Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/he

A combination of multi-scale calculations with machine learning for investigating hydrogen storage in metal organic frameworks

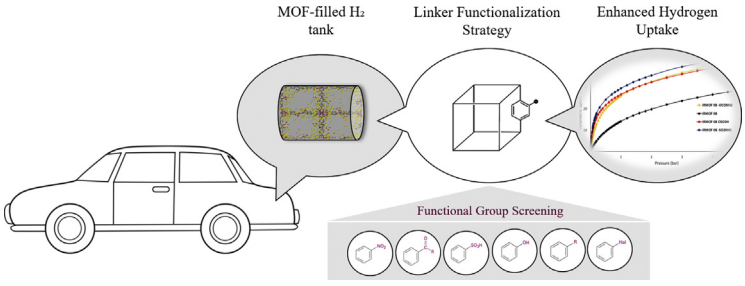
Rafaela Maria Giappa [a], *Emmanuel Tylianakis* [a,c], *Marco Di Gennaro* [d], *Konstantinos Gkagkas* [b], *George E. Froudakis* [a,*]

[a] *Department of Chemistry, University of Crete, Voutes Campus, GR-70013 Heraklion, Crete, Greece*
[b] *Material Engineering Division, Toyota Motor Europe NV/SA, Technical Center, 1930 Zaventem, Belgium*
[c] *Department of Materials Science and Technology, University of Crete, Voutes Campus, GR-70013, Heraklion, Crete, Greece*
[d] *Nanomat, Q-Mat, CESAM, European Theoretical Spectroscopy Facility, Universite de Liege, B-4000 Liege, Belgium*

HIGHLIGHTS

- Bottom up design of hydrogen storage materials.
- Functionalization of MOFs for enhanced hydrogen storage.
- Combination of multi-scale calculations with machine learning.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 18 March 2021
Received in revised form
2 June 2021
Accepted 3 June 2021
Available online 27 June 2021

Keywords:
Hydrogen storage
MOFs
Multi-scale

ABSTRACT

Combining multi-scale calculations with machine learning, we investigate how the ligand functionalization affects the hydrogen storage profile of Metal Organic Frameworks. The binding energy of hydrogen with 58 strategically selected functionalized benzenes was calculated with accurate ab-initio methods. Our results show that many functional groups (e.g. $-OPO_3H_2$, $-OCONH_2$) increase the interaction strength up to 15–25% compared to benzene while $-OSO_3H$ holds the most promise with an enhancement up to 80%. Grand Canonical Monte-Carlo calculations with interatomic potentials derived from the ab-initio calculations, verify the trend obtained from the meticulous screening. In addition, a proof of principle Machine Learning analysis is performed on the ab-initio results showing a good prediction of the H_2 binding energies even with a limited amount of data. The results from our bottom-up approach lead us to conclude that this functionalization strategy can be

* *Corresponding author.*
E-mail address: frudakis@uoc.gr (G.E. Froudakis).
https://doi.org/10.1016/j.ijhydene.2021.06.021

Ab-initio
GCMC
ML

applied to various porous materials in order to enhance their hydrogen storage performance.

Introduction

Nowadays humanity is facing two urgent, interconnected problems; the exhaust of conventional energy resources and the need to reduce CO_2 emissions and move towards a sustainable, carbon emission free economy. Within the scope of sustainability, hydrogen economy presents a possible solution to both problems. Hydrogen is considered as an ideal energy carrier since it is a completely clean, non-toxic burning fuel that can store a significant amount of energy. Fuel-cell technology is currently under intensive research and development in view of the expected benefits in facing the environmental problems and the gradual depletion of conventional fuel reserves.

One key challenge for the commercial development of fuel-cell powered vehicles is the efficient hydrogen storage [1]. On a mass basis, hydrogen has nearly three times the energy content of gasoline, with 120 MJ/kg versus 44 MJ/kg respectively. On the volume basis, however, the situation is reversed; the same amount of hydrogen requires more storage space than gasoline with liquid hydrogen having a density of 8 MJ/L whereas gasoline has 32 MJ/L [2]. Current fuel cell vehicles, manufactured and sold by automotive companies such as Honda, Hyundai and Toyota, store hydrogen in high pressure (~700 bar) carbon-fiber-reinforced fuel tanks [3]. Still, even at such high pressures, 6 kg of liquid hydrogen require a tank of 150 L volume. The goal is to design low-cost, light-weight materials that can reversibly and rapidly store hydrogen near ambient conditions. The U.S. Department of Energy (DOE) has set the ultimate technical targets of onboard hydrogen storage for light-duty fuel cell vehicles to be 79.2 MJ/kg for the useable gravimetric and 6.1 MJ/L for the useable volumetric capacities [4]. To overcome the challenges in the existing technologies as well as meet the DOE targets, several strategic pathways have been proposed and are currently an active area of multidisciplinary research [5]. Among the existing technologies, an alternative way of storing hydrogen is the adsorption in nanoporous materials, such as Metal Organic Frameworks (MOFs). MOFs are attractive candidates possessing high surface areas so that hydrogen could be stored in smaller, more manageable volumes and at relatively low pressures.

Introduced by O. Yaghi in 1995 [6], MOFs are a class of crystalline materials consisting of inorganic metal ions or metal clusters linked by multidentate organic linkers. The structure of MOFs is characterized by an open, porous framework and is associated with significant properties such as low density, high surface area and high porosity. Due to their structural diversity and tunability, their properties can be tailored in an attempt to identify the optimal composition for the desirable application. After O. Yaghi first reported the high H_2 uptake capacity of MOF-5 in 2003 [7], MOFs have been intensively researched as hydrogen adsorbents. In order to be promising for H_2 storage applications, a porous material should show an uptake that is ideally large and the adsorption should be strong and reversible.

Many different strategies have been up-to-date employed towards enhancing the hydrogen storage capacity of MOFs. Some of them include control of the pore size (e.g. increase of BET surface area, impregnation, catenation), introduction of open metal sites in the metal oxide part and in the organic linker, doping of alkali elements onto the organic linker, functionalization of the organic linker and hydrogen spill-over [8]. S. Yu and co-workers performed a combined DFT and GCMC study of the electronic structures and hydrogen adsorption properties of the Zn-based MOF-650 [9]. As $B(OH)_2$ linker substituted the original linkers as ligand, and the active metals Mg/Ca substituted Zn as the center metal, the H_2 storage amount of the resulting MOF was increased by approximately 20 wt%, concluding that H_2 molecules tend to be attracted by electron donors and metal atoms with low electronegativity.

For practical hydrogen storage applications, volumetric and gravimetric hydrogen capacities are the key factors that determine the size and weight of the MOF-filled tank required to store hydrogen for a reasonable driving range. These factors must be optimized so the tank is neither too large nor too heavy. Recently, several high-throughput studies have searched databases of real and hypothetical MOFs in order to identify candidates that simultaneously exhibit high volumetric and gravimetric hydrogen uptakes. Notably, in the work of Siegel and coworkers [10], nearly 500,000 MOFs were screened computationally and the most promising candidates were assessed experimentally afterwards, with three MOF candidates, SNU-70, UMCM-9, and PCN-610/NU-100, found to surpass the capacities of IRMOF-20, the up-to-date record holder for balanced hydrogen capacity [10].

As a way to increase the H_2 physisorption energy of MOFs, the functionalization of organic linkers shows a good effect and is the focus of this study. Currently, organic linkers of many MOF structures consist of aromatic backbones such as benzene and naphthalene. Thus, a simple, viable aromatic system that can be employed and studied extensively with computationally demanding ab-initio calculations is the singly substituted benzene ring. According to ab-initio calculations performed so far, the H_2 binding energy to aromatic organic linkers increases with the introduction of $-NH_2$, $-CH_3$, and $-OH$ groups due to their ability to enrich the aromatic system electronically [11]. In the work of Xia et al. the introduction of $-OH$, $-NO_2$, $-CH_3$, $-CN$, $-I$ functional groups on the phenyl ring of the organic linker of MOF-808 was found to have a favorable impact on the hydrogen adsorption profile of the framework at cryogenic

temperatures, with MOF-808-CN showing the highest capacity of 3.15 wt% at 77 K and 1 bar [12]. One of the most promising functional groups proposed is the -OLi alkoxide with interaction energies between the hydrogen molecules and this functional group to be up to three times larger compared to the unmodified MOF [13]. The gravimetric capacity of the Li-modified MOFs was calculated to 10 wt % at 77 K and 100 bar, while the corresponding values show great promise also at room temperature with an uptake of 4.5 wt %.In the -OLi alkoxide group, the oxygen atom serves as a promising attaching atom. Another promising attaching atom was found to be the sulfate anion [14] which can be introduced using the sulfonic acid (H−SO_3H) group. Calculations have shown that −SO_3H and −SO_3Li modified MOF candidates exhibit significantly increased hydrogen adsorption capacities [15].

Following the pathway of the linker functionalization, we investigated the effect of a series of aromatic substitutions by means of multiscale theoretical techniques. In an attempt to identify a promising functional group and also shed light on the effect of the functional group composition to the hydrogen uptake profile, we calculated the binding energy of H_2 to 58 singly functionalized phenyl rings. Subsequently, following the trend obtained by the ab-initio screening, Grand Canonical Monte Carlo (GCMC) simulations were employed for three candidates of the IRMOF series; IRMOF-8, IRMOF-14 and IRMOF-16, modified with some of the best functional group candidates. The total gravimetric and volumetric uptake profile was obtained both at cryogenic (77 K) and ambient temperature (300 K). This functionalization strategy can be applied in numerous porous materials and guide both the post-synthetic modification for the already synthesized as well as the targeted synthesis of new framework candidates. In parallel, we used our multiscale results as a database for a Machine Learning (ML) analysis. The full database of local minima (1138 data points) was analyzed with various ML algorithms and representations in order to show that starting with an atomistic description and a large enough database, one can reach chemical accuracy, i.e. k_BT or 0.598 kcal/mol or ~10^{-4}Ha. This offers an important alternative to costly quantum mechanical calculations.

Methodology

By virtue of consistency, all geometry optimizations were held at the MP2/def2-TZVPP level of theory using the approximate resolution of the identity (RI-MP2) [16] and the frozen-core approximation. Calculations were performed with the TURBOMOLE software package [17]. Numerical frequency calculations were performed to all the different optimized structures arisen, to verify them as stationary points on the potential energy surface. Binding energy values were corrected for the Basis Set Superposition Error (BSSE) by the full counterpoise method proposed by Boys and Bernardi [18]. Electrostatic potential maps of the monomers and electron density redistribution plots of the complexes were generated in an attempt to gain insights on the nature of the interaction of hydrogen with the functionalized benzenes.

In order to investigate how linker functionalization impacts the H_2 adsorption in MOFs, we functionalized three MOF candidates; IRMOF-8, IRMOF-14, and IRMOF-16; and performed classical Monte Carlo simulations in the Grand Canonical ensemble (GCMC) [19]. We studied the hydrogen uptake profile of both the parent structures and the ones functionalized with some of the best candidates. The different organic linkers were doubly functionalized. Positions of the IRMOF framework atoms were obtained from crystallographic information files available at the Cambridge Structural Database (CSD) [20].

Interactions of hydrogen guest molecules with all framework other than the functional group (FG) atoms were calculated employing DREIDING parameters [21] for framework atoms and empirically derived parameters for H_2. For the description of the intermolecular interactions between H_2 molecules and the FG atoms, we used ab-initio derived parameters. Details on the procedure of the fitting can be found in the SI section. Coulomb interactions were taken into account by calculating the partial atomic charges for atoms in each model cluster at the B3LYP/def2-TZVP level of theory employing the CHELPG method [22]. The Darkrim-Levesque (D-L) model [23] was used to represent for H_2 and the H−H bond length is held fixed at 0.741 Å.

GCMC calculations were performed with the RASPA software package [24]. Periodic boundary conditions were applied to all three dimensions. All Lennard-Jones interactions were calculated up to a cut-off distance of 12.8 Å. All Coulomb interactions were handled using the Ewald summation technique [25] and quantum effects were taken into account by the Feynman-Hibbs effective potential [26]. Simulations were performed in supercells incorporating enough repeat units such that all edge lengths were greater than 25.6 Å; i.e. twice the LJ cut-off radius. The framework was considered rigid with all framework atoms held fixed throughout the simulations. For each simulation point, 50,000 cycles were performed for system equilibration followed by additional 100,000 cycles for sampling over the ensemble averages. Each cycle comprised of N steps, where N was the number of molecules in the system at the current loading. Monte Carlo moves attempted are translation, reinsertion, creation and deletion. The total gravimetric and volumetric uptake isotherms at 77 and 300 K were obtained.

In parallel, accounting for the full dataset of local minima calculated for the functionalized benzene molecules with the H_2 binding in different positions, we tested different types of ML techniques. The full database accounts for 1138 local minima with energies calculated at the MP2/def2-TZVPP level of theory. For ML we tested three different regression methods, namely Kernel Ridge, Random Forest and Support Vector algorithms (dubbed respectively as KRR, RFR and SVR respectively) as implemented in Scikit-learn [27] (more details are given in the SI). Three state of the art descriptors from literature were employed to map the chemical information into machine readable string of data; Coulomb Matrix (CM) [28], Bag of Bonds (BOB) [29] and Smooth Overlap of Atomic Potentials (SOAP) [30]. The descriptors were obtained with the DScribe package [31] and the QML code [32]. Splitting our dataset into a training and a test set for validation, we employed a leave-one-out cross validation scheme with $c = 5$ folds and optimized the kernel at each step. Hyperparameters were optimized with an in-house genetic algorithm search.

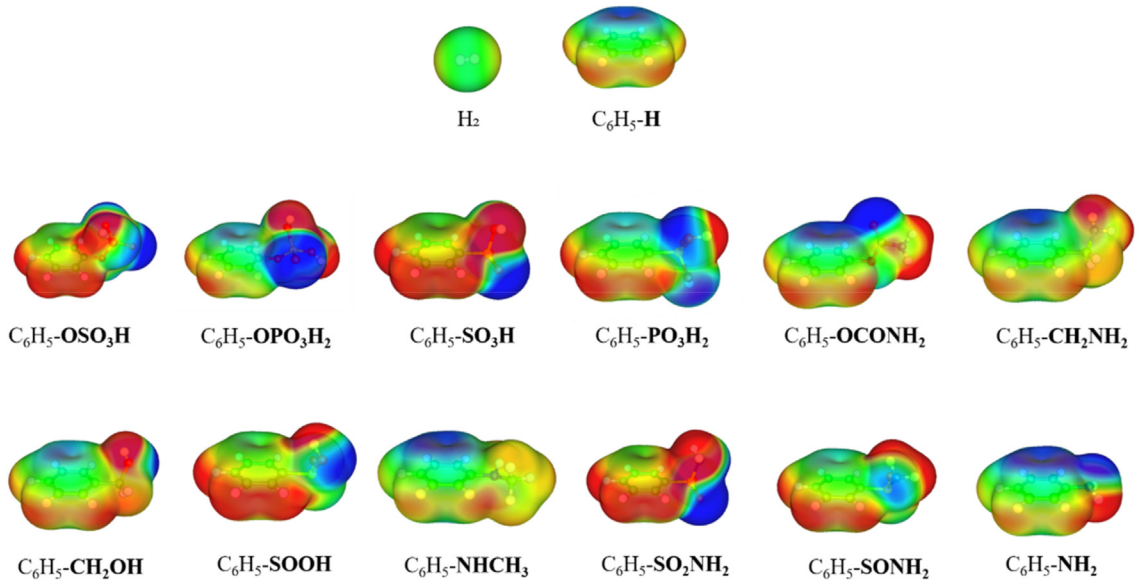


Fig. 1 – The electrostatic potential maps of the best performing systems under study, calculated at the MP2/def2-TZVPP level of theory. With blue and red the regions of low and high electrostatic potential respectively, ranging from −0.03 to +0.03 Hartree·e^{-1}. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Results & discussion

The selection of the FGs was based on electrophilic aromatic substitution pathways. It is known that electrophilic substituents can exert resonance and inductive effects that can be either electron withdrawing (e.g. $-CH_3$) or electron donating (e.g. -CF3). These effects can be visualized by electrostatic potential maps. In Fig. 1 we present the electrostatic potential maps of the best performing systems in the current study. The electrostatic potential maps of all the systems can be found in the SI section. An extensive conformational search was held for each of the systems; the shown structures are the energetically most favorable configurations for each system optimized at the MP2/def2-TZVPP level of theory.

In Table 1 we present the most promising systems under study along with the corresponding binding energy values in kJ/mol and the binding strength enhancement with respect to benzene. Fig. 2 shows the global minimum configuration of these systems. Results for all 58 systems can be found in the SI section.

In Fig. S2 found in the SI section, it can be seen that for the majority of the systems studied, the most energetically favorable position for H_2 is above the ring; hydrogen interacts with the π system of the functionalized rings and the strength of the interaction correlates with the ability of the substituents to enrich the aromatic system electronically. From the strongest interacting FGs shown in Fig. 2, the $-SO_3H$, $-PO_3H_2$, $-OPO_3H_2$ and $-OSO_3H$ candidates were able to outperform the delocalized aromatic electron distribution and thus exhibit an increased interaction, with binding energies >20% enhanced when compared to benzene. From these, $-OSO_3H$ stands out with an enhancement of ~80%. This is a special case as it combines the interaction with the delocalized cloud and the interaction of the polarized hydrogen of the FG with the quadrupole moment of the H_2 molecule. An interesting case of a well performing FG is also $-OCONH_2$ in which hydrogen seems to take advantage of both the delocalized π system of the ring and the FG's influence.

Aiming to provide a visualization of the changes in the electron density upon interaction with hydrogen, electron density difference plots were calculated at the MP2/def2-TZVPP level of theory and are shown in Fig. 3 for the best performing systems and in Fig. S3 of the SI section for all the systems under study.

Table 1 – [a] Sorted binding energies (kJ/mol) of the strongest interacting $H_2\cdots C_6H_5$-X systems, calculated at the RI-MP2/def2-TZVPP level of theory. All interaction energy values have been corrected for the Basis Set Superposition Error (BSSE) by the full counterpoise method [18].[b] Percentage of binding energy enhancement with the introduction of the FG compared to benzene.

System	Binding Energy (kJ/mol)[a]	Binding Energy Enhancement (%)[b]
$H_2\cdots C_6H_5-OSO_3H$	−7.1	81%
$H_2\cdots C_6H_5$-OPO_3H_2	−4.9	26%
$H_2\cdots C_6H_5-SO_3H$	−4.8	23%
$H_2\cdots C_6H_5$-PO_3H_2	−4.8	22%
$H_2\cdots C_6H_5-OCONH_2$	−4.7	20%
$H_2\cdots C_6H_5-CH_2NH_2$	−4.7	19%
$H_2\cdots C_6H_5-CH_2OH$	−4.7	19%
$H_2\cdots C_6H_5-SOOH$	−4.7	19%
$H_2\cdots C_6H_5-NHCH_3$	−4.7	19%
$H_2\cdots C_6H_5-SO_2NH_2$	−4.6	18%
$H_2\cdots C_6H_5-SONH_2$	−4.6	16%
$H_2\cdots C_6H_5-NH_2$	−4.5	16%
$H_2\cdots C_6H_5-H$	−3.9	(ref)

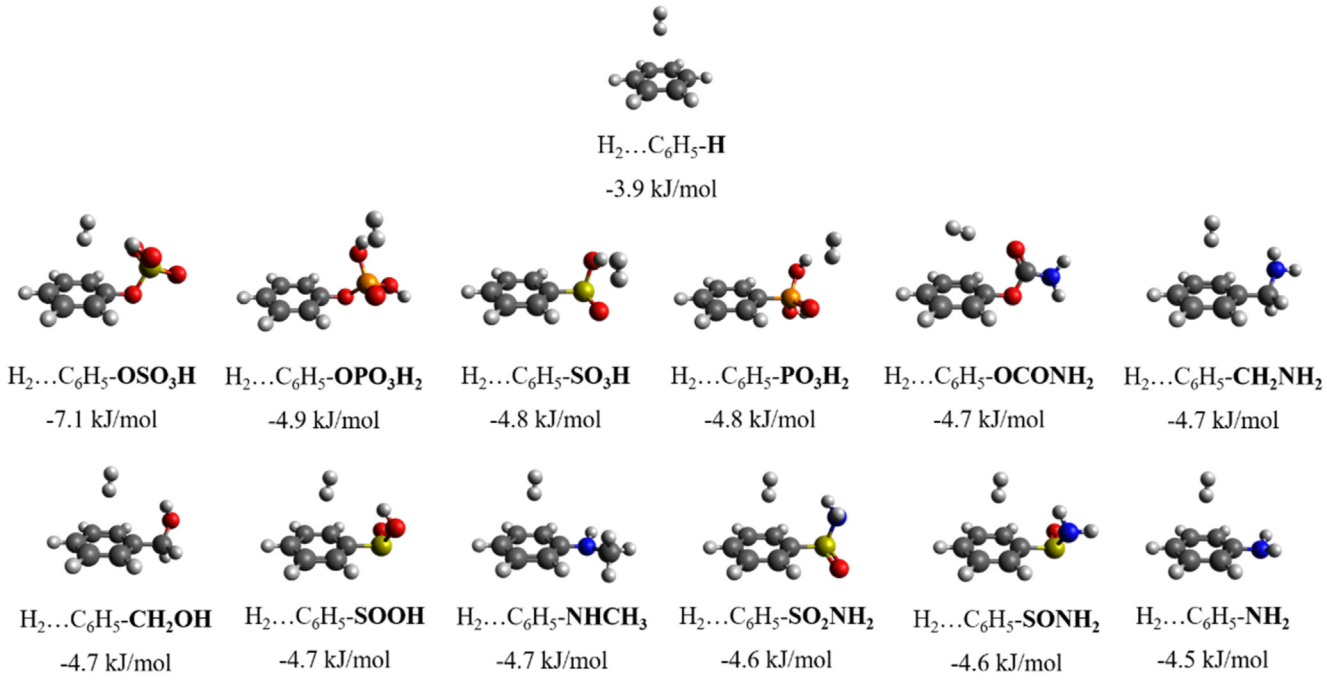


Fig. 2 – Global minima of the optimized geometries for the best performing FGs.

The largest electron density redistributions are observed for the strongest interacting functional groups; $H_2\cdots C_6H_5$-SO_3H, $H_2\cdots C_6H_5$-PO_3H_2, $H_2\cdots C_6H_5$-OPO_3H_2 and $H_2\cdots C_6H_5$-OSO_3H exhibit polarization of the electron cloud of hydrogen and prominent polarization of the electron density of the atoms of the FGs. The strong polarization effect comes with the fact that these FGs draw the hydrogen molecule from the π system of the ring towards their side. For the cases of $H_2\cdots C_6H_5$–$OCONH_2$, $H_2\cdots C_6H_5$–SOOH, and $H_2\cdots C_6H_5$–SO_2NH_2, hydrogen has changed its orientation in such a way that

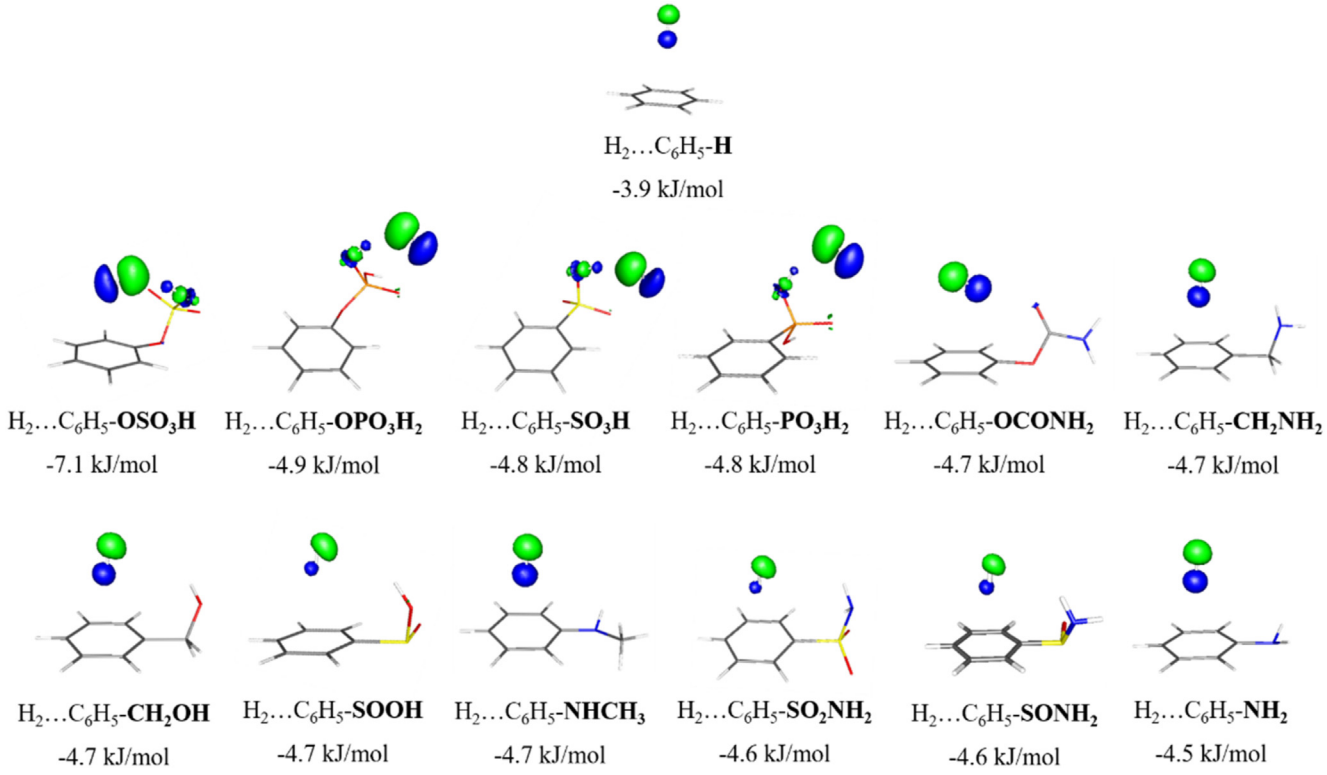


Fig. 3 – Electron-density redistribution plots of the optimized geometries of the strongest interacting $H_2\cdots C_6H_5$-X complexes. With blue and green the regions that gain and lose electron density upon the formation of the complex, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

exploits both the π system of the ring and the FG's field. Thus, although the electron energy redistribution is not that big, the binding strength is high.

From the ab-initio screening of the functional groups, four of the best performing candidates were selected for further GCMC studies, namely $-OCONH_2$, $-SO_2NH_2$, $-OPO_3H_2$ and $-OSO_3H$. For applying our linker functionalization strategy, numerous MOF structures as well as other porous materials' structures (e.g. COFs, ZIFs, etc.) are suitable. Herein, we choose IRMOF-8, IRMOF-14 and IRMOF-16 as sample cases with their organic linker part being appropriate for chemical modification. As it can be seen in Fig. 4, IRMOF-8 was doubly functionalized with $-OCONH_2$, $-SO_2NH_2$ and $-OSO_3H$, while IRMOF-14 and IRMOF-16 with the bulkier $-OPO_3H_2$ and $-OSO_3H$ FGs.

The most interesting results from the GCMC simulations were obtained at T = 77 K and are summarized in Figs. 5 and 6, while all GCMC results, both at T = 77 K and T = 300 K, can be found in the SI section.

At T = 77 K, the volumetric uptake is enhanced in the case of the functionalized IRMOF-8 compared to the parent structure, with an enhancement up to ~60% for the case of IRMOF-8-SO_2NH_2 at P = 5 bar. At low loadings (i.e. P < 5 bar) the first hydrogen molecules fill the structure and "feel" the effect of the stronger binding sites that FGs introduce. At higher loadings, where more and more hydrogen molecules enter the pore, their interaction with the FGs is screened by the H_2-H_2 interactions. Thus, the percentage of enhancement drops but is still present; in the case of IRMOF-8-SO_2NH_2 is 10–20%. Overall, the strength of the introduced binding sites of all three different FGs prevails

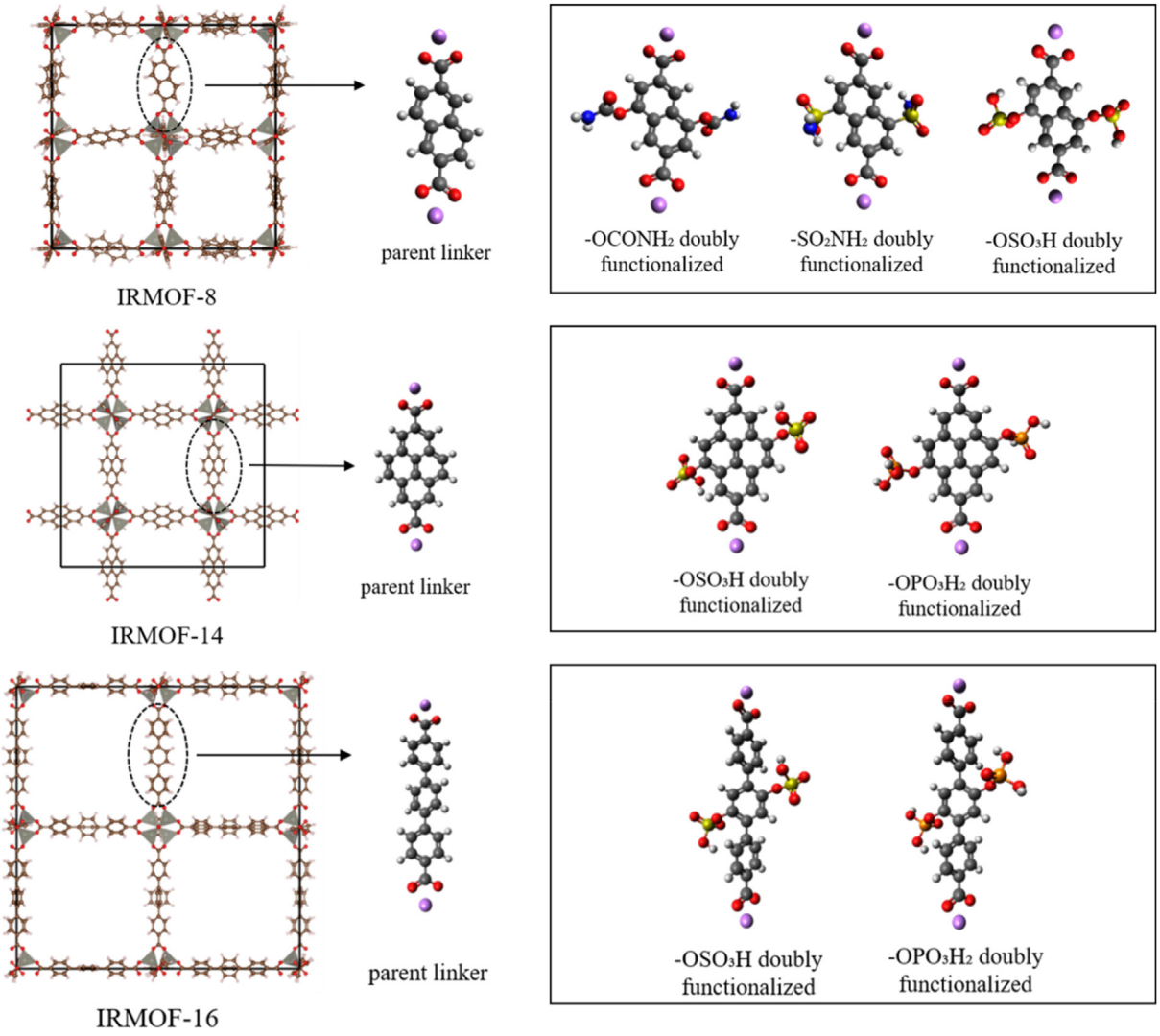


Fig. 4 – The IRMOFs selected and functionalized for this study along with the cluster model chosen to represent their organic linker part. Hydroxyl groups are terminated with a Li atom to represent the effect of the charge density of the metal cluster. All structures shown are optimized at the B3LYP/def2-TZVP level of theory. C, H, O, S, P, Li atoms as grey, white, red, yellow, orange and purple spheres respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

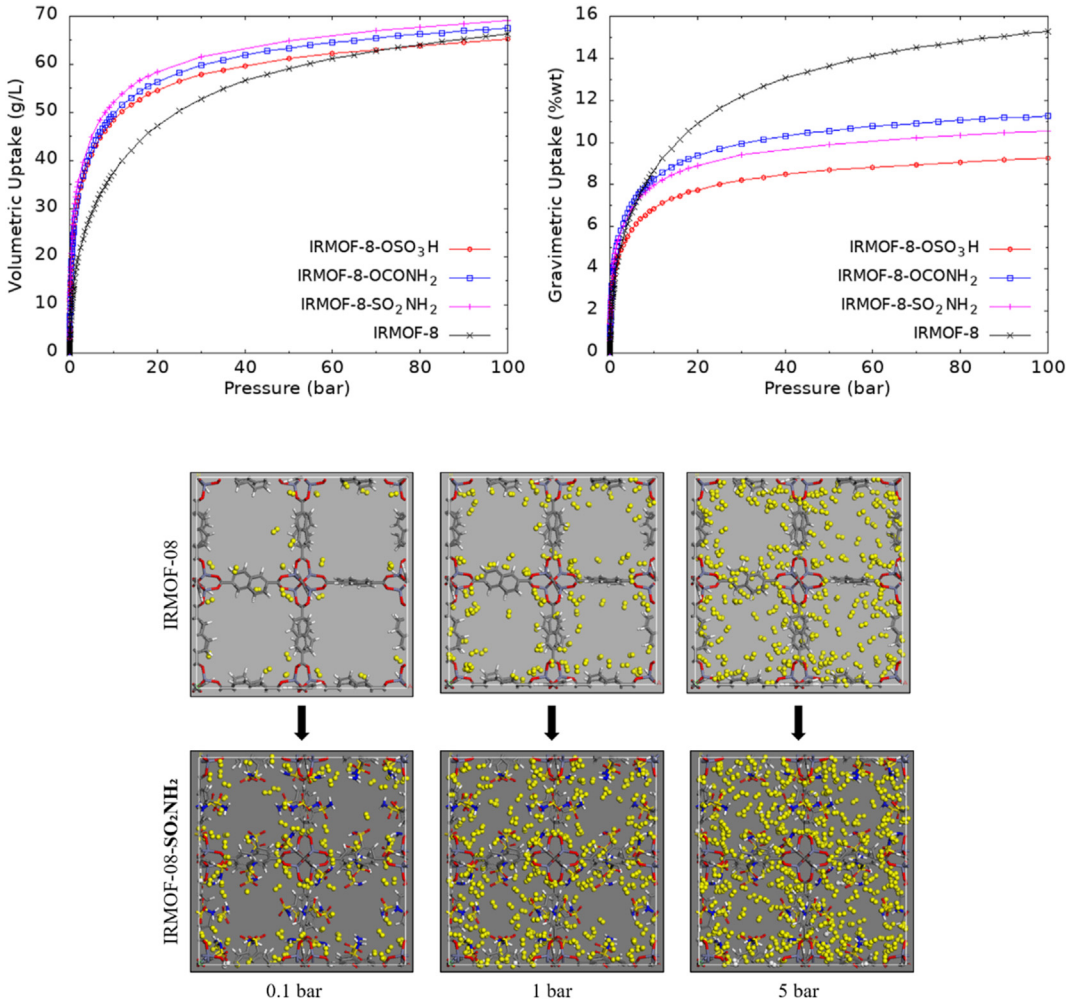


Fig. 5 – (A) Volumetric (left) and gravimetric (right) hydrogen uptake isotherms for IRMOF-8 and IRMOF-8-n (n: $-OSO_3H$, $-OCONH_2$, $-SO_2NH_2$) at T = 77 K. (B) Snapshots of IRMOF-8 and IRMOF-8-SO_2NH_2 taken from the GCMC simulations at T = 77 K and 0.1, 1, 5 bar. Guest hydrogen molecules with yellow. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

over the effect of their bulkiness for the whole pressure range, with $-SO_2NH_2$ being the most promising FG.

The gravimetric uptake shows that the strength of the introduced binding sites is not able to compensate for the weight of the FGs, with the unmodified structure clearly outperforming the functionalized counterparts at all pressures.

The volumetric uptake is enhanced in the case of the functionalized IRMOF-16 compared to the parent structure,

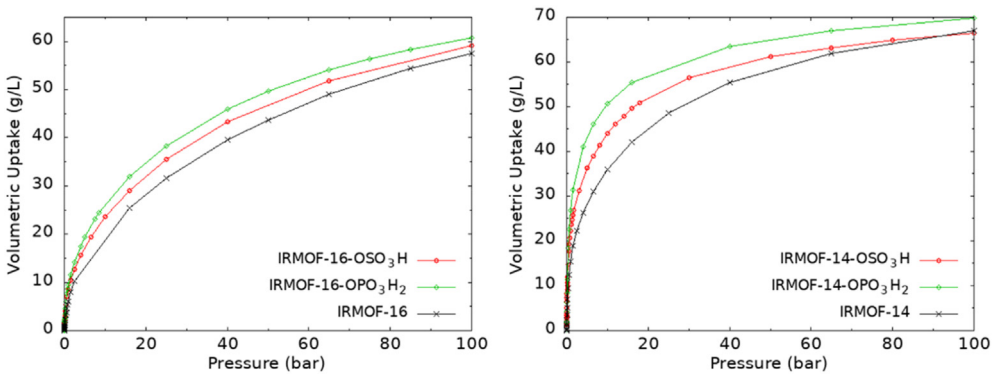


Fig. 6 – Volumetric hydrogen uptake isotherms for IRMOF-16 and IRMOF-16-n (n: $-OSO_3H$, $-OPO_3H_2$) (left) and IRMOF-14 and IRMOF-14-n (n: $-OSO_3H$, $-OPO_3H_2$) (right) at T = 77 K.

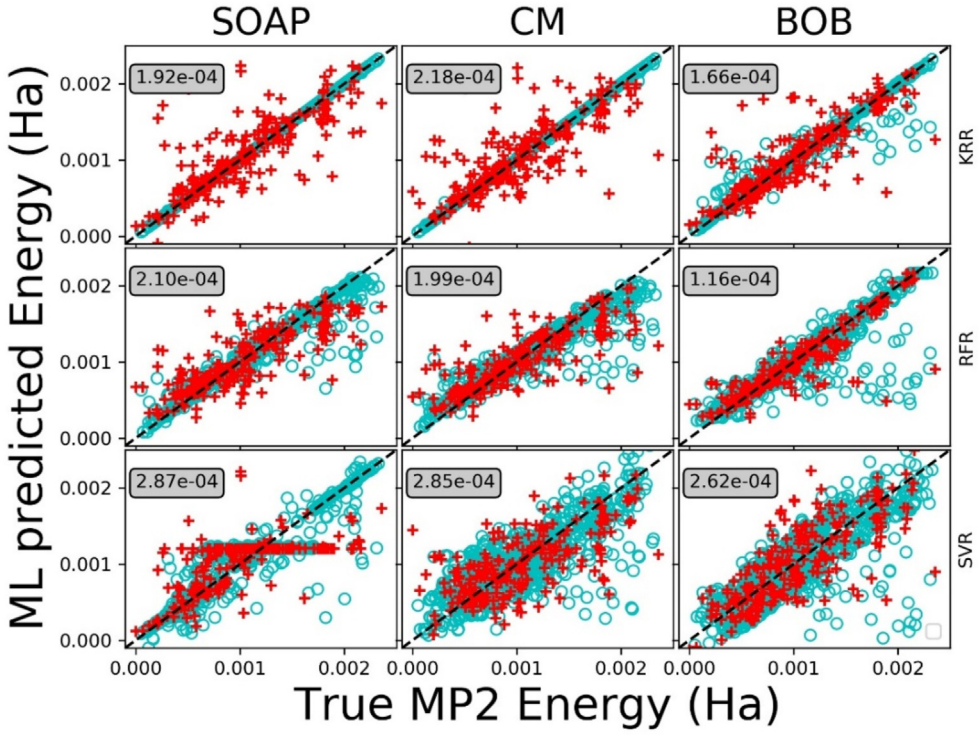


Fig. 7 – Scatter plot for different algorithms and descriptors analyzed: ML predicted energy vs true MP2 energy. Empty cyan circles: training set, Red crosses: test set. Mean absolute errors (in Hartree) are reported in grey boxes. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

with an enhancement of more than 50% for both FGs at P = 5 bar and cryogenic temperature. From the two FGs, -OPO_3H_2 has a greater uptake for all the range of pressures. In the case of the functionalized IRMOF-14 there is a ~20% enhancement for OSO_3H- and more than ~40% for OPO_3H_2-functionalized at P = 5 bar. From the two FGs, -OPO_3H_2 shows greater uptake for all the range of pressures.

In order to prove that ab-initio interaction energies can be learned by different types of machine learning techniques, we performed Supervised Machine Learning on the MP2/def2-TZVPP interaction energies of the optimized structures and show that chemical accuracy ($1k_BT$ (@300 K) ~ 10^{-4} Ha/25 meV/2.5 kJmol^{-1}) could be achieved with as few as 1000 entries. This indicates that the interaction energy between the H_2 molecule and the functionalized benzene in MOFs is an appropriate quantity for ML investigation. Analogous ML simulations had been carried out only for metal hydrides in a dataset of hydrogen storage materials provided by the US Department of Energy [33, 34]. Their ML simulations suggested that the complex material class, followed by Mg is applicable for a wide range of operating temperatures.

Our ML results are summarized in Fig. 7. The prediction results for both the training set (cyan empty circles) and the test set (red crosses) are depicted. The test set is a subset of the full database (20%) and comprises of randomly selected instances, unseen to the training procedure. The columns correspond to the three different representations and the rows correspond to the three different algorithms implemented in this study. The Mean Absolute Error (MAE) on the test set is listed in the grey box of each subplot. In Fig. 8, the learning curves for KRR combined with the different descriptors are represented. Both the MAE and the R^2 coefficient are plotted versus the training set size. The dashed black lines correspond to the chemical accuracy (top panel) and $R^2 = 1$ (bottom panel) and represent the deviation from the quantum mechanical calculated values. All ML algorithms show an increase in the learning rate with a larger dataset size, meaning a larger confidence with increased experience, reaching energy prediction accuracy comparable with quantum mechanical methods at around 1000 datapoints. It is therefore

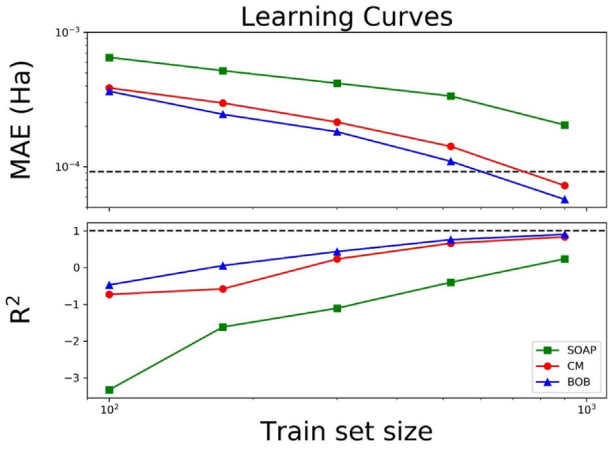


Fig. 8 – Learning Curves for KRR and different descriptors. Log-log plot of MAE (top) and semi-log plot of R^2 (bottom) vs training set size. The dashed line in the top plot represents the chemical accuracy of $1K_BT$ (@300 K) ~ 10^{-4} Ha.

expected that larger databases can further reduce the error on prediction of binding energies for unseen structures.

Conclusions

In this work, we have studied the effect of linker functionalization of MOF structures on the enhancement of their hydrogen uptake capacities, by means of multi-scale molecular simulations. In the first place, we studied with ab-initio methods the binding strength towards hydrogen of 58 functionalized benzenes of the form C_6H_5-X (X: the different FGs). From this set, 37 FGs showed higher binding energies compared to benzene, with 12 of them presenting a binding enhancement over 15%. Notably, $-OSO_3H$ FG stands out with an enhancement of ~80%. In the context of surface modification, the FGs that exhibit an enhancement over 20% are recommended as promising for potential linker functionalization of MOF candidates.

In order to describe accurately the intermolecular interactions between the guest hydrogen molecules and the MOF atoms, we used ab-initio derived results and fitted the parameters of the DREIDING potential onto these results. All of our studied cases indicate that available Force-Fields fail to describe accurately our systems, highlighting the importance of the fitting procedure.

The effect of linker functionalization on the H_2 adsorption uptake of MOFs was studied employing Grand Canonical Monte Carlo simulations at cryogenic and ambient conditions. We functionalized three different MOF candidates from the IRMOF family (IRMOF-8, IRMOF-16 and IRMOF-14) with the FGs that showed strong binding towards hydrogen according to our ab-initio calculations (-$OCONH_2$, $-SO_2NH_2$, $-OSO_3H$ and -OPO_3H_2). A significant enhancement of the volumetric uptake at 77 K was observed for all cases. While structural features clearly influence the packing of H_2 molecules within an adsorbent, of equal importance are the specific binding locations within the pores and the relative binding strength at these sites. Taking into consideration the weight of the FGs as well as other important factors (e.g. pore environment), the adsorption is in agreement with the strength of interaction between the adsorbate (H_2) and the organic parts of the adsorbent (MOF). However, within the scope of practical applications, the absolute uptake percentages illustrate clearly that considerable work is still required in the field.

Our ML results show a fair prediction of the H_2 binding energies even with a limited amount of data. ML can confidently identify low laying structures in the interaction energy landscape, and thus it is possible to extend the search to a larger chemical space. This paves the road for more extensive calculations, able to bypass a large part of the computational effort and extrapolate valuable physical-chemical insights. Possible improvements of our ML model include an increased database, a classification according to charge transfer effects and orbital shapes, classifying whether a minimum is local or global by learning its shape, and insertion of symmetry information into the descriptor.

Overall, in this work a systematic and transferable database is set up and can be used for experimental and further computational and ML studies on the enhancement of hydrogen uptake capacities of porous materials in general (i.e. MOFs, COFs, ZIFs, etc.).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Marco Di Gennaro acknowledges Klanik SA. George Froudakis acknowledges that this research has been co-financed by Toyota Motor Europe NV/SA and by the European Regional Development Fund of the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH – CREATE – INNOVATE (project code: T1EDK-00770).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijhydene.2021.06.021.

REFERENCES

[1] Broom DP, Webb CJ, Fanourgakis GS, Froudakis GE, Trikalitis PN, Hirscher M. Concepts for improving hydrogen storage in nanoporous materials. Int J Hydrogen Energy 2019;44:7768–79. https://doi.org/10.1016/j.ijhydene.2019.01.224.

[2] Engineering-ToolBox. Fossil and alternative fuels - energy content. 2008. https://www.engineeringtoolbox.com/fossil-fuels-energy-content-d_1298.html. [Accessed 18 September 2020].

[3] Toyota Mirai Fuel cell. https://www.toyota.com/content/ebrochure/2018/mirai_FuelCellTech.pdf.

[4] DOE technical targets for onboard hydrogen storage for light-duty vehicles. https://www.energy.gov/eere/fuelcells/doe-technical-targets-onboard-hydrogen-storage-light-duty-vehicles.

[5] Eberle U, Felderhoff M, Schüth F. Chemical and physical solutions for hydrogen storage. Angew Chem Int Ed 2009;48:6608–30. https://doi.org/10.1002/anie.200806293.

[6] Yaghi OM, Li G, Li H. Selective binding and removal of guests in a microporous metal–organic framework. Nature 1995;378:703–6. https://doi.org/10.1038/378703a0.

[7] Rosi NL, Eckert J, Eddaoudi M, Vodak DT, Kim J, O'Keeffe M, et al. Hydrogen storage in microporous metal-organic frameworks. Science 2003;300(80). 1127 LP – 1129. https://doi.org/10.1126/science.1083440.

[8] Rowsell JLC, Yaghi OM. Strategies for hydrogen storage in metal–organic frameworks. Angew Chem Int Ed 2005;44:4670–9. https://doi.org/10.1002/anie.200462786.

[9] Yu S, Jing G, Li S, Li Z, Ju X. Tuning the hydrogen storage properties of MOF-650: a combined DFT and GCMC simulations study. Int J Hydrogen Energy 2020;45:6757–64. https://doi.org/10.1016/j.ijhydene.2019.12.114.

[10] Ahmed A, Seth S, Purewal J, Wong-Foy AG, Veenstra M, Matzger AJ, et al. Exceptional hydrogen storage achieved by screening nearly half a million metal-organic frameworks. Nat Commun 2019;10. https://doi.org/10.1038/s41467-019-09365-w.
[11] Hübner O, Glöss A, Fichtner M, Klopper W. On the interaction of dihydrogen with aromatic systems. J Phys Chem 2004;108:3019–23. https://doi.org/10.1021/jp031102p.
[12] Xia L, Liu Q, Wang F, Lu J. Improving the hydrogen storage properties of metal-organic framework by functionalization. J Mol Model 2016;22. https://doi.org/10.1007/s00894-016-3129-3.
[13] Klontzas E, Mavrandonakis A, Tylianakis E, Froudakis GE. Improving hydrogen storage capacity of MOF by functionalization of the organic linker with lithium atoms. Nano Lett 2008;8:1572–6. https://doi.org/10.1021/nl072941g.
[14] Mulfort KL, Farha OK, Stern CL, Sarjeant AA, Hupp JT. Post-synthesis alkoxide formation within metal-organic framework materials: a strategy for incorporating highly coordinatively unsaturated metal ions. J Am Chem Soc 2009;131:3866–8. https://doi.org/10.1021/ja809954r.
[15] Mavrandonakis A, Klontzas E, Tylianakis E, Froudakis GE. Enhancement of hydrogen adsorption in metal-organic frameworks by the incorporation of the sulfonate group and Li cations. A multiscale computational study. J Am Chem Soc 2009;131:13410–4. https://doi.org/10.1021/ja9043888.
[16] Weigend F, Häser M, Patzelt H, Ahlrichs R. RI-MP2: optimized auxiliary basis sets and demonstration of efficiency. Chem Phys Lett 1998;294:143–52. https://doi.org/10.1016/S0009-2614(98)00862-8.
[17] Balasubramani SG, Chen GP, Coriani S, Diedenhofen M, Frank MS, Franzke YJ, et al. TURBOMOLE: modular program suite for ab initio quantum-chemical and condensed-matter simulations. J Chem Phys 2020;152:184107. https://doi.org/10.1063/5.0004635.
[18] Boys SF, Bernardi F. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. Mol Phys 1970;19:553–66. https://doi.org/10.1080/00268977000101561.
[19] Eddaoudi M, Kim J, Rosi N, Vodak D, Wachter J, Keeffe M, et al. Systematic design of pore size and functionality in isoreticular MOFs and their application in methane storage. Science 2002;295(80). https://doi.org/10.1126/science.1067208. 469 LP – 472.
[20] Groom CR, Bruno IJ, Lightfoot MP, Ward SC. The Cambridge structural database. Acta Crystallogr B 2016;72:171–9. https://doi.org/10.1107/S2052520616003954.
[21] Mayo SL, Olafson BD, Goddard WA. DREIDING: a generic force field for molecular simulations. J Phys Chem 1990;94:8897–909. https://doi.org/10.1021/j100389a010.
[22] Breneman CM, Wiberg KB. Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. J Comput Chem 1990;11:361–73. https://doi.org/10.1002/jcc.540110311.
[23] Darkrim F, Levesque D. Monte Carlo simulations of hydrogen adsorption in single-walled carbon nanotubes. J Chem Phys 1998;109:4981–4. https://doi.org/10.1063/1.477109.
[24] Dubbeldam D, Calero S, Ellis DE, Snurr RQ. RASPA: molecular simulation software for adsorption and diffusion in flexible nanoporous materials. Mol Simulat 2016;42:81–101. https://doi.org/10.1080/08927022.2015.1010082.
[25] Adams DJ. On the use of the Ewald summation in computer simulation. J Chem Phys 1983;78:2585–90. https://doi.org/10.1063/1.445014.
[26] Feynman RP, Hibbs AR. Quantum mechanics and path integrals. New York: McGraw-Hill; 1965.
[27] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. J Mach Learn Res 2011;12:2825–30.
[28] Rupp M, Tkatchenko A, Müller K-R, von Lilienfeld OA. Fast and accurate modeling of molecular atomization energies with machine learning. Phys Rev Lett 2012;108:58301. https://doi.org/10.1103/PhysRevLett.108.058301.
[29] Huang B, von Lilienfeld OA. Communication: understanding molecular representations in machine learning: the role of uniqueness and target similarity. J Chem Phys 2016;145:161102. https://doi.org/10.1063/1.4964627.
[30] De S, Bartók AP, Csányi G, Ceriotti M. Comparing molecules and solids across structural and alchemical space. Phys Chem Chem Phys 2016;18:13754–69. https://doi.org/10.1039/C6CP00415F.
[31] Himanen L, Jäger MOJ, Morooka EV, Federici Canova F, Ranawat YS, Gao DZ, et al. DScribe: library of descriptors for machine learning in materials science. Comput Phys Commun 2020;247:106949. https://doi.org/10.1016/j.cpc.2019.106949.
[32] Christensen AS, Faber FA, Huang B, Bratholm LA, Tkatchenko A, Müller K-R, et al. qmlcode/qml: Releasevol. 3; 2017. https://doi.org/10.5281/zenodo.817332.
[33] Rahnama A, Zepon G, Sridhar S. Machine learning based prediction of metal hydrides for hydrogen storage, part I: prediction of hydrogen weight percent. Int J Hydrogen Energy 2019;44:7337–44. https://doi.org/10.1016/j.ijhydene.2019.01.261.
[34] Rahnama A, Zepon G, Sridhar S. Machine learning based prediction of metal hydrides for hydrogen storage, part II: prediction of material class. Int J Hydrogen Energy 2019;44:7345–53. https://doi.org/10.1016/j.ijhydene.2019.01.264.