules exhibit non-negligible quantum effects that cannot be captured with classical methods. In this work, we use the Lennard-Jones (LJ) model

to describe the classical component of intermolecular interactions. The quantum effects are accounted for by modifying the pair potential with the 4th-order approximation to the Feynman-Hibbs (FH) equation[27]

$$U(r)=U^{LJ}(r)+\frac{\beta\hbar^2}{24\mu}\left[U^{LJ}{}'(r)+\frac{2U^{LJ}{}''(r)}{r}\right]+\frac{\beta^2\hbar^4}{1152\mu^2}\left[U^{LJ}{}''''(r)+\frac{4U^{LJ}{}'''(r)}{r}+\frac{15U^{LJ}{}'(r)}{r^3}\right] \quad (1)$$

where *r* denotes the center-to-center distance, U^{LJ} represents the 12-6 LJ potential, $\beta = 1/k_B T$, k_B is the Boltzmann constant, T is the absolute temperature, $\hbar$ is the reduced Planck constant, $\mu = m_i m_j / (m_i + m_j)$ is the reduced mass for the interacting particles *i* and *j* with mass m_i and m_j, respectively. In Eq.(1), the number of primes represents the order of derivatives of the LJ potential with respect to the distance. The LJ parameters, $\sigma = 0.296$ nm and $\epsilon / k_B = 34.2$ K, are the same for H_2 and D_2[29]; the isotopologues are distinguished only by the quantum corrections related to their molecular weights. The universal force field (UFF) is used to represent all MOF atoms and a cutoff distance of 12.9 Å is applied to the LJ potential[30]. The Lorentz-Berthelot mixing rule is employed for describing interaction between different atoms.

Figure S1 illustrates the pair potential between hydrogen isotopologues according to the quantum-corrected LJ models (Supporting Information). While the 4th order approximation of the FH equation is computationally more expensive than the quadratic approximation commonly used in molecular simulations, addition of the higher order terms is important to fully capture the quantum effect in particular for hydrogen gases in a confined space[27]. As temperature falls, we may see more discrepancy among the intermolecular potentials for different hydrogen isotopologues because the quantum effect becomes more pronounced. Consistent with a previous report by Beenakker et al[11], a heavier hydrogen isotopologue has a smaller size but a higher attractive energy than a lighter isotopologue. The size difference can be attributed to the fact that

a heavier isotope has a narrower translational wave function thus a smaller de Broglie wavelength. On the other hand, the energy difference arises from the disparity in the zero-point energies of H_2 and D_2 molecules.

2.2 Ideal Adsorption Selectivity

We evaluate the separation efficiency for D_2/H_2 adsorption in MOF materials based on the ideal adsorption solution theory (IAST)[31]. For binary gas mixtures at low pressure with an equimolar composition in the bulk phase, IAST predicts a selectivity the same as that from Henry's law. The adsorption selectivity is thus measured by the ratio of Henry's constant for the heavier isotopologue relative to that for the lighter isotopologue

$$\alpha^{IM} = K_{h,2} / K_{h,1} \tag{2}$$

where superscript "IM" stands for ideal systems, i.e., gas adsorption at extremely low pressure such that interactions between gas molecules are negligible. For gas adsorption, Henry's constant is calculated from

$$K_h = \frac{1}{k_B T V} \int \exp\left[-\beta \varphi^{ext}\left(\mathbf{r}\right)\right] d\mathbf{r} \tag{3}$$

where **r** stands for the position of a gas molecule, V is the system volume, and $\varphi^{ext}\left(\mathbf{r}\right)$ represents the external potential for a gas molecule at **r** due to its interaction with the MOF atoms. With the quantum effects accounted for by the FH equation, we have different adsorbate-adsorbent interaction energies for H_2 and D_2 thereby different Henry's constants.

2.3 Ideal Membrane Selectivity

MOFs can be used for D_2/H_2 separation either as an adsorbent or as a porous membrane. In Henry's law region (*viz*., at low pressure), the membrane selectivity is determined by Henry's constant and the self-diffusivity coefficient of gas molecules at infinite dilution[17, 32, 33]:

$$k^{IM} = \frac{K_{h,2}}{K_{h,1}} \frac{D_{0,2}}{D_{0,1}} = \frac{P_2}{P_1} \tag{4}$$

where permeability P is defined as the product of Henry's constant, K_h, and the self-diffusivity coefficient at infinite dilution, D_0[33]. Unlike adsorption selectivity, the membrane selectivity depends on both the thermodynamic and the transport properties of individual gas compounds.

As reported in an earlier work[34], the diffusion coefficient at infinite dilution can be predicted by using the transition-state theory (TST):

$$D_0 = \frac{1}{2} \varpi a^2 \tag{5}$$

where a stands for the distance between the equilibrium positions of the gas molecule in two neighboring cages (the initial and final states of transmission), and ϖ is the gas hopping rate. The latter can be calculated from the potential energy along the diffusion path[35]

$$\varpi = \sqrt{\frac{k_B T}{2\pi m}} \frac{\exp\left[-\beta \varphi^{ext}\left(s^*\right)\right]}{\int_0^1 \exp\left[-\beta \varphi^{ext}\left(s\right)\right] ds} \tag{6}$$

where m denotes the molecular mass, the integral is performed along the reaction coordinate of gas hopping, and superscript * represents the transition state of a gas molecule hoping between neighboring cages. Similar to our previous work[26], the minimum energy path can be calculated by using a simplified string method. Among the minimum paths of all three possible directions, the one with highest diffusion coefficient is used as the diffusion path of the gas molecules because it is the most likely where molecular hopping would take place inside the MOF.

Although the above procedure is based on the classical theory, it has been shown that the quantum-corrected potential is able to capture not only the equilibrium quantum sieving effect but also the kinetic quantum sieving (KQS) consistent with experimental observations[36, 37]. In contrast to CAQS, KQS accounts for the quantum effects on gas transport arising from the

difference in the de Broglie wavelength. Because a lighter isotope has a smaller de Broglie wavelength, it experiences a higher energy barrier passing through the transition state than a heavier isotope thus a smaller diffusivity coefficient. The KQS is most significant for materials with ultra-small pores or at low temperature when the difference between the de Broglie wavelengths of isotopes is more pronounced.

It should be noted that, in addition to quantum sieving, quantum tunneling may contribute to gas transport at the cryogenic temperature. In principle, the quantum tunneling effect can be incorporated by adding a tunneling correction factor Q to the right side of Eq. (5),

$$Q = \frac{e^{\alpha}}{\gamma - \alpha}\left(\gamma e^{-\alpha} - \alpha e^{-\gamma}\right) \tag{7}$$

where $\alpha = E / k_B T$, $\gamma = 2a\pi^2 \left(2mE\right)^{1/2} / h$, E is the energy barrier for gas diffusion, $2a$ stands for the hopping distance, and h represents the Planck constant[38]. It has been shown that quantum tunneling is significant for H_2/D_2 separation only when the energy barrier is much larger than the difference between the zero-point energies[38]. Because such condition is rarely satisfied for D_2 and H_2 in MOF materials, we neglect the correction factor affiliated with the quantum tunneling effect for all calculations reported in this work.

2.4 Machine Learning Models

In addition to the physics-based models described above, we use data-based methods (a.k.a. machine learning methods) to identify the structural features of MOF membranes with good separation efficiency. Specifically, we have tested the performance of four types of machine learning models that may be able to reveal the desirable features of nanoporous materials for H_2/D_2 separation: *support-vector machine* (SVM), *random forest* (RF), *gradient-boosted trees* (GBT) and *deep neural network* (DNN). SVM is one of the most prevailing machine learning models for

classification; it projects the original data into a higher-dimensional space with the help of hyperplane construction[34, 35]. In comparison with alternative models, SVM is expected to achieve a better distinction of different MOFs within the original data because the kernel method is able to well recognize data patterns. RF is a machine-learning model consisting of a large amount of decision trees ("if-then" logic consequence) to extract useful information from the input data[39, 40]. The final result given by RF is the average vote of all decision trees in the model. In general, RF is able to avoid overfitting better than a single decision tree because the vote of biased decision trees would be averaged in the end. Similar to RF, GBT employs an ensemble of decision trees (weak learners) to achieve a strong learning power. Different from RF, decision trees in GBT are shallow trees with high bias and low variance instead of fully grown decision trees. Finally, DNN is well known for its excellent performance in predicting new features with a large number of input parameters. It is based on the idea of *artificial neural network* (ANN) that mimics the function of a human brain[41-43]. While ANN uses only a layer of neurons (transfer functions), DNN incorporates multiple layers of neurons thereby having a better interpretation power.

For all machine learning methods considered in this work, we optimize the parameters with the *k*-fold cross-validation procedure. It has a single parameter (*k*) referring to the number of groups that a given data sample is to be split into. In this work, we take a *k*-fold value of 17, the first prime number that is divisible to the number of MOFs in the training dataset.

3. Results and Discussion

3.1 Adsorption versus Permeation

Nanoporous materials can be used as an effective medium for gas separation based on either adsorption or permeation. The latter often achieves a higher selectivity because permeation in a membrane is related to both thermodynamic and transport properties. However, computational

studies of MOFs for gas separations are mostly focused on the adsorption behavior because calculation of transport properties renders additional challenges. In a previous work[34], we proposed an efficient theoretical procedure to calculate the diffusivity by using a simplified string method to calculate the minimum-energy path. The theoretical procedure allows us to calculate adsorption isotherms and diffusion coefficients with negligible computational cost in comparison with conventional simulation methods.

Figure 1 shows the selectivity for a large library of MOF materials (CoRE MOF 2019) versus their capacity for D_2/H_2 separation both by adsorption and by membrane permeation at 77 K. For separation of the hydrogen isotopologues by gas adsorption, the selectivity quickly declines as Henry's constant increases, implying that MOFs with high D_2/H_2 selectivity are limited by a low separation capacity. Approximately, the ideal adsorption selectivity falls exponentially with Henry's constant, and the latter provides a direct measure of adsorption capacity.

Table S2 presents the top 5 MOFs with the highest adsorption selectivity. Compared with the MOFs recommended for membrane separation to be discussed in the following section, the MOFs with the highest adsorption selectivity are often affiliated with small gas diffusion coefficients. As a result, the separation efficiency may be further compromised by the adsorption-desorption hysteresis as observed in recent experiments[18, 44]. Conversely, membrane separation exhibits no negative correlation between the selectivity and capacity.

As shown in Figure 1B, most MOFs are able to achieve an ideal membrane selectivity significantly higher than that suggested by the Robeson boundary (details are provided in Supporting Information). The high selectivity values indicate that a MOF membrane would have a much better performance than a conventional polymer membrane for separating D_2/H_2. Figure 1B also shows that most MOFs are able to achieve k_{IM} much higher than the classical limit of the

Knudsen diffusion (~ 0.7), albeit a few MOFs have k_{IM} around 0.7 because of the lack of sufficiently small pores for the KQS effect to take place.

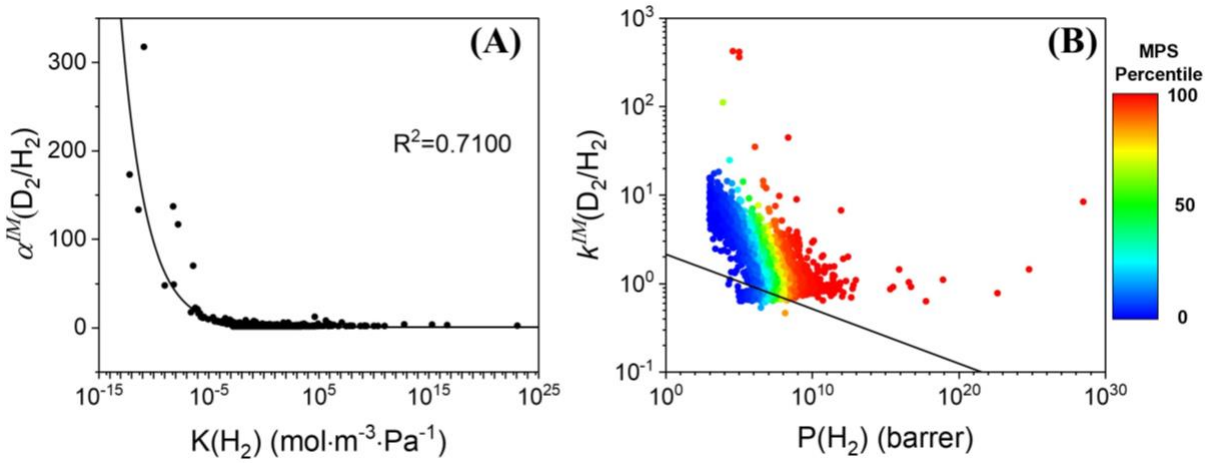


Figure 1. (A) Ideal selectivity for D_2/H_2 separation at 77 K by MOF adsorption versus Henry's constant for H_2. The points are calculated from Eq.(2) for 12,723 MOFs from the CoRE 2019 library, and the solid line is empirically fitted with $\alpha^{IM} = 139.7e^{-3694000K} + 1.33$. (B) Ideal membrane selectivity verse permeability of H_2 at 77 K. The points are predicted from Eq.(4) for the same CoRE MOF materials, and the solid line represents the Robeson boundary. The color in (B) denotes the percentile of Membrane Performance Score (MPS): the red and blue colors represent the highest and lowest MPS, respectively, and the green color represents the intermediate MPS.

Table 1 lists the properties of top 5 MOFs from the CoRE 2019 library with the highest membrane selectivity k_{IM}. In Supporting Information (Table S3), the selectivity and the structural features of top 5 MOFs identified according to k_{IM} are compared with other MOFs that have been reported in the literature for D_2/H_2 separation. The selectivity of MOFs identified in this work is higher than those previously reported nanoporous materials by almost one order of magnitude. As

discussed later, the top performing MOFs significantly enhance the KQS effect owing to their small pore limit diameters.

Table 1. Properties of top MOFs for D_2/H_2 separation at 77 K identified according to their ideal membrane selectivity k_{IM} ranking.

MOF	P(D_2) (barrer)	D_0(D_2) ($m^2 \cdot s^{-1}$)	K_h(D_2) ($mol \cdot m^{-3} \cdot Pa$)	k_{IM}(D_2/H_2)
ROQFUA07	1.689×10^{7}	5.47×10^{-18}	1.03×10^{9}	421.5
ROQNES05	4.506×10^{7}	6.75×10^{-18}	2.23×10^{9}	413.6
ROQFUA08	3.943×10^{7}	1.35×10^{-17}	9.77×10^{8}	360.6
ZOJWAY	9.232×10^{5}	2.54×10^{-15}	1.22×10^{5}	110.7
ECIVUH	1.057×10^{10}	8.93×10^{-12}	3.96×10^{5}	44.4

3.2 Structural features of highly selective MOF membranes

To understand molecular mechanisms underpinning the efficiency of D_2/H_2 selectivity, we have calculated the geometric features of the MOFs with top 5% ideal membrane selectivity, i.e., the largest cavity diameter (LCD), the pore limit diameter (PLD), and the void fraction. All geometry calculations are based on Zeo++ software with UFF[45]. As MOF membranes are able to achieve better performance than MOF adsorbents for D_2/H_2 separation in terms of both selectivity and capacity, the top ranked MOF adsorbents are not considered in our structural analysis.

Figure 2 shows the distributions of the pore limit diameter (A), the largest cavity diameter along the diffusion path (B), and the void fraction (C) of the top ranked MOF membranes. Distributions of the void volume and surface area of these materials are presented in Figure S2. In Figure 2A, we see a significant enhancement on the PLD distribution between 2 to 3 Å, suggesting that small pores are a desirable feature for D_2/H_2 separation. Intuitively, MOFs with PLD about 2-3 Å would better sieve D_2/H_2 at 77 K because the effective LJ diameter for hydrogen molecules is around 3 Å (see Table S1). As over 90% MOFs with top 5% ideal membrane selectivity have a PLD between 2 to 4 Å, it is important to replace the background distribution from the CoRE

database for MOFs with PLD of 2-4 Å to avoid misinterpretation of how the structural features correlate with the membrane performance. Figure S3 shows the distribution of unit cell volumes of the MOF membranes with top 5% ideal membrane selectivity compared with the same distribution but for all MOFs from the CoRE 2019 library and for MOFs with PLD in the range of 2-4 Å. While modifying the unit cell volume could not improve the performance for D_2/H_2 separation, one might erroneously conclude that the top MOF membranes would correlate with the unit cell volume if its distribution for the top 5% MOFs is directly compared with that for all MOFs in the CoRE 2019 library. Figure S3 shows an enhanced distribution of the unit cell volume at 0-2000 $Å_3$ simply because MOFs with PLD of 2-4 Å have a relatively small unit cell volume (0-2000 $Å_3$).

Figure 2 shows that the PLD and LCD distributions for the top ranked MOF membranes are noticeably different from the background distributions. While PLD plays a major role in determining the diffusion barrier, LCD affects gas permeation along the rest of diffusion coordinate. The smaller the LCD, the less favorable would be the minimum energy path for molecular hopping, leading to a smaller diffusion coefficient and gas permeability. Over 70% of MOFs with high ideal membrane selectivity have LCD in the range of 3-5 Å.

In addition to PLD and LCD, the void fraction is an important parameter to characterize the non-occupied space inside the MOF materials. It has been shown that a small change in the void fraction would affect the permeability of hydrogen molecules by orders of magnitude[46, 47]. However, permeability does not monotonically increase with the void fraction but depends also on the lattice type (e.g., fcc, bcc or simple cubic). Although a smaller void fraction would lead to a higher ratio of Henry's constants, it may also result in a smaller adsorption capacity. On the other hand, a larger void fraction may lead to larger pores but with diminishing quantum sieving effects

for D_2 and H_2 separation. Figure 2C shows that nearly half of MOFs with top 5% ideal membrane selectivity have an intermediate void fraction around 0.45.

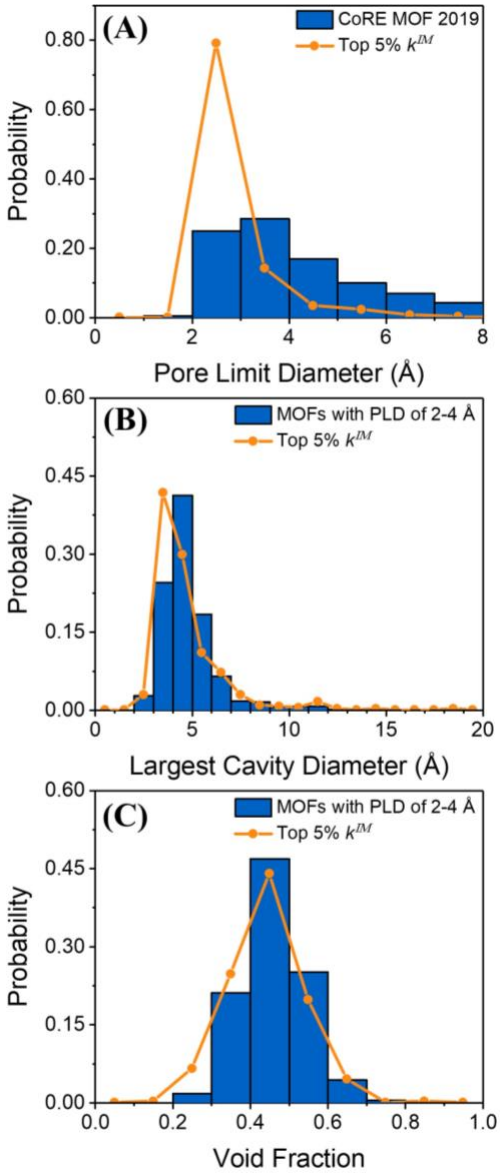


Figure 2. Distributions of the pore limit diameter (A), the largest cavity diameter along the diffusion path (B), and the void fraction (C) for MOFs in the CoRE 2019 library with the top 5% ideal membrane selectivity for D_2/H_2 separation at 77 K.

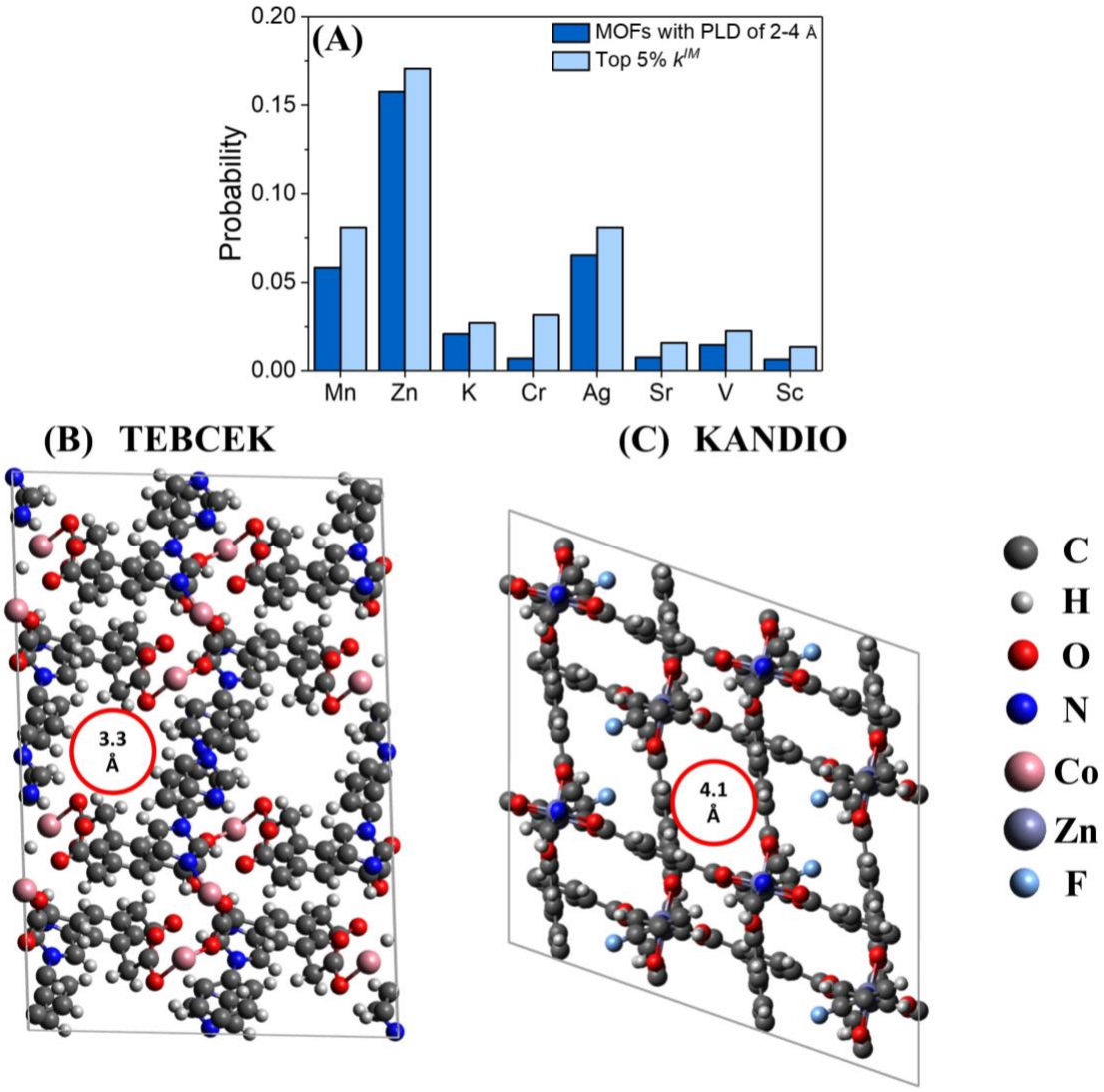


Figure 3. (A) Distributions of metal elements in the top 5% MOFs in terms of the ideal membrane selectivity and those with PLD in the range of 2-4 Å. All these elements exhibit a probability difference by at least 0.5% between the two groups of MOF structures. Example structures of MOFs containing metal elements with high (B) and low (C) energy parameters. The red circles denote the pore limit diameters.

In addition to the structural features, we have investigated the types of metal elements in the top ranked MOFs promising as a membrane for D_2/H_2 separation and their role in determining the separation efficiency. Because it would be computationally prohibitive to generate the force-

field parameters of open metal sites for high-throughput purpose, we assume in this work all metal elements are coordinatively saturated. Figure 3 shows the percentages of a few major metal elements in MOFs with top 5% ideal membrane selectivity and in those MOFs with PLD in the range of 2-4 Å. The results for all other metal elements are listed in Table S4. Most metal elements have only minor percentage difference between MOFs with top 5% ideal membrane selectivity and MOFs with PLD in the range of 2-4 Å.

Table 2. The LJ parameters for major metal elements in MOFs from the CoRE 2019 library with top 5% ideal membrane selectivity (elements with bolded fonts are considered as metal sites with high energy parameter)

Element	ε/k_B (K)	σ (Å)	Element	ε/k_B (K)	σ (Å)
Mn	6.54	2.64	Ag	18.11	2.80
Zn	**62.40**	**2.46**	**Sr**	**118.26**	**3.24**
K	17.61	3.40	V	8.05	2.80
Cr	7.55	2.69	Sc	9.56	2.94

The metal elements shown in Figure 3A have a probability in MOFs with top 5% ideal membrane selectivity at least 0.5% higher than that in MOFs with PLD in the range of 2-4 Å. Table 2 presents the LJ parameters for all these elements. Surprisingly, only a few metal elements have the LJ energy parameters higher than those for hydrogen molecules. As the Lorentz-Berthelot mixing rule is used to predict attraction between different species, metal elements with higher energy parameters would be more favored for D_2/H_2 separation. In other words, a larger energy parameter for the adsorbent is preferred because the small difference between the chemical species to be separated would be magnified. As each MOF is consisted of organic linkers and metal nodes, a.k.a. secondary building units (SBUs), quantum sieving can still be achieved for metal sites with lower energy parameter by pairing it with linkers of appropriate choice. For metal sites with a

lower energy parameter, a smaller linker would then be needed to make confinement small enough to distinguish H_2/D_2 and enable effective quantum sieving.

Figures 3B and 3C present representative structures of MOF membranes containing metal sites with high and low energy parameters, respectively. TEBCEK contains cobalt atoms ($\epsilon / k_B = 7.05$ K) while KANDIO contains zinc atoms ($\epsilon / k_B = 62.40$ K). While these two structures have similar ideal membrane selectivity, KANDIO has larger pore characteristics (LCD = 6.0 Å and PLD = 4.1 Å) in comparison to TEBCEK (LCD = 3.3 Å and PLD = 2.6 Å). In order to have equivalent membrane selectivity (KANDIO vs. TEBCEK, 5.7 vs. 4.6), a smaller pore size is required for MOFs containing metal sites with a lower binding energy such that their impacts on gas diffusivity and adsorption amount can be compensated with each other.

3.3 Membrane Performance Score

The adsorbent performance score has been widely used to evaluate porous materials for gas separation because it takes into account both selectivity and separation capacity. However, a similar metric was missing for membrane separations. Here we define the *Membrane Performance Score* (MPS) according to the membrane selectivity and permeability:

$$MPS = S_{fast/slow} \Re P_{fast} \tag{8}$$

where $S_{fast/slow}$ is the membrane selectivity of the fast diffusing species over the slow diffusing species, P_{fast} is the permeability of the fast specie. As shown in Figure 1B, a high MPS can be achieved by either a high selectivity or a high permeability. For separation of H_2 and D_2, most MOFs with high MPS can be attributed to high permeability but intermediate selectivity.

Table 3. Top 5 MOFs for D_2/H_2 separation at 77 K according to the membrane performance score

MOF	P(D_2) (barrer)	D_0(D_2) ($m^2 \cdot s^{-1}$)	k_{IM}(D_2/H_2)	MPS (barrer)
RUBLEH	2.94×10^{29}	3.97×10^{-10}	8.3	2.44×10^{30}
ERANAO	9.96×10^{24}	2.40×10^{-8}	1.42	1.42×10^{25}

YEGKIG	3.71×10^{22}	1.78×10^{-9}	0.77	6.30×10^{22}
XOPVOO	9.93×10^{18}	2.11×10^{-8}	1.09	1.08×10^{19}
FEKBED	3.94×10^{17}	5.88×10^{-12}	0.62	1.02×10^{18}

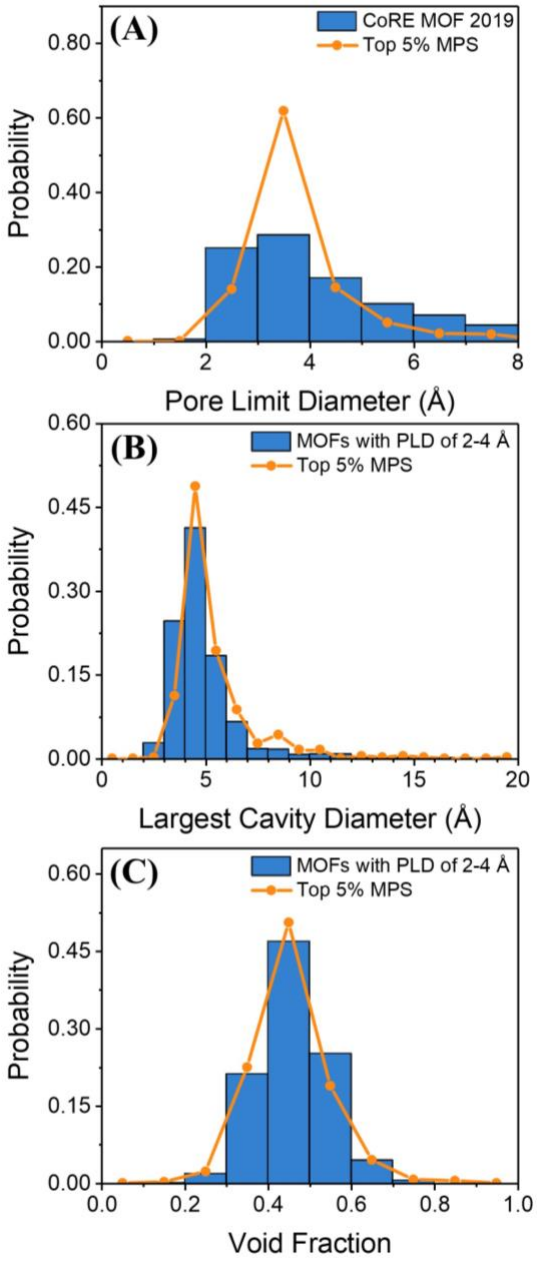


Figure 4. Distributions of the pore limit diameter (A), the largest cavity diameter along the diffusion path (B), and the void fraction (C) for MOFs with top 5% MPS for D_2/H_2 separation.

Table 3 lists the theoretical values of permeability, diffusivity coefficient at infinite dilution, and membrane selectivity for top 5 MOFs with the highest MPS values. Figure 4 and Figure S4 provide an analysis of the structural features of the top 5% MOFs according to the MPS ranking. Compared with MOFs with top 5% ideal membrane selectivity, MOFs with top 5% MPS have enhanced distributions for both PLD and LCD at a larger pore diameter, and the void fraction distribution shifts to the direction of less confinement. For D_2/H_2 separation, it is much easier to have a high permeability rather than a high membrane selectivity. Therefore, MOFs with top 5% MPS have less confined structures in order to achieve high permeability. Indeed, all MOFs listed in Table 3 have an extremely high permeability but with only moderate membrane selectivity.

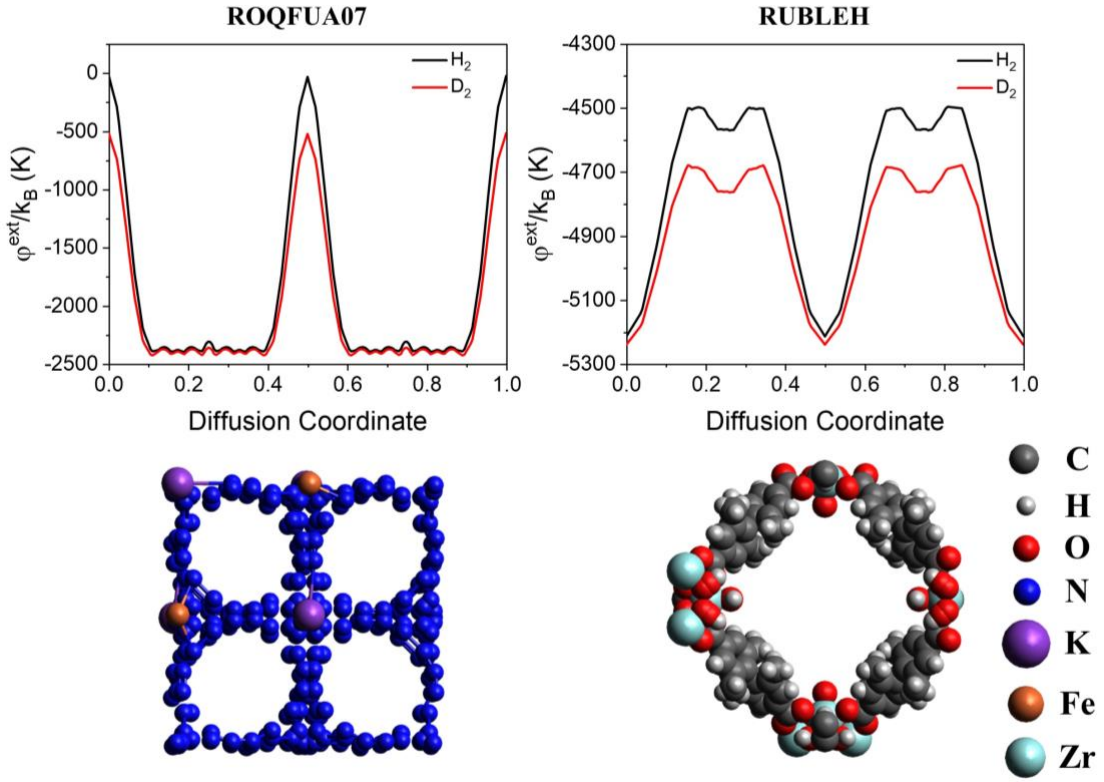


Figure 5. The energy landscape along the diffusion coordinate and the structures of MOFs with the highest ideal membrane selectivity (left panel - ROQFUA07) and those with the highest membrane performance score (right panel - RUBLEH).

Figure 5 shows the energy landscapes and the structures of MOFs with the highest ideal membrane selectivity (ROQFUA07) and with the highest membrane performance score (RUBLEH). We may identify significant differences in the energy landscapes for ROQFUA07 and RUBLEH. As discussed above, the highest ideal membrane selectivity is affiliated with an energy barrier that distinguishes hydrogen isotopologues through quantum sieving. However, the highest membrane performance score is attributed to high permeability and intermediate membrane selectivity. Not only is the difference between the energy barriers for H_2 and D_2 diffusion in ROQFUA07 much higher than those in RUBLEH, but the energy landscape along the minimum energy path differs greatly between these two MOF materials as well. RUBLEH is strongly attractive to both hydrogen isotopologues along the minimum energy path, leading to a high permeability and a high membrane performance score. By contrast, ROQFUA07 imposes little attraction at the diffusion barrier, implying that the process is dominated by repulsive interactions. The difference in the energy landscape results in the permeability of hydrogen isotopologues in RUBLEH much higher than that in ROQFUA07.

The structural features of MOF with the highest ideal membrane selectivity (ROQFUA07) are similar to those recommended by Nguyen et al.[36] In order to achieve good D_2/H_2 separation, PLD (2.587 Å) should be significantly smaller than σ_{H_2} while LCD (4.936 Å) should be slightly larger than σ_{H_2}. Among the 10 MOFs with the highest k_{IM} and MPS, ZOJWAY and FEKBED have relatively large PLDs compared to those for other top MOF candidates. Figure S5 shows the energy landscape along the diffusion path for these materials. ZOJWAY achieves an excellent selectivity through the extremely dense atomic coordination in its crystal structure, which leads to a large difference in the D_2/H_2 diffusion barriers. Conversely, H_2 has a smaller energy barrier in FEKBED and thus diffuses faster than D_2. For MOFs with such large pores, the KQS effect

becomes insignificant because H_2 diffusion in large pores does not suffer from its larger molecular size while D_2 is slowed down by stronger attraction energy and heavier molecular weight.

3.4 Promising Features of MOF Membranes Predicted by Machine Learning

While physics-based modelling allows us to analyze the materials behavior based on atomistic details, machine learning provides a complementary route to discovering new materials that satisfy predefined specific practical needs. A first step toward the so-called inverse design is to identify patterns or promising features from a large amount of data relating materials performance to the atomic constituents.

As discussed above, structural features, such as PLD, LCD and void fraction, are intrinsically related to the performance of MOF membranes for D_2/H_2 separation. In principle, these properties may be directly used as input parameters (*viz*., descriptors) for machine learning. Although the general trend can be captured by regression analysis, we found that these structural features are oversimplified and insufficient to establish quantitative correlations between the MOF structures and the physical properties underlying their performance for D_2/H_2 separation (shown in Figure S6). Understandably, the chemical effects including the metal elements and the kinetic effects such as overlapping of neighboring pores would have a major influence on the separation selectivity. However, even with additional descriptors reflecting the metal elements, a quantitative correlation is still beyond the capacity of regression analysis. Therefore, in the following, we use classification instead of regression methods for data-based modelling.

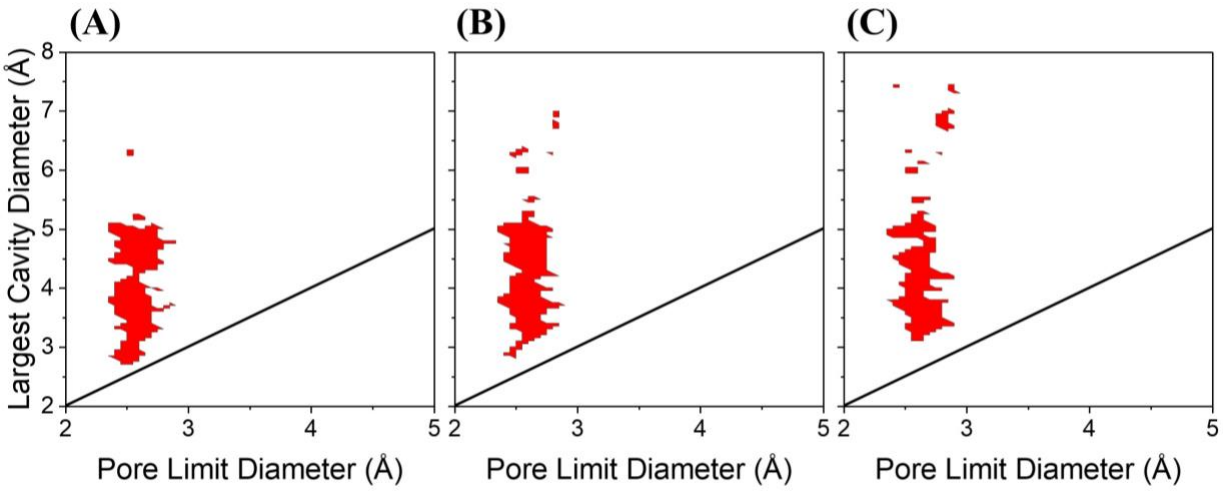


Figure 6. Structural features predicted by machine learning models for MOFs with top 10% ideal membrane selectivity (red color). Here the void fraction is fixed at (A) 0.4, (B) 0.45 and (C) 0.5. The black lines show conditions where PLD equals to LCD. Because PLD must be larger than LCD, the area below the black line is physically impossible.

We have analyzed the structural features of MOFs using *support-vector machine* (SVM), *random forest* (RF), *gradient-boosted trees* (GBT), and *deep neural network* (DNN). Table S5 compares the accuracy of different machine learning models. All four classification models are able to recognize MOFs with top 10% ideal membrane selectivity. With additional information on metal elements, DNN shows most significant improvement in accuracy. On the other hand, RF, GBT and SVM methods provide satisfactory classification with only structural features, e.g., in terms of the PLD, LCD and void fraction of each MOF structure. With the kernel method projecting data to a higher dimensional space, SVM is able to distinguish small differences between MOFs better than RF and GBT, which employ many decision trees of the same dimension. Surprisingly, RF, GBT and SVM methods show no improvement of the classification accuracy with the extra information for metal elements as the input. Because high-quality classification can be readily achieved with SVM, RF or GBT using the structural features as the input, additional

information on metal elements might be considered merely as a noise. Such information is too discrete and scarce and becomes detrimental to recognize the data patterns with SVM, RF or GBT[48].

Among the four classification methods, SVM has the highest accuracy and it is thus used to predict structural features for MOFs with the best membrane selectivity. The classification accuracy, which is defined as the percentage of MOFs correctly predicted by the four machine learning methods for the entire database, is presented in Supporting Information (Table S5). Figure 6 shows the predicted results. Here the red color denotes PLD and LCD values predicted by SVM for MOFs with top 10% ideal membrane selectivity at three different void fractions. The optimum PLD is approximately the same as the LJ diameter of the hydrogen molecules. In other words, MOFs with PLD comparable to the molecular size would yield a large difference in the energy barriers in H_2 or D_2 diffusion thus promoting separation. According to SVM, LCD should be slightly larger than the molecular size of isotopic hydrogen in order to utilize the KQS effect for D_2/H_2 separation, and the range of LCD shifts to larger diameters when the void fraction increases. While MOFs with PLD comparable to the molecular size of hydrogen would impose a nearly repulsive interaction, those with LCD slightly larger than the molecular size would be able to provide more attraction along the minimum energy path thereby increasing the molecular hopping rate. Because a larger LCD would accommodate a larger void fraction, the specific range of LCD depends on the void fraction of the MOFs to be designed to achieve optimal membrane performance.

4. Conclusions

In this work, we used physics-based models to evaluate the performance of 12,723 MOFs from the CoRE 2019 database for D_2/H_2 separation. The selectivity and separation capacity were

calculated for each MOF when it is used either as adsorbent or membrane material. We find that excellent D_2/H_2 selectivity can be achieved through gas adsorption but the MOF performance is compromised by low separation capacity. By contrast, MOFs can also be used as a membrane material for D_2/H_2 separation with a good balance of selectivity and capacity. Even at relatively high temperature (77 K), the D_2/H_2 selectivity for best MOF membranes identified in this work (Table 1) is almost one order of magnitude higher than those previously reported in the literature. Because all MOF structures in the CoRE database have been experimentally synthesized, our theoretical results are directly testable with experimental measurements.

An analysis of the structural features and the metal compositions of promising MOF membranes indicates that high D_2/H_2 selectivity can be achieved when the pore limit diameter (PLD) is comparable to the Lennard-Jones (LJ) diameter of hydrogen molecules. In particular, PLD plays an important role in determining the diffusion barrier, which is closely affiliated with the kinetic quantum sieving (KQS) effects. When the gas-MOF interaction is strongly affected by dense atom coordination in the crystal structure or unsaturated metal sites important in chemical affinity quantum sieving, the separation performance will also depend on other structural features such as the types of unsaturated metal sites, the largest cavity diameter (LCD) and the void fraction. For practical applications, we introduced the membrane performance score to evaluate the overall performance of the MOF membranes in terms of both selectivity and permeability. While MOFs with high membrane selectivity are characterized with extremely small pores, those with high membrane performance scores are less confined in order to achieve high permeability. In addition to PLD, gas permeation in MOFs depends on the LCD and the metal composition. The smaller the LCD, the less favorable would be the minimum energy path for molecular hopping, leading to a smaller diffusion coefficient and gas permeability. For MOFs with different metal compositions,

metal elements with a lower binding energy would require a smaller secondary building unit in order to achieve the same KQS effect.

Combination of physics-based models for high-throughput screening and data-based modelling for identification of useful geometric features facilitates the inverse design of MOFs for better separation performance. With the extensive data generated from physics-based modelling, we have identified useful features of MOF materials by exploring four different machine learning models. Although none of these models gives a quantitative correlation of membrane selectivity or the membrane performance scores, satisfactory results can be obtained by using support-vector machine (SVM) to reproduce the structural features of promising MOF membranes. For effective D_2/H_2 separation, SVM predicts that the KQS effects are most significant when the PLD of MOFs is comparable to, while LCD is slightly larger than, the molecular dimeter of isotopic hydrogens. The specific range of LCD depends on the void fraction of the MOF materials. Those structural features predicted by integrating physics-based modeling with machine learning provide useful insights into the rational design of new MOF structures for more efficient D_2/H_2 separation.

Acknowledgement

This material is based upon work supported by the National Science Foundation Harnessing the Data Revolution Big Idea under Grant No. NSF 1940118.

Reference

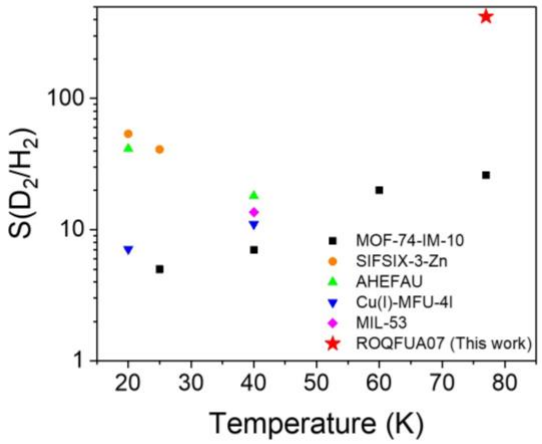
[1] G. Zaccai, How soft is a protein? A protein dynamics force constant measured by neutron scattering, Science, 288 (2000) 1604-1607.
[2] A. Machida, H. Saitoh, H. Sugimoto, T. Hattori, A. Sano-Furukawa, N. Endo, Y. Katayama, R. Iizuka, T. Sato, M. Matsuo, S. Orimo, K. Aoki, Site occupancy of interstitial deuterium atoms in face-centred cubic iron, Nat Commun, 5 (2014) 5063.
[3] I.V. Stiopkin, C. Weeraman, P.A. Pieniazek, F.Y. Shalhout, J.L. Skinner, A.V. Benderskii, Hydrogen bonding at the water surface revealed by isotopic dilution spectroscopy, Nature, 474 (2011) 192-195.

[4] F. Keppler, J.T. Hamilton, W.C. McRoberts, I. Vigano, M. Brass, T. Rockmann, Methoxyl groups of plant pectin as a precursor of atmospheric methane: evidence from deuterium labelling studies, New Phytol, 178 (2008) 808-814.
[5] P.P. Povinec, H. Bokuniewicz, W.C. Burnett, J. Cable, M. Charette, J.F. Comanducci, E.A. Kontar, W.S. Moore, J.A. Oberdorfer, J. de Oliveira, R. Peterson, T. Stieglitz, M. Taniguchi, Isotope tracing of submarine groundwater discharge offshore Ubatuba, Brazil: results of the IAEA-UNESCO SGD project, J Environ Radioact, 99 (2008) 1596-1610.
[6] P.C. Souers, Hydrogen properties for fusion energy, University of California Press, Berkeley, 1986.
[7] K. Sanderson, Big interest in heavy drugs, Nature, 458 (2009) 269.
[8] H.K. Rae, Selecting Heavy Water Processes, in: Separation of Hydrogen Isotopes, 1978, pp. 1-26.
[9] P.A. Bokhan, V.V. Buchanov, N.V. Fateev, M.M. Kalugin, M.A. Kazaryan, A.M. Prokhorov, D.E. Zakrevskĭi, Laser Isotope Separation in Atomic Vapor, 2006.
[10] T.R. Mazur, B. Klappauf, M.G. Raizen, Demonstration of magnetically activated and guided isotope separation, Nature Physics, 10 (2014) 601-605.
[11] J.J.M. Beenakker, V.D. Borman, S.Y. Krylov, Molecular-Transport in Subnanometer Pores - Zero-Point Energy, Reduced Dimensionality and Quantum Sieving, Chem Phys Lett, 232 (1995) 379-382.
[12] S.A. FitzGerald, C.J. Pierce, J.L. Rowsell, E.D. Bloch, J.A. Mason, Highly selective quantum sieving of D2 from H2 by a metal-organic framework as determined by gas manometry and infrared spectroscopy, J Am Chem Soc, 135 (2013) 9458-9464.
[13] J.Y. Kim, R. Balderas-Xicohtencatl, L. Zhang, S.G. Kang, M. Hirscher, H. Oh, H.R. Moon, Exploiting Diffusion Barrier and Chemical Affinity of Metal-Organic Frameworks for Efficient Hydrogen Isotope Separation, J Am Chem Soc, 139 (2017) 15135-15141.
[14] J.R. Li, R.J. Kuppler, H.C. Zhou, Selective gas adsorption and separation in metal-organic frameworks, Chemical Society Reviews, 38 (2009) 1477-1504.
[15] R.B. Getman, Y.S. Bae, C.E. Wilmer, R.Q. Snurr, Review and analysis of molecular simulations of methane, hydrogen, and acetylene storage in metal-organic frameworks, Chemical Reviews (Washington, D. C.), 112 (2012) 703-723.
[16] C.E. Wilmer, M. Leaf, C.Y. Lee, O.K. Farha, B.G. Hauser, J.T. Hupp, R.Q. Snurr, Large-scale screening of hypothetical metal–organic frameworks, Nature Chemistry, 4 (2011) 83.
[17] S. Keskin, J.C. Liu, J.K. Johnson, D.S. Sholl, Atomically detailed models of gas mixture diffusion through CuBTC membranes, Micropor Mesopor Mat, 125 (2009) 101-106.
[18] J.Y. Kim, H. Oh, H.R. Moon, Hydrogen Isotope Separation in Confined Nanospaces: Carbons, Zeolites, Metal-Organic Frameworks, and Covalent Organic Frameworks, Adv Mater, 31 (2019) e1805293.
[19] G. Garberoglio, Quantum sieving in organic frameworks, Chem Phys Lett, 467 (2009) 270-275.
[20] H. Oh, K.S. Park, S.B. Kalidindi, R.A. Fischer, M. Hirscher, Quantum cryo-sieving for hydrogen isotope separation in microporous frameworks: an experimental study on the correlation between effective quantum sieving and pore size, J Mater Chem A, 1 (2013) 3244-3248.
[21] D. Cao, H. Huang, Y. Lan, X. Chen, Q. Yang, D. Liu, Y. Gong, C. Xiao, C. Zhong, S. Peng, Ultrahigh effective H2/D2 separation in an ultramicroporous metal–organic framework material through quantum sieving, J Mater Chem A, 6 (2018) 19954-19959.

[22] J.Y. Kim, L. Zhang, R. Balderas-Xicohtencatl, J. Park, M. Hirscher, H.R. Moon, H. Oh, Selective Hydrogen Isotope Separation via Breathing Transition in MIL-53(Al), J Am Chem Soc, 139 (2017) 17743-17746.
[23] S.R. Challa, D.S. Sholl, J.K. Johnson, Light isotope separation in carbon nanotubes through quantum molecular sieving, Phys Rev B, 63 (2001).
[24] G. Garberoglio, J.K. Johnson, Hydrogen isotope separation in carbon nanotubes: calculation of coupled rotational and translational States at high densities, ACS Nano, 4 (2010) 1703-1715.
[25] Y. Tian, W.Y. Fei, J.Z. Wu, Separation of Carbon Isotopes in Methane with Nanoporous Materials, Ind Eng Chem Res, 57 (2018) 5151-5160.
[26] M. Zhou, Y. Tian, W. Fei, J. Wu, Fractionation of Isotopic Methanes with Metal–Organic Frameworks, The Journal of Physical Chemistry C, 123 (2019) 7397-7407.
[27] A.V. Kumar, S.K. Bhatia, Quantum effect induced reverse kinetic molecular sieving in microporous materials, Phys Rev Lett, 95 (2005) 245901.
[28] Y.G. Chung, E. Haldoupis, B.J. Bucior, M. Haranczyk, L. S., V.K. D., L. S., M. Milisavljevic, H. Zhang, J. Camp, B. Slater, J.I. Siepmann, D.S. Sholl, R.Q. Snurr, Computation-Ready Experimental Metal-Organic Framework (CoRE MOF) 2019 Dataset, in, 2019.
[29] V. Buch, Path integral simulations of mixedpara - D2andortho - D2clusters: The orientational effects, The Journal of Chemical Physics, 100 (1994) 7610-7629.
[30] A.K. Rappe, C.J. Casewit, K.S. Colwell, W.A. Goddard, W.M. Skiff, Uff, a Full Periodic-Table Force-Field for Molecular Mechanics and Molecular-Dynamics Simulations, Journal of the American Chemical Society, 114 (1992) 10024-10035.
[31] A.L. Myers, J.M. Prausnitz, Thermodynamics of mixed-gas adsorption, AIChE Journal, 11 (1965) 121-127.
[32] S. Keskin, D.S. Sholl, Efficient methods for screening of metal organic framework membranes for gas separations using atomically detailed models, Langmuir, 25 (2009) 11786-11795.
[33] E. Haldoupis, S. Nair, D.S. Sholl, Efficient calculation of diffusion limitations in metal organic framework materials: a tool for identifying materials for kinetic separations, Journal of the American Chemical Society, 132 (2010) 7528-7539.
[34] Y. Tian, X. Xu, J. Wu, Thermodynamic Route to Efficient Prediction of Gas Diffusivity in Nanoporous Materials, Langmuir, 33 (2017) 11797-11803.
[35] D. Frenkel, B. Smit, Understanding Molecular Simulation: from Algorithms to Applications, 2nd ed., Academic Press, San Diego, 2002.
[36] T.X. Nguyen, H. Jobic, S.K. Bhatia, Microscopic observation of kinetic molecular sieving of hydrogen isotopes in a nanoporous material, Phys Rev Lett, 105 (2010) 085901.
[37] M.J. Murphy, G.A. Voth, A.L.R. Bug, Classical and quantum transition state theory for the diffusion of helium in silica sodalite, J Phys Chem B, 101 (1997) 491-503.
[38] E.V. Anslyn, Modern physical organic chemistry, University Science/Sausalito CA, Sausalito CA, 2006.
[39] H. Tin Kam, The random subspace method for constructing decision forests, IEEE Transactions on Pattern Analysis and Machine Intelligence, 20 (1998) 832-844.
[40] L. Rokach, O. Maimon, Data Mining with Decision Trees, 2014.
[41] G.B. Goh, N.O. Hodas, A. Vishnu, Deep learning for computational chemistry, J Comput Chem, 38 (2017) 1291-1307.
[42] Z. Waszczyszyn, Fundamentals of Artificial Neural Networks, in, Springer Vienna, Vienna, 1999, pp. 1-51.

[43] G.G. Towell, J.W. Shavlik, Knowledge-based artificial neural networks, Artificial Intelligence, 70 (1994) 119-165.
[44] K.M. Thomas, Adsorption and desorption of hydrogen on metal-organic framework materials for storage applications: comparison with other nanoporous materials, Dalton Trans, (2009) 1487-1505.
[45] T.F. Willems, C. Rycroft, M. Kazi, J.C. Meza, M. Haranczyk, Algorithms and tools for high-throughput geometry-based analysis of crystalline porous materials, Micropor Mesopor Mat, 149 (2012) 134-141.
[46] L. Perez-Mas, A. Martin-Molina, M. Quesada-Perez, A. Moncho-Jorda, Maximizing the absorption of small cosolutes inside neutral hydrogels: steric exclusion versus hydrophobic adhesion, Phys Chem Chem Phys, 20 (2018) 2814-2825.
[47] W.K. Kim, M. Kanduč, R. Roa, J. Dzubiella, Tuning the Permeability of Dense Membranes by Shaping Nanoscale Potentials, Physical Review Letters, 122 (2019) 10.1103/PhysRevLett.1122.108001.
[48] T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning, Springer Nature Switzerland AG. , 2009.

TOC graph


100
10
1
S(D2/H2)
20
30
40
50
60
70
80
Temperature (K)
MOF-74-IM-10
SIFSIX-3-Zn
AHEFAU
Cu(I)-MFU-4l
MIL-53
ROQFUA07 (This work)