Check for updates

How to cite: *Angew. Chem. Int. Ed.* **2021**, *60*, 7828 – 7837
International Edition: doi.org/10.1002/anie.202015250
German Edition: doi.org/10.1002/ange.202015250

MOF Simulations

Effect of Metal–Organic Framework (MOF) Database Selection on the Assessment of Gas Storage and Separation Potentials of MOFs

*Hilal Daglar, Hasan Can Gulbalkan, Gokay Avci[+], Gokhan Onder Aksu[+], Omer Faruk Altundal[+], Cigdem Altintas, Ilknur Erucar, and Seda Keskin**

Abstract: *Development of computation-ready metal–organic framework databases (MOF DBs) has accelerated high-throughput computational screening (HTCS) of materials to identify the best candidates for gas storage and separation. These DBs were constructed using structural curations to make MOFs directly usable for molecular simulations, which caused the same MOF to be reported with different structural features in different DBs. We examined thousands of common materials of the two recently updated, very widely used MOF DBs to reveal how structural discrepancies affect simulated CH_4, H_2, CO_2 uptakes and CH_4/H_2 separation performances of MOFs. Results showed that DB selection has a significant effect on the calculated gas uptakes and ideal selectivities of materials at low pressure. A detailed analysis on the curated structures was provided to isolate the critical elements of MOFs determining the gas uptakes. Identification of the top-performing materials for gas separation was shown to strongly depend on the DB used in simulations.*

Introduction

Due to the development of novel design strategies in reticular chemistry,[1,2] there is a remarkable increase in the number of synthesized metal-organic frameworks (MOFs) uploaded into the Cambridge Structural Database (CSD).[3] Experimental identification of the top performing materials for gas adsorption and separation among thousands of MOFs is a challenge.[4] High-throughput computational screening (HTCS) studies based on Grand Canonical Monte Carlo (GCMC) simulations are performed to find out the most promising MOFs for gas storage (CO_2,[5] CH_4,[6] H_2[7]) and separation (CO_2/CH_4,[8] CO_2/H_2,[9] CO_2/N_2,[10] and CH_4/H_2[11]) in a time-efficient manner.[12–14] The only experimental input of GCMC simulations is the crystallographic information file (CIF) of the synthesized MOFs, which contains the atomic coordinates and the symmetry information for atoms. These CIFs generally contain residual solvent molecules and/or disordered atoms that should be removed/fixed before using the structure files in molecular simulations.

Computation-ready MOF databases (DBs) have applied a series of curations to MOF structures for direct usage of the CIFs in molecular simulations and significantly accelerated the HTCS of MOFs. Computation-ready, experimental MOFs (CoRE MOF 2014), consisting of 5109 structures, was the first publicly available experimental MOF DB.[15] This DB includes 3D MOFs with pore limiting diameter > 2.4 Å, which have been curated by checking disordered atoms, removing bound and unbound solvents, adding missing hydrogens into the frameworks, and editing manually when required. In 2017, the CSD non-disordered MOF subset was published consisting of 55547 non-disordered MOFs and it is continuously being updated along with the CSD while disordered structures are fixed and maintained.[16,17] This MOF subset consists of 1D, 2D, 3D non-disordered MOFs, and structures should be further curated using the Python script provided with the dataset to clean out the residual solvents.[16] Both DBs have recently undergone major updates. CoRE MOF 2014 is updated by fixing the previously disordered MOFs and 14142 solvent-free MOFs are reported in the CoRE MOF 2019.[18] For the CSD non-disordered MOF subset, several guidelines were introduced to make a MOF search such as checking for the polymeric bonds and introducing stringent requirements where each structure went through many automated tests to be classified as a non-disordered MOF.[17] Although these MOF datasets have been updated and improved both in terms of the number of MOFs and the accuracy of the curation procedures used for the refinement of structures, there are still improperly reported CIFs in DBs.

The ultimate aim of HTCS studies is to identify the top performing MOFs for a specific application, however, it is challenging to identify problematic MOFs in a HTCS study when thousands of materials are taken from DBs. Therefore, accurate representation of MOFs in computation-ready DBs is crucial. Chen et al.[19] investigated the misbonded atoms in

[*] H. Daglar, H. C. Gulbalkan, G. O. Aksu,[+] O. F. Altundal,[+] C. Altintas, S. Keskin
Department of Chemical and Biological Engineering, Koc University
Rumelifeneri Yolu, Sariyer, 34450, Istanbul (Turkey)
E-mail: skeskin@ku.edu.tr
G. Avci[+]
Department of Materials Science and Engineering, Koc University
Rumelifeneri Yolu, Sariyer, 34450, Istanbul (Turkey)
I. Erucar
Department of Natural and Mathematical Sciences, Faculty of Engineering, Ozyegin University
Cekmekoy, 34794, Istanbul (Turkey)
[+] These authors contributed equally to this work.
Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under: https://doi.org/10.1002/anie.202015250.

© 2021 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

MOFs retrieved from CoRE MOF 2019 and reported that 4901 MOFs have structural problems, such as isolated and overlapping atoms which may affect the simulation results. We recently focused on why the structures having the same refcode (a unique, six letter code to index MOFs) were differently reported in different computation-ready DBs and addressed the discrepancies in structural curations such as treatment of charge balancing ions (CBIs) and removal of solvents.[20] Velioglu and Keskin[21] examined the effect of structural curations on CO_2/CH_4, CO_2/N_2, and CO_2/H_2 mixture separation performances of a small number of MOFs (68) and showed that different procedures used for the removal of coordinated solvents and CBIs led to significant differences in simulated CO_2 uptakes of the same MOFs taken from different DBs.

In this study, we provide a detailed analysis of MOFs in the two updated, currently available DBs: all solvents removed subset of CoRE MOF 2019 (will be referred as CoRE DB) and CSD non-disordered MOF subset of November 2019, stripped from solvents (will be referred as CSDSS DB). First, all the common MOFs having the same refcodes in these computation-ready DBs were identified and chemical formulas of these curated common MOFs were compared to determine the ones reported differently in two DBs as a result of structural curations. We specifically focused on the question of how structural differences between common MOFs in two DBs affect the (i) simulated uptakes of CH_4, H_2, CO_2, (ii) calculated ideal selectivities, (iii) calculated adsorbent performance evaluation metrics for CH_4/H_2 mixture separation at vacuum swing adsorption (VSA) and pressure swing adsorption (PSA) conditions. We finally performed a detailed atomistic analysis on the chemical formula differences of MOFs which reveals that gas uptakes are significantly affected by the type of elements removed during curations. Overall, we addressed how the use of computation-ready MOFs in molecular simulations affects the assessment of gas storage and separation potentials of MOFs.

Results and Discussion

Effect of structural discrepancies on the simulated gas uptakes in MOFs

After analyzing 3543 common MOFs of CoRE and CSDSS DBs using the computational methodology shown in Scheme S1 of Supporting Information (SI), we concluded that 2434 MOFs have the same chemical formula when taken from the two DBs; the ones we refer as chemical formula matched-MOFs (CFM-MOFs), and 1109 MOFs have different chemical formulas; the ones we refer as chemical formula unmatched-MOFs (CFU-MOFs). We provided the list of CFM-MOFs and CFU-MOFs with their calculated physical properties as a separate file in the SI. To easily compare both the computed physical properties, such as pore limiting diameter (PLD), the largest cavity diameter (LCD), accessible surface area (S_{acc}), pore volume (PV), porosity (ϕ), and density (ρ), and the simulated performance metrics, such as ideal selectivity (S_{ideal}), mixture selectivity (S_{mix}), CH_4 working capacity (ΔN_{CH_4}), adsorbent performance score (APS), and percent regenerability (R%) of MOFs taken from CoRE and CSDSS DBs, we defined a term called property(or performance)$_{ratio}$ as follows:

$$\text{property(or performance)}_{ratio} = \frac{\text{physical property (or performance metric) of a MOF taken from CoRE DB}}{\text{physical property (or performance metric) of a MOF taken from CSDSS DB}} \quad (1)$$

This term demonstrates the calculated property of a MOF used from CoRE DB relative to that of the same MOF used from CSDSS DB. For example, if the ratio of gas uptake ($N_{i,ratio}$) for a MOF is greater (smaller) than 1, CoRE MOF DB has a higher (lower) computed gas uptake (N_i) than the CSDSS DB. We first focused on 2434 CFM-MOFs for which representative examples are given in Table S1 with their chemical formula, $N_{i,ratio}$, and $S_{ideal,ratio}$ with respect to the ratio of their structural properties (ϕ_{ratio}, PV_{ratio}, $S_{acc,ratio}$, PLD_{ratio}) and the ratio of heat of adsorption values of gases ($Q^0_{st,i,ratio}$) at infinite dilution. As Table S1 shows each CFM-MOF has the same refcodes, chemical formulas, and calculated structural properties in the two DBs. As a result, GCMC simulations gave the same gas uptakes (Figure S1). Therefore, utilization of the chemical formula together with the refcode successfully identifies the structures that are the same in CoRE and CSDSS DBs.

The ultimate aim of our work is to reveal how the simulation results of 1109 CFU-MOFs change depending on the DB used. In most HTCS studies for gas separation, molecular simulations are performed either at infinite dilution or for single-component gas adsorption at different pressures.[10,22–24] Single-component gas uptakes are also used to define the deliverable capacities for H_2 and CH_4 storage.[25,26] Therefore, we compared the two DBs based on the molecular simulation results for Henry's constants, K_{H_2}, K_{CH_4} computed at infinite dilution and N_{H_2}, N_{CH_4} computed at 1, 20, and 65 bar for CFU-MOFs in Figure 1. Simulations of MOFs taken from CoRE DB generally result in higher Henry's constants and uptakes for H_2 and CH_4 compared to MOFs taken from CSDSS DB at all conditions. This can be explained by the molecular origin of the structural differences due to the different curation methods of DBs: CSDSS DB only considers the removal of unbound and bound solvents remaining in the pores of the MOFs whereas CoRE DB adds the missing hydrogen atoms and treats the CBIs in addition to solvent removal. For example, retention of CBIs in CoRE MOFs depends on the existence of CBI formula within the chemical formula of the MOF reported in the CSD, while the solvent removal script of CSDSS does not have an additional step for treating CBIs. These different procedures lead to variances in chemical and physical properties of MOFs in two DBs which remarkably affect their simulated gas uptakes.

Figure 1 a shows that K_{CH_4} and K_{H_2} of MOFs significantly change depending on the DB used. As an example, the CoRE version of a MOF (refcode: IBABUL) has a K_{CH_4} which is almost 50 times larger than its CSDSS version due to the accidental removal of unbound sulfonyldibenzoic acid and

 © 2021 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH

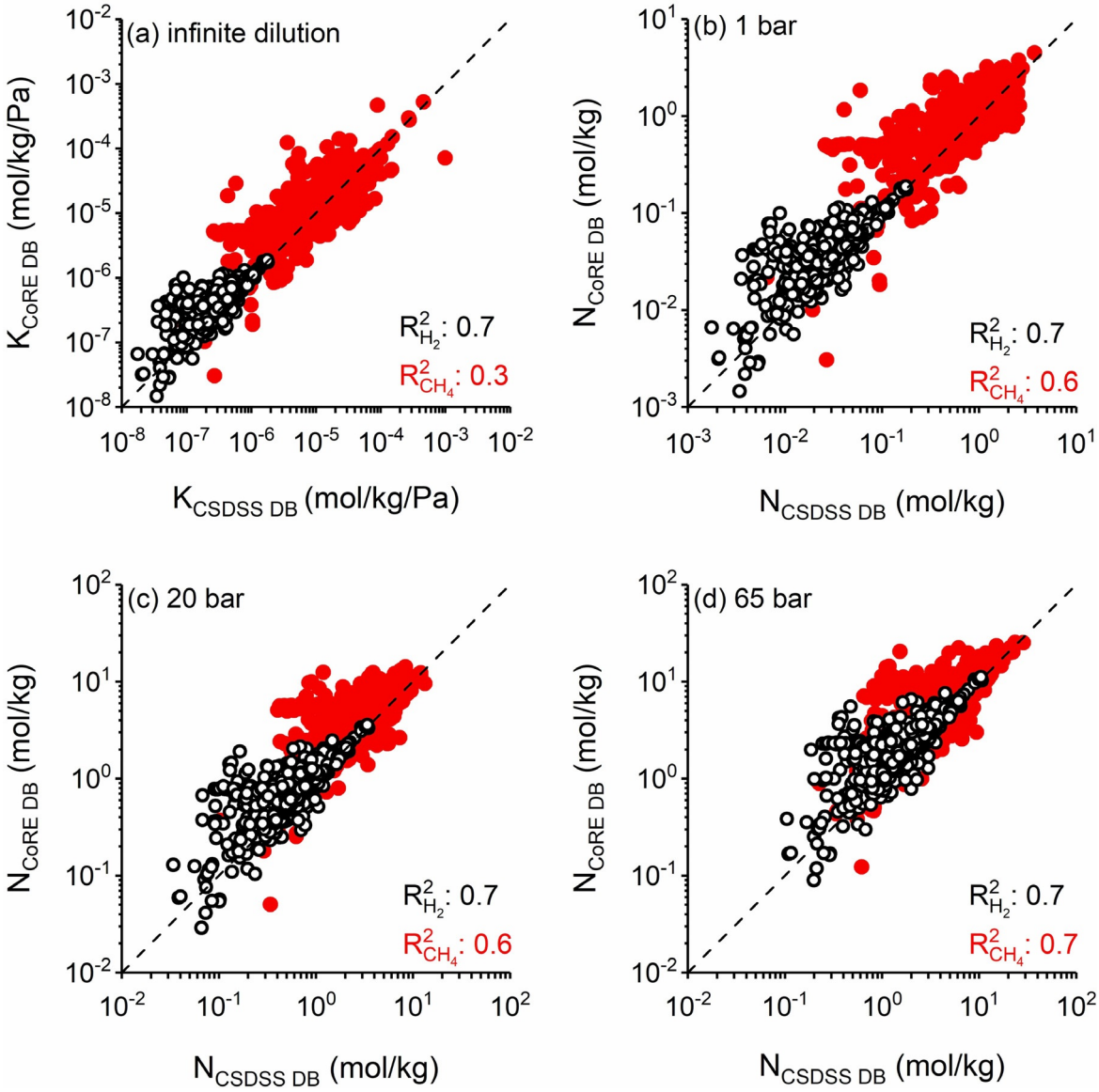


Figure 1. K_{H_2}, K_{CH_4} computed at a) infinite dilution, N_{H_2}, N_{CH_4} computed at b) 1 bar, c) 20 bar, and d) 65 bar for 1109 CFU-MOFs of two DBs. Red symbols represent CH_4, black symbols represent H_2 data.

hydronium ion by CoRE DB.[27] This is important because several HTCS studies use infinite dilution condition as the initial screening step to identify the best MOFs,[23,28] and our results show that the choice of DB plays a vital role at this condition. As pressure increases up to 65 bar, the deviation between N_{CH_4} of MOFs in two DBs becomes less pronounced (Figure 1b–d). At high pressures, gas-gas interactions also contribute to the gas uptakes in addition to the gas-MOF interactions, hence the effect of structural curations on the simulated gas uptakes decrease at high pressures. Overall, our results highlight that using either the CoRE or CSDSS DB in molecular simulations significantly affects the identification of the materials with high gas storage capacity.

To explore the differences of CFU-MOFs in detail, we divided the MOFs into four groups based on the trends observed in Figure 1. Ten representative CFU-MOFs from these groups are shown in Table S2 with their unit cell images, chemical formulas, ϕ_{ratio}, PV_{ratio}, $S_{acc,\ ratio}$, PLD_{ratio}, $Q^0_{st,i,ratio}$, $N_{i,\ ratio}$, and $S_{ideal,\ ratio}$. Group A shows the CFU-MOFs with $N_{i,\ ratio} < 0.9$, indicating that CSDSS MOFs are computed to have higher uptakes both for CH_4 and H_2 than CoRE MOFs. One of the interesting examples is PACZUQ. CSDSS version of PACZUQ has higher N_{CH_4} and N_{H_2} values at 1 bar (0.6 and 0.02 mol kg^{-1}, respectively) compared to the CoRE version (0.2 and 0.01 mol kg^{-1}, respectively). This is attributed to the increase in S_{acc} (586 m^2 g^{-1} for CoRE, 954 m^2 g^{-1} for CSDSS) and PV (0.31 cm^3 g^{-1} for CoRE, 0.38 cm^3 g^{-1} for CSDSS) due to the removal of bound water molecules by CSDSS DB.

Group B ($0.9 < N_{i,\ ratio} < 1.1$) represents CFU-MOFs having very similar gas uptakes. A representative MOF is FUTDII in which N_{CH_4} and N_{H_2} are similar for both DBs, although the Cl^- ion which should reside in the channels based on its experimental structure,[29] is accidentally removed by the CoRE DB. Group C shows CFU-MOFs having

 © 2021 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH

an $N_{i,\ ratio} > 1.1$, where CoRE MOFs are computed to have higher gas uptakes than CSDSS MOFs for both gases. QOZFUG is a representative for which N_{CH_4} and N_{H_2} in CoRE DB (0.9 and 0.04 mol kg^{-1} at 1 bar) are calculated to be higher than the ones in CSDSS DB (0.2 and 0.01 mol kg^{-1}). This difference is due to the removal of hydroxyl groups and H_2O molecules (reported without hydrogens) by CoRE DB. As a result, S_{acc} (PV) is computed as 1946 and 496 m^2 g^{-1} (0.7 and 0.4 cm^3 g^{-1}) for the CoRE and CSDSS versions, respectively.

Group D shows the CFU-MOFs with $N_{i,\ ratio} < 0.9$ for CH_4 and > 1.1 for H_2, or vice versa. For example, CSDSS version of PUZMAZ is calculated to have a higher N_{CH_4} (1.1 mol kg^{-1} at 1 bar) than the CoRE version (0.8 mol kg^{-1}). This is attributed to the removal of the guest diphenylamine in CoRE DB, leading to less favorable adsorption sites for CH_4, as supported by the lower $Q^0_{st,\ CH_4}$ in CoRE (10 kJ mol^{-1}) than that in CSDSS (19 kJ mol^{-1}). On the other hand, N_{H_2} is computed to be higher in the CoRE version (0.1 mol kg^{-1}) than the CSDSS version (0.03 mol kg^{-1}). Removal of the guest does not significantly affect the framework-H_2 interaction as $Q^0_{st,\ H_2}$ values of PUZMAZ in both DBs are low (< 5 kJ mol^{-1}), indicating that available space in CoRE version is significant for H_2 adsorption.

To better understand how structural differences in two DBs affect the structure-performance relations of CFU-MOFs, we show the gas uptakes with respect to PV_{ratio} and $S_{acc,\ ratio}$ in Figure 2 at four different conditions. As shown in Figure 2a,b, K_{CH_4} at infinite dilution and N_{CH_4} at 1 bar does not depend on PV_{ratio} and $S_{acc,\ ratio}$, but as pressure increases to 20 and 65 bar (Figure 2c,d), there is a strong correlation of PV_{ratio} and $S_{acc,\ ratio}$ with $N_{CH_4,\ ratio}$. The relation between Q^0_{st,CH_4} and $K_{CH_4,\ ratio}$ in Figure S2(a) shows that Q^0_{st,CH_4} is the main driving factor for CH_4 adsorption at infinite dilution. S_{acc} and PV play a more important role on gas uptakes at higher pressures.[30,31] Similar analyses for H_2 (Figures S2(b) and S3)

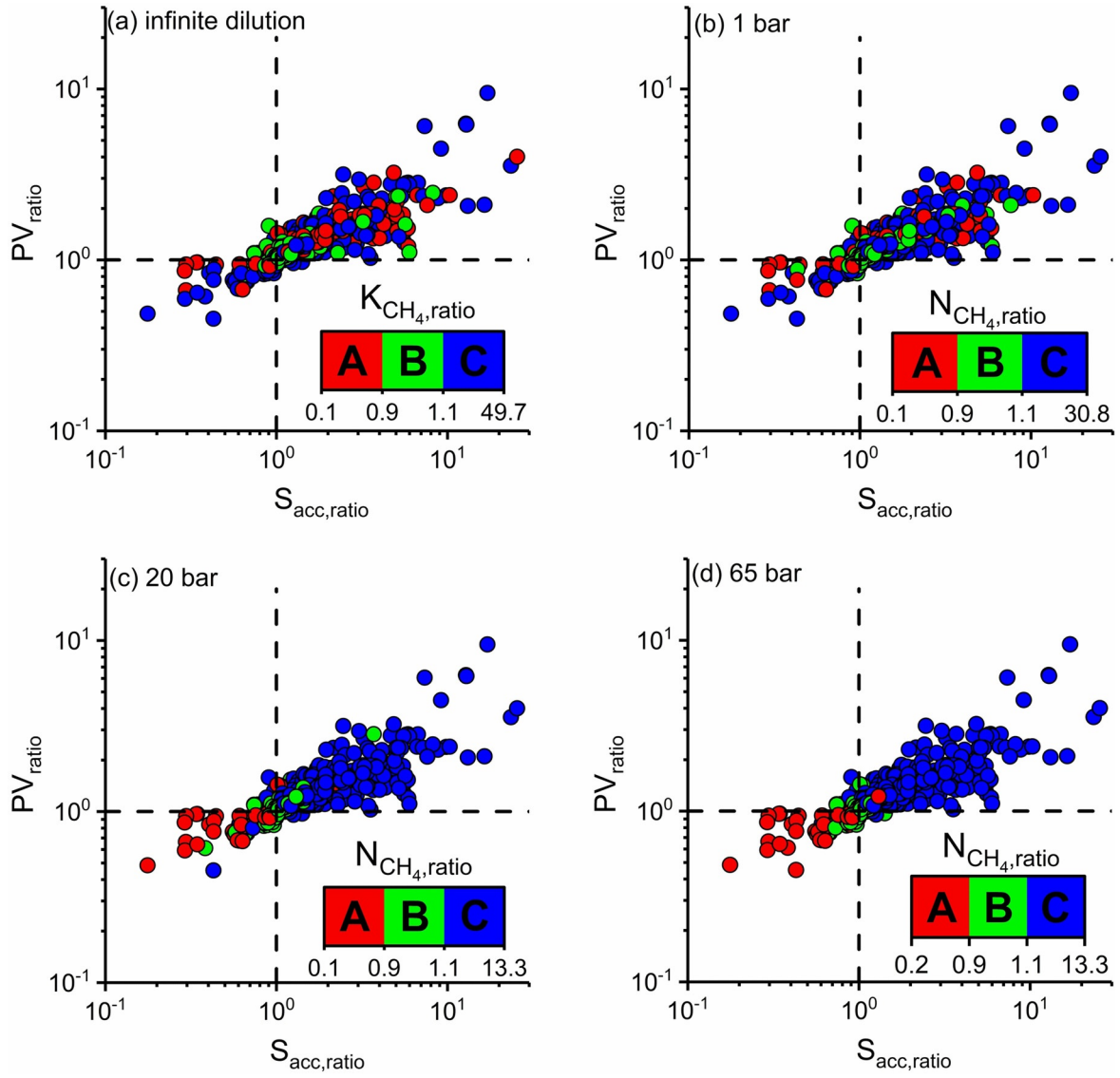


Figure 2. Relation between PV_{ratio} and $S_{acc,\ ratio}$ of 1109 CFU-MOFs. Data is color-coded according to $K_{CH_4,\ ratio}$ computed at a) infinite dilution, $N_{CH_4,\ ratio}$ computed at b) 1 bar, c) 20 bar, d) 65 bar. A, B and C on color code represents the groups of MOFs whose examples are shown in Table S2.

 © 2021 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH

shows that S_{acc} and PV are effective on H_2 adsorption even at low pressures and trends do not significantly change with increasing pressure, since H_2 does not strongly interact with the pore walls at room temperature.[32] These results highlight that $S_{acc,ratio}$ and PV_{ratio} of CFU-MOFs can be used to predict if the simulated CH_4 uptake of a MOF differs depending on DB used at high pressures.

Figure S4(a-b) shows the comparison of K_{CO_2} and N_{CO_2} at 1 bar for CFU-MOFs in two DBs when electrostatic interactions between MOF and CO_2 (EICM) are not considered. Similar to CH_4 and H_2, MOFs taken from CoRE DB generally have higher K_{CO_2} and N_{CO_2} compared to MOFs taken from CSDSS DB. For example, the uptake of a MOF taken from CoRE DB is 2 mol kg^{-1} whereas the CSDSS version of the same MOF has an uptake of 0.13 mol kg^{-1} which indicates that the identification of MOFs with promising CO_2 capture potential is significantly dependent on the DB choice. Figure S5 shows the relation between computed Q^0_{st,CO_2} for CFU-MOFs and $K_{CO_2,ratio}$. K_{CO_2} is in agreement with Q^0_{st,CO_2} which shows that the latter is the determining factor for adsorption at infinite dilution. Relations between $K_{CO_2,ratio}$ or $N_{CO_2,ratio}$ and PV_{ratio} and $S_{acc,ratio}$ are shown in Figure S6. Among the three conditions (infinite dilution, 1 bar, and 20 bar), the impact of PV and S_{acc} on N_{CO_2} is the highest at 20 bar due to the increased number of CO_2 molecules inside the MOFs at high pressures. To examine the impact of electrostatic interactions on CO_2 adsorption, we compared K_{CO_2} and N_{CO_2} of 128 representative CFU-MOFs with and without considering EICM in Figure S4(c,d). When the partial charges are assigned to frameworks, the differences between the computed gas uptakes of CFU-MOFs of the two DBs are more pronounced. For example, a MOF has a $K_{CO_2,ratio}$ ($N_{CO_2,ratio}$) of 8.2 (8.3) when EICM are neglected, but this increases to 22.3 (20.2) when EICM are considered. Thus, CO_2 uptakes drastically differ based on the DB used when EICM is considered in simulations.

Effect of structural discrepancies on adsorbent evaluation metrics of MOFs

The initial assessment for gas separation performances of MOFs is usually done using S_{ideal} of MOFs in HTCS studies.[13,33] In Figure S7, we compared ideal CH_4/H_2 selectivities of 1109 CFU-MOFs at infinite dilution, 1, 20, and 65 bar. The highest differences are observed at infinite dilution. For example, a CFU-MOF taken from the CSDSS DB is computed to have an S_{ideal} of 1369 at infinite dilution, but when the same MOF is used from CoRE DB, its selectivity is 92. Therefore, a MOF identified to be highly promising when CSDSS DB is used can be eliminated because of its low selectivity when CoRE DB is used. At 65 bar, selectivities of MOFs in both DBs are in a very narrow range, 1–5, hence selectivities computed at high pressures are less sensitive to the DB selection. More importantly, even if the selectivity of a MOF at 65 bar would change based on the DB used, a promising MOF would not be missed due to low selectivity values. Therefore, DB selection has paramount importance on the accurate identification of the MOFs with useful selectivities at low pressures.

Computationally demanding mixture simulations are required to evaluate the exact gas separation performances of MOFs.[11,34] Therefore, we compared mixture gas uptakes of CFU-MOFs obtained from CH_4/H_2:50/50 adsorption simulations at 0.1, 1, and 10 bar in Figure S8. Similar to the results of single-component gas simulations, CFU-MOFs retrieved from CoRE DB generally give higher uptakes than CSDSS DB. If molecular simulations of CFU-MOFs are performed at low pressure to compute the mixture gas uptakes, DB selection would be more critical compared to high pressure. The sensitivity of adsorbent performance evaluation metrics S_{mix}, APS, and R% (given in Table S4) to the MOF DB is then examined. We compare all 3543 MOFs (CFU-MOFs and CFM-MOFs) based on their S_{mix} at PSA condition (P_{ads}:10 bar and P_{des}:1 bar) in Figure 3a. As expected, results for all metrics of CFM-MOFs have a good agreement. On the other hand, calculated metrics of CFU-MOFs alter based on the DB used. One of the striking examples is that a CFU-MOF in CSDSS DB (refcode: HIDCUU) can be considered as very promising due to its high S_{mix} (380), while its CoRE DB version has a moderate S_{mix} (39) at 10 bar. CSDSS version includes dimethylammonium cations (which should reside within the anionic structure to create the charge balance[35]) that leads to narrower pores and higher selectivities compared to its CoRE version. Figure 3b also shows that the range of calculated APS for CFU-MOFs in CoRE DB is much larger (2.5×10^{-2}–709 mol kg^{-1}) than the one for CSDSS DB (0.1–431.9 mol kg^{-1}), which affect the identification of the top adsorbents.

Materials having R% > 85% are considered as potentially the best candidates because of their reusability.[36] Figure 3c shows that many CFU-MOFs in CSDSS DB are estimated to have a high R% (85%) whereas they are predicted to have moderate R% (60%) when CoRE DB is used, and vice versa. We identified the top 10 CFU-MOFs of both DBs offering R% > 85% with the highest APSs in Figure 3c. The most critical result is that there is no common CFU-MOF in these two top MOFs lists. For example, FABGUM is one of the top MOFs identified in CoRE DB (R%:86.3%, APS:91.4 mol kg^{-1}). However, when this MOF is taken from CSDSS DB, R% is calculated as 61%, making it non-promising. Thus, different MOFs would be identified as the best materials based on the DBs used when R% and APS are used as evaluation metrics. To better understand which metric is more affected by the selection of DB, we compared $R\%_{ratio}$ and APS_{ratio} of CFU-MOFs. Figure 3d shows that $R\%_{ratio}$ is in a narrow range of 0.5–12.4, whereas the APS_{ratio} varies from 3.47×10^{-2} to 90.3, indicating that APS is more sensitive to DB selection than R%. We observed the similar results when CFU-MOFs were studied at VSA condition (P_{ads}:1 bar and P_{des}:0.1 bar) as shown in Figure S9. APSs of CFU-MOFs in two DBs are remarkably different suggesting that working capacity and selectivity used to compute APSs are more affected by structural curations at low pressures.

 © 2021 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH

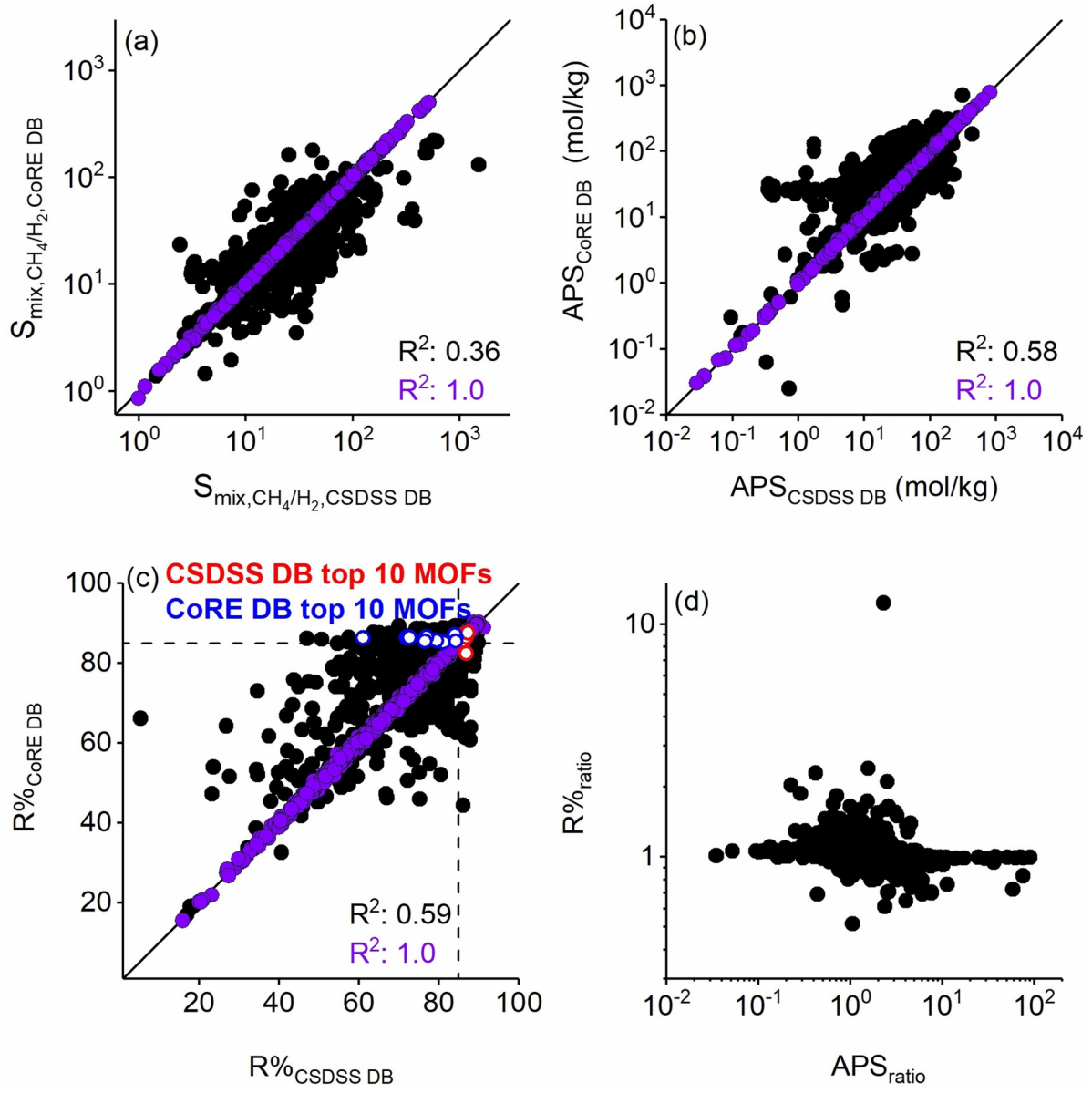


Figure 3. a) S_{mix} computed at 10 bar, b) APS, c) R% of 1109 CFU-MOFs (black), and 2434 CFM-MOFs (purple) at PSA condition. Blue and red symbols in (c) represent the top 10 MOFs of two DBs. d) APS_{ratio} and $R\%_{ratio}$ of 1109 CFU-MOFs computed at PSA condition.

Elemental analysis of structural discrepancies

To provide an atomic-level understanding on the structural discrepancies, we categorized CFU-MOFs into 13 groups considering the most common differences in chemical formulas: only carbon (C), only hydrogen (H), only nitrogen (N), only oxygen (O), only metal (M), only halogen (X), (O + M), (O + H), (C + H), (O + X), (C + H + N), (C + H + O), and (C + H + N + O). The CFU-MOFs in each group are mutually exclusive.[37] Colors in Figure 4 a–c represent the $N_{i,ratio}$ of MOFs for H_2, CH_4, CO_2 uptakes at 1 bar. Red (blue) bars represent CFU-MOFs in CSDSS (CoRE) DB whose gas uptakes are higher than the ones in CoRE (CSDSS) DB, while green bars represent CFU-MOFs in both DBs with similar gas uptakes. The number of MOFs with removed atoms is generally more in CoRE DB due to their solvent removal procedure such as detecting and removing coordinated solvents. However, CSDSS DB has more MOFs with missing atoms for two categories, (H) and (O + H). This may be due to two reasons: (i) there may be MOFs reported without Hs in CSDSS DB but Hs were added to the structures before they were reported in CoRE DB, (ii) CoRE DB did not remove non-neutral hydroxyl groups when detected whereas solvent removal script of CSDSS removed these groups.

The largest number of CFU-MOFs in Figure 4 is in the (O) group, and Os are mostly missing in CoRE MOFs. This may be attributed to many different types of O groups available in experimental structures, such as free O atoms, double-bonded O atoms, and O of solvent (water) with missing H atoms. Another major group in Figure 4 is (C + H + O) group which was mostly removed by CoRE DB and

 © 2021 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH

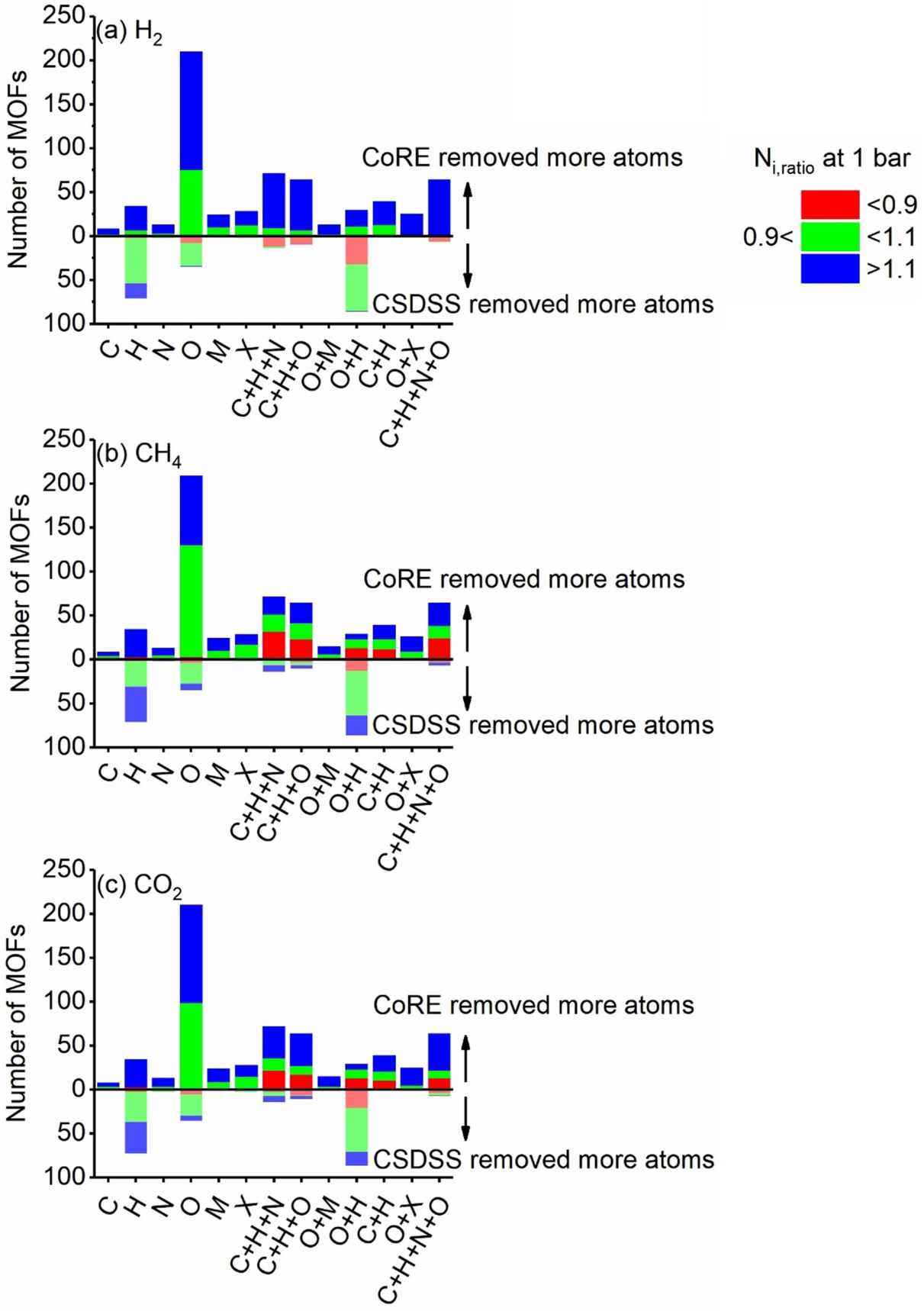


Figure 4. Elemental differences between 1109 CFU-MOFs of two DBs. Colors represent a) $N_{H_2,\ ratio}$, b) $N_{CH_4,\ ratio}$, c) $N_{CO_2,\ ratio}$ of MOFs at 1 bar.

 © 2021 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH

their removal has a remarkable impact on the gas uptakes since these groups (bound or unbound) are observed to be inside the pores. Overall, curations on groups including O moieties may have a notable effect on the simulated gas uptakes if O atoms reside as the main adsorption sites in the pores.[38]

Removal of atoms/groups inside the pores have different effects such as a change in chemical composition, pore sizes, number of favorable adsorption sites, and all have different impacts on simulated gas uptakes. Figure 4a shows that N_{H_2} is predicted mostly higher in CFU-MOFs where CoRE DB removed more atoms/groups than CSDSS. The increase in pore size upon the removal of those atoms has a more pronounced effect on H_2 uptake than the change in chemical composition of the CFU-MOF. For the MOFs in (C + H + N) group, N_{CH_4} and N_{CO_2} (Figure 4b,c) are computed to be higher in many CSDSS MOFs although those in CoRE DB have more removed groups and more free space. The removed groups in those MOFs were observed to be aromatic rings or ions which may act as adsorption sites.[39]

Figure 4 shows that structural differences of CFU-MOFs in two DBs may be due to the removal of only a single type or a group of few types of elements. Thus, we categorized oxygens, halogens, hydrogens, and metals into subgroups and first investigated how the removed oxygens affect CH_4 and H_2 uptakes of CFU-MOFs. The majority of CFU-MOFs whose chemical formulas are different in two DBs due to the number of Os were categorized into four different subgroups in Figure 5. (i) Os connected to the metal atoms in the framework were removed (O-), (ii) Os of water connected to the metal atoms in the framework were removed (O of H_2O-), (iii) double-bonded O to the metal was removed (O=) or (iv) Os of unbound H_3O^+ ions were removed (O of H_3O^+) in either of the two DBs. For case (i), while $N_{CH_4,\,ratio}$ and $N_{H_2,\,ratio}$ of CFU-MOFs are in the range of 0.8-3 and 0.8-2, respectively, the average $N_{CH_4,\,ratio}$ and $N_{H_2,\,ratio}$ values are almost 1. In other cases (ii, iii, and iv), the same trend is observed, indicating that removal of only O atom does not significantly alter uptakes.

In Figure 5, we studied halogen including subgroups, F^-, Cl^-, Br^-, SbF_6^-, PF_6^-, CF_3O_3S, BF_4^-, ClO_4^-, and $Cl^- + NO_3^-$, which were removed by either of the DBs during curations. Among the subgroups containing fluorine, SbF_6^- has the highest average $N_{CH_4,\,ratio}$ (3.3), since the removal of bulky fluorine groups by CoRE DB provided more available space for gases. Cl^- ion removal does not have an effect on the average $N_{CH_4,\,ratio}$ (1.1). If CFU-MOFs have a difference in terms of multiple ion groups (such as $Cl^- + NO_3^-$), we observed a high average $N_{CH_4,\,ratio}$ (5) and $N_{H_2,\,ratio}$ (2.3). Overall, removal of single ion groups such as F^-, Cl^-, or Br^- does not cause a major difference on gas uptakes but the absence or presence of bulky ion groups containing these halogen moieties may dramatically affect the gas uptakes of MOFs. The high average of $N_{CH_4,\,ratio}$ shows that the presence/absence of bulky fluorine and chlorine groups in MOFs should be checked before molecular simulations.

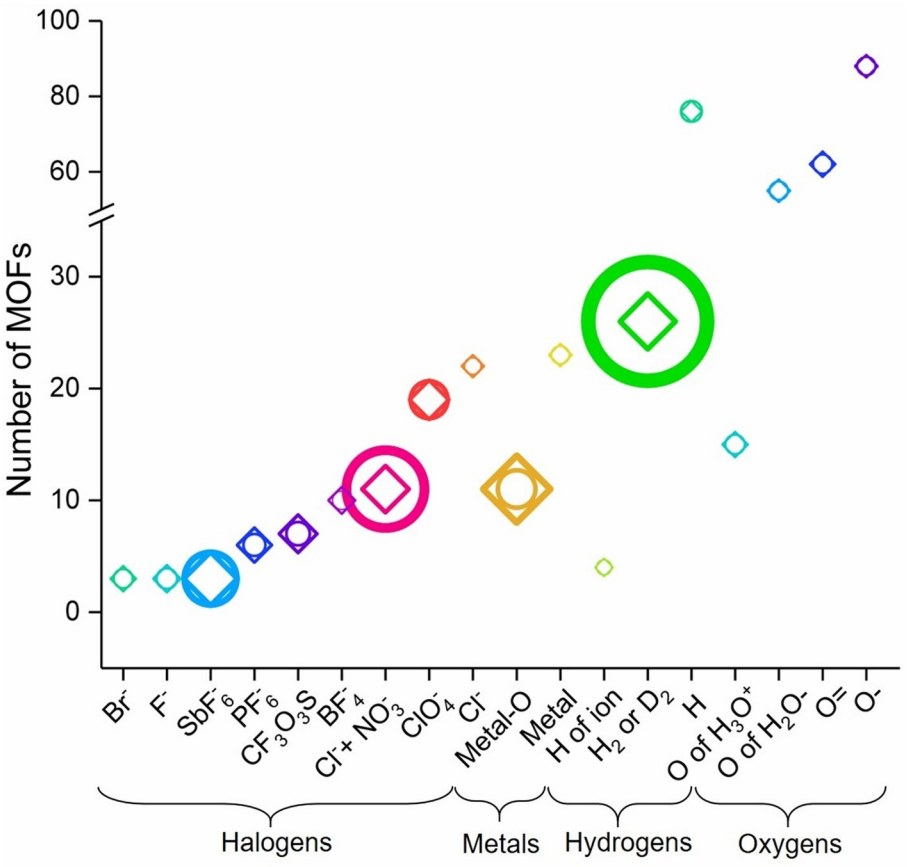
There are three subgroups in which only Hs were removed by one of the DBs: (i) H atom, (ii) adsorbed H_2 or D_2 molecules, and (iii) Hs of ions. For case (i), $N_{CH_4,\,ratio}$ is close to unity for most MOFs. However, if none of the Hs were reported in the framework, $N_{CH_4,\,ratio}$ can be high, up to 4.6. In case (ii), removal of adsorbed H_2 or D_2 molecules by CoRE DB inside the pores results in a high average $N_{CH_4,\,ratio}$ (7.6) and $N_{H_2,\,ratio}$ (2.8) due to the availability of more adsorption sites. In (iii), removal of Hs from ion groups does not have an effect on the average $N_{CH_4,\,ratio}$ (1). While differences in the number of Hs is not a determining factor for most MOFs, adsorbed H_2 or D_2 remaining inside the pores leads to significant differences in N_{CH_4}.

There is a variety of metal types in MOFs including alkaline (Li, Na, Mg), transition (Cd, Co, Cu, Mn, Ni, Zn), and post-transition (Pb), and removal of only metal atoms does not have a considerable impact on the average $N_{CH_4,\,ratio}$ (1.2) and $N_{H_2,\,ratio}$ (1.1). When both metal and oxygen atoms are removed, the average $N_{CH_4,\,ratio}$ (2.4) and $N_{H_2,\,ratio}$ (3.4) values are high. We finally note that the number of MOFs studied for the aforementioned groups is quite low, which may not be enough to make a general statement for all MOFs. However, the change in gas uptakes due to structural curations indicates the significance of the absence/presence of certain elemental subgroups in MOFs.

Conclusion

We showed that a significant number of MOFs (1109) in two available, updated computation-ready DBs, CoRE DB and CSDSS DB, are reported with the same refcode but with different chemical composition due to the procedures followed by the DBs to make MOFs directly usable in molecular simulations. Simulated CH_4, H_2, and CO_2 uptakes of MOFs especially at low pressure (infinite dilution and 1 bar) strongly depend on the selection of the computation-ready MOF DB, which may critically change our assessment about the gas storage potential of MOFs. Atomic-level examination of structures reveals that the choice of DB is crucial for MOFs having guest H_2 or D_2, metal-O complexes, and/or bulky halogen groups since these materials have significantly different gas uptakes depending on from which DB they are taken. The choice of the DB leads to large variations in calculated ideal selectivities of MOFs at low pressure but not at high pressures. Mixture selectivities and APS of MOFs dramatically change depending on the DB used in molecular simulations, the rankings of MOFs based on these metrics, hence identification of the top materials, completely changes when different DBs are used. HTCS of MOFs is very valuable to direct the experimental efforts and resources to the best materials and our results highlight that the selection of the computation-ready MOF DB in HTCS studies have an extremely important role for the accurate assessment of gas storage and separation potentials of MOFs. We finally close by noting that having an internally consistent, accurate, and continuously updated computation-ready MOF DB, which is established based on a standardized, well-defined protocol that accurately treats missing hydrogen atoms, CBIs, guest/solvent molecules, and/or other structural features that should be fixed to obtain the precise representation of MOFs, would be extremely useful. We believe that this can be possible by

 © 2021 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH



Group	Subgroup	Number of MOFs	Range of $N_{CH_4,\ ratio}$ ($N_{H_2,\ ratio}$)	Average of $N_{CH_4,\ ratio}$ ($N_{H_2,\ ratio}$)
oxygens	O of H_3O^+	15	0.95-2.16 (1.01-2.04)	1.24 (1.25)
	O of H_2O	55	0.72-2.76 (0.79-1.67)	1.17 (1.14)
	O=	62	0.68-5.56 (1.00-2.70)	1.21 (1.27)
	O-	88	0.76-2.96 (0.80-1.98)	1.17 (1.14)
halogens	Br^-	3	1.04-1.46 (1.13-1.44)	1.22 (1.25)
	F^-	3	1.04-1.90 (1.0-2.05)	1.32 (1.35)
	SbF_6^-	3	0.99-7.40 (1.53-4.57)	3.35 (2.68)
	PF_6^-	6	1.03-2.10 (1.30-2.01)	1.41(1.72)
	CF_3O_3S	7	0.97-2.03 (1.27-2.61)	1.50 (1.89)
	BF_4^-	10	0.72-1.61 (1.09-1.67)	1.11(1.34)
	$Cl^-+NO_3^-$	11	4.50-5.40 (2.07-2.45)	5.00 (2.32)
	ClO_4^-	19	0.85-6.50 (1.17-3.86)	2.38 (1.95)
	Cl^-	22	0.92-1.70 (0.95-1.70)	1.07 (1.12)
metals	Metal-O	11	0.99-6.56 (1.01-11.42)	2.42 (3.36)
	Metal	23	1.00-1.78 (1.01-1.39)	1.16 (1.14)
hydrogens	H of ion	4	0.56-1.36 (0.66-1.30)	0.97 (0.91)
	H_2 or D_2	26	1.13-19.26 (1.13-5.02)	7.57 (2.74)
	H	76	0.40-4.56 (0.56-1.32)	1.34 (1.04)

Figure 5. Analysis of the relationship between structural differences of MOFs in two DBs and their effects on CH_4 and H_2 uptakes of CFU-MOFs at 1 bar. — represents the single bond = represents the double bond. Size of circles represents the average of $N_{CH_4,\ ratio}$ Size of diamonds represents the average($N_{H_2,\ ratio}$).

 © 2021 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH

the collaboration of DB producers, simulators using DBs, and experimentalists and development of such a MOF DB provides a great opportunity to unlock the potential of these fascinating materials for various applications.

Acknowledgements

S.K. acknowledges ERC-2017-Starting Grant. This study has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (ERC-2017-Starting Grant, grant agreement No 756489-COSMOS).

Conflict of interest

The authors declare no conflict of interest.

Keywords: adsorption · database · high-throughput screening · metal–organic frameworks (MOFs) · molecular modelling

[1] A. Helal, Z. H. Yamani, K. E. Cordova, O. M. Yaghi, *Natl. Sci. Rev.* **2017**, *4*, 296 – 298.
[2] O. M. Yaghi, *ACS Cent. Sci.* **2019**, *5*, 1295 – 1300.
[3] F. H. Allen, *Acta Crystallogr. Sect. B* **2002**, *58*, 380 – 388.
[4] D. Ongari, L. Talirz, B. Smit, *ACS Cent. Sci.* **2020**, *6*, 1890.
[5] Z. Hu, Y. Wang, B. B. Shah, D. Zhao, *Adv. Sustainable Syst.* **2019**, *3*, 1800080.
[6] Y. He, W. Zhou, G. Qian, B. Chen, *Chem. Soc. Rev.* **2014**, *43*, 5657 – 5678.
[7] H. W. Langmi, J. Ren, B. North, M. Mathe, D. Bessarabov, *Electrochim. Acta* **2014**, *128*, 368 – 392.
[8] J. Rogacka, A. Seremak, A. Luna-Triguero, F. Formalik, I. Matito-Martos, L. Firlej, S. Calero, B. Kuchta, *Chem. Eng. J.* **2021**, *403*, 126392.
[9] S. Nandi, P. De Luna, T. D. Daff, J. Rother, M. Liu, W. Buchanan, A. I. Hawari, T. K. Woo, R. Vaidhyanathan, *Sci. Adv.* **2015**, *1*, e1500421.
[10] H. Daglar, S. Keskin, *J. Phys. Chem. C* **2018**, *122*, 17347 – 17357.
[11] C. Altintas, I. Erucar, S. Keskin, *ACS Appl. Mater. Interfaces* **2018**, *10*, 3668 – 3679.
[12] H. Daglar, S. Keskin, *Coord. Chem. Rev.* **2020**, *422*, 213470 – 213490.
[13] I. Erucar, S. Keskin, *Front. Mater.* **2018**, *5*, 004.
[14] J. Jiang, R. Babarao, Z. Hu, *Chem. Soc. Rev.* **2011**, *40*, 3599 – 3612.
[15] Y. G. Chung, J. Camp, M. Haranczyk, B. J. Sikora, W. Bury, V. Krungleviciute, T. Yildirim, O. K. Farha, D. S. Sholl, R. Q. Snurr, *Chem. Mater.* **2014**, *26*, 6185 – 6192.
[16] P. Z. Moghadam, A. Li, S. B. Wiggin, A. Tao, A. G. P. Maloney, P. A. Wood, S. C. Ward, D. Fairen-Jimenez, *Chem. Mater.* **2017**, *29*, 2618 – 2625.
[17] A. Li, R. Bueno-Perez, S. Wiggin, D. Fairen-Jimenez, *CrystEngComm* **2020**, 1 – 10.
[18] Y. G. Chung, E. Haldoupis, B. J. Bucior, M. Haranczyk, S. Lee, H. Zhang, K. D. Vogiatzis, M. Milisavljevic, S. Ling, J. S. Camp, *J. Chem. Eng. Data* **2019**, *64*, 5985 – 5998.
[19] T. Chen, T. A. Manz, *RSC Adv.* **2020**, *10*, 26944 – 26951.
[20] C. Altintas, G. Avci, H. Daglar, A. Nemati Vesali Azar, I. Erucar, S. Velioglu, S. Keskin, *J. Mater. Chem. A* **2019**, *7*, 9593 – 9608.
[21] S. Velioglu, S. Keskin, *Mater. Adv.* **2020**, *1*, 341 – 353.
[22] Z. Qiao, K. Zhang, J. Jiang, *J. Mater. Chem. A* **2016**, *4*, 2105 – 2114.
[23] S. Li, Y. G. Chung, R. Q. Snurr, *Langmuir* **2016**, *32*, 10368 – 10376.
[24] G. Avci, S. Velioglu, S. Keskin, *ACS Appl. Mater. Interfaces* **2018**, *10*, 33693 – 33706.
[25] A. Ahmed, S. Seth, J. Purewal, A. G. Wong-Foy, M. Veenstra, A. J. Matzger, D. J. Siegel, *Nat. Commun.* **2019**, *10*, 1 – 9.
[26] C. M. Simon, J. Kim, D. A. Gomez-Gualdron, J. S. Camp, Y. G. Chung, R. L. Martin, R. Mercado, M. W. Deem, D. Gunter, M. Haranczyk, *Energy Environ. Sci.* **2015**, *8*, 1190 – 1199.
[27] S. Bhattacharya, S. Pal, S. Natarajan, *ChemPlusChem* **2016**, *81*, 733 – 742.
[28] Y. G. Chung, P. Bai, M. Haranczyk, K. T. Leperi, P. Li, H. Zhang, T. C. Wang, T. Duerinck, F. You, J. T. Hupp, *Chem. Mater.* **2017**, *29*, 6315 – 6328.
[29] S. Ma, J. M. Simmons, D. Sun, D. Yuan, H.-C. Zhou, *Inorg. Chem.* **2009**, *48*, 5263 – 5268.
[30] M. Fernandez, T. K. Woo, C. E. Wilmer, R. Q. Snurr, *J. Phys. Chem. C* **2013**, *117*, 7681 – 7689.
[31] S. Ma, H.-C. Zhou, *Chem. Commun.* **2010**, *46*, 44 – 53.
[32] J. Sculley, D. Yuan, H.-C. Zhou, *Energy Environ. Sci.* **2011**, *4*, 2721 – 2735.
[33] Y. J. Colón, R. Q. Snurr, *Chem. Soc. Rev.* **2014**, *43*, 5735 – 5749.
[34] J. A. Swisher, L.-C. Lin, J. Kim, B. Smit, *AIChE J.* **2013**, *59*, 3054—3064.
[35] M.-L. Ma, C. Ji, S.-Q. Zang, *Dalton Trans.* **2013**, *42*, 10579 – 10586.
[36] C. Altintas, G. Avci, H. Daglar, A. Nemati Vesali Azar, S. Velioglu, I. Erucar, S. Keskin, *ACS Appl. Mater. Interfaces* **2018**, *10*, 17257 – 17268.
[37] For example, the molecular formula of TEFWIL in CSDSS DB ($C_{240}H_{96}In_{24}N_{96}O_{192}$) differs from that of the same MOF in CoRE DB ($C_{240}H_{72}In_{24}N_{96}O_{192}$) only by 24 H atoms and this MOF was only counted in the (H) group. However, the molecular formula of OXALIJ in CSDSS DB ($C_{70}H_{46}Cd_3N_2O_{24}$) and in CoRE DB ($C_{68}H_{58}Cd_3N_2O_{24}$) differ by 2 C and 12 H atoms and this MOF was only counted in the (C + H) group.
[38] J. Liu, R. Zou, Y. Zhao, *Tetrahedron Lett.* **2016**, *57*, 4873 – 4881.
[39] K. Konstas, T. Osl, Y. Yang, M. Batten, N. Burke, A. J. Hill, M. R. Hill, *J. Mater. Chem.* **2012**, *22*, 16698 – 16708.

Manuscript received: November 15, 2020
Accepted manuscript online: January 14, 2021
Version of record online: March 1, 2021

 © 2021 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH