

Development and Application of a Single Neural Network Potential for IRMOF-n (n=1,4,6,7,10)

*Omer Tayfuroglu**, *Abdulkadir Kocak*, *Yunus Zorlu*

Department of Chemistry, Gebze Technical University, 41400 Gebze, Kocaeli, Turkey

ABSTRACT

Metal-organic frameworks (MOFs) with their exceptional porous and organized structures have been subject of numerous applications. Predicting macroscopic properties from atomistic simulations require the most accurate force fields, which is still a major problem due to MOFs' hybrid structures governed by covalent, ionic and dispersion forces. Application of ab-initio molecular dynamics to such large periodic systems are thus beyond the current computational power. Therefore, alternative strategies must be developed to reduce computational cost without losing reliability. In this work, we describe the construction of a neural network potential (NNP) for IRMOF-n series (n=1,4,7,10) trained by PBE-D4/def2-TZVP reference data of MOF fragments. We validated the resulting NNP on both fragments and bulk MOF structures by prediction of properties such as equilibrium lattice constants, phonon density of states and linker orientation. The energy and force RMSE values for the fragments are only 0.0017 eV/atom and

* Corresponding Author

E-mail: otayfuroglu@gtu.edu.tr Phone:+902626053083

0.15 eV/Å, respectively. The NNP predicted equilibrium lattice constants of bulk structures, which are not included in training, are off by only 0.2-2.4% from experimental results. Moreover, our fragment trained NNP greatly predicts phenylene ring torsional energy barrier, equilibrium bond distances and vibrational density of states of bulk MOFs. Furthermore, NNP allows us to investigate unusual behaviors of selected MOFs such as the thermal expansion properties and the effect of mechanical strain on the adsorption of hydrogen and methane molecules. The NNP based molecular dynamics (MD) simulations suggest the IRMOF-4 and IRMOF-7 to have positive-to-negative thermal expansion coefficients while the rest to have only negative thermal expansion under the studied temperatures of 200 K to 400 K. The deformation of bulk structure by reduction of unit cell volume has shown to increase volumetric methane uptake in IRMOF-1 but decrease in IRMOF-7 due to the steric hindrance.

INTRODUCTION

Metal-organic frameworks (MOFs) with their highly ordered microscopic structure have attracted scientific interest due to their large surface areas and pore volumes, high flexibilities, and capabilities of reversible processes. For the past decade, they have been the subject of numerous applications ranging from energy storage materials to industrial gas adsorption [1–5], drug delivery [6–10], separation [11–14], and catalysis [15–17]. Quick and easy modification of metal sites and organic linkers allows rational design and versatile functionalization.

Most of the applications require well defined process conditions such as pressure and temperature. For instance, thermal expansion due to higher temperatures might lead to deformation of plasticity in materials [18–20]. In particular, adsorption and desorption processes of guest molecules without heat control system might cause drastic temperature changes in MOF structures. On the other hand, adjusting temperature can be beneficial to controlled adsorption and desorption

guest molecules for some applications such as temperature swing adsorption (TSA) [21,22]. Thus, thermal properties of MOFs are still interest of investigation. The most interesting thermal property is that most frameworks show negative thermal expansion (NTE), the decrease in the total volume upon heating. This property becomes critical and can be utilized when designing on-board energy related (CH_4 and H_2) storage materials [18]. It also creates thermal pressure that plays important role in the mechanical stability of the frameworks.

Adsorption and desorption properties have been extensively studied by both experimental and theoretical methods [23–26]. However, there are few theoretical studies on mechanical behaviors of MOFs such as applied temperature and applied pressure [27]. Computational studies on adsorption/desorption properties as well as thermal behaviors of MOFs rely on classical force fields. In order to understand the atomistic details of host-guest interaction, the reactivity and flexibility of MOFs, more accurate methods such as Density Functional Theory (DFT) are required. Therefore, alternative strategies must be introduced to caveat the problem.

The bottleneck with DFT methods is that MOF structures may include thousands of atoms in unit cells and first principles definition of such large systems are quite costly in terms of computation [28]. In addition, the existence of multiple transition metals with open shell structures (in inorganic subunits) further complicates the calculations. Moreover, dynamic definition of physical and chemical events requires extensive sampling of the system with MOFs and/or guest-molecules. Application of ab-initio molecular dynamics to such systems are thus beyond the current computational power. Therefore, alternative strategies must be developed to reduce computational cost without losing reliability.

Over the last decade, Machine Learning (ML) has been greatly applied in many different systems [28–34]. Highly accurate and reactive ML potentials can be built to describe MOF systems with

both organic and inorganic subunits for simulations. Rather than building electronic structure from the first principles at every time step, their great capabilities of learning of multidimensional potential energy surfaces (PES) allows one to compute much faster simulations [32,35–40].

Behler and coworkers were first to construct neural network potentials (NNPs) for MOF-5 (IRMOF-1) and successfully reproduced the electronic structure calculation [28]. They used the idea of fragmentation subsets of bulk MOF to train PES at the DFT level. The DFT reference energies belonging to non-equilibrium geometries of MOF-5 fragments have been used to train NNPs. Although they only used the non-equilibrium fragments rather than using bulk structures in the construction of NNPs, they could accurately predict bulk properties such as lattice parameters, elastic constants and negative thermal expansion. This study showed the applicability of NNPs on MOF like organic-inorganic hybrid systems. Furthermore, the NNPs were constructed by affordable DFT calculations of fragments rather than costly bulk structures [28].

In the construction of potentials, there are several artificial neural network (ANN) models that are mostly based on atomic contribution to the target property (e.g. potential energy surface) such as ANI-torch [28], HDNN [43], deep tensor neural network (DTNN) [28] and SchNet [32].

In this work, we describe the construction and validation of a generic neural network potential for IRMOF-n series (n=1,4,7,10) based on DFT reference data and predict bulk properties such as equilibrium lattice constants, phonon density of states and linker orientation. Our developed NNP enables to reveal unusual behaviors of MOFs such as the thermal expansion properties and the effect of mechanical strain on the adsorption of hydrogen and methane molecules for the selected MOFs.

MATERIAL AND METHODS

Our NNP has been constructed via fragmentation of MOFs with an iterative process as illustrated in **Figure 1**. In the construction process, we used SchNet [32], which is an end-to-end machine learning model (open-source code is available in SchNetPack toolbox [43]). SchNet is a continuous-filter convolutional neural network model designed to predict molecular properties such as energies and atomic forces etc. modified version of deep tensor neural network (DTNN) framework [44]. It is a multilayer neural network consisting of atom embedding, interaction blocks and prediction blocks. At the embedding layer, each atom type is mapped to feature vectors $x_i^l = \mathbb{R}^F$ to create a randomly initialized representation. Here, F denotes the feature dimension and l is the number of layers for i 'th atom. This representation is refined during training by the interaction block. The building block computes pairwise interactions between all x_i^l within a certain cutoff and then refines atomic representations by a subsequent t steps up to T times in total. These atom-wise representations are processed to obtain the desired molecular property by the summation of all atomic contributions in the prediction block. In addition to calculating energies by reliable neural network (NN) models, SchNet also provides the force models obtained by differentiation of energy with respect to atomic positions. Rotationally and translationally invariant description of energy model allows conservation of both force field and energy. Details of the algorithm can be found in Schutt's works [32,43,44]

Herein, we introduce a generic SchNet model for IRMOF-n (n=1, 4, 6, 7, 10) series using DFT energies and forces based on fragments of the MOFs. According to hyper-parametrization on metal organic hybrid molecular systems, 96 number of feature maps for atom embedding, a radial cutoff of 6.0 Å and T=3 interaction refinements have been used. The model was trained with a learning rate of 10^{-4} using ADAM optimizer. Among the dataset, 80% was used for training while 10%

was used for the validation and the remaining 10% was reserved for testing. All training processes have been performed using a Tesla P100 GPU.

The DFT energies and forces for all non-equilibrium geometries were calculated using the ORCA software package (version 4.2) [45] at the PBE/def2-TZVP level for all atoms. SCF calculations were accelerated with the RIJDX option available in the software to notably increase computational efficiency [46]. Tight SCF criteria were applied for all calculations and long range interactions were corrected by the use of density dependent atom-pairwise dispersion (D4) functions [47].

A custom-built python code was prepared using the Atomic Simulation Environment (ASE) [48] framework to perform calculation and store all geometries and their DFT results in an ASE database format. UCSF Chimera [49] and VESTA [50] software packages were used for visualization of 3D structures.

The simulations of fragments used to train NNP were carried out using SchNetPack software with the standard velocity Verlet integrator [43]. The time constant for the simulations was set to 0.5 fs, and Langevin thermostat was used. The final simulations of periodic MOF structures were performed to investigate bulk properties for phonon modes, thermal mismatch and mechanical stress using SchNetPack and ASE interchangeably. Phonons were produced by calculation of vibrational modes using finite differences as implemented in ASE. For thermal studies, MD simulations at different temperatures were performed ranging from 200 K to 400 K by 50 K increments. For each temperature value, the energy of the periodic bulk system was minimized. After minimization, MD simulations of 500 ps with 0.5 fs time steps (total of 1 million steps) were carried out using NPT ensemble at 1 atm reference pressure. Langevin thermostat along with

Nosé-Hoover barostats were used. Damping constants for both thermostats were set to 0.1 ps. Last 300 ps of the simulations was used for the analysis.

Thermal expansion coefficient (α_0), bulk modulus (β_0) and thermal pressure coefficients (γ) are the main parameters that can be calculated for the thermal behavior [18].

Thermal expansion coefficient (α_0) is defined as:

$$\alpha_0 = \frac{1}{V} \left(\frac{\partial V}{\partial T} \right)_P$$

This can be re-written by integrating for volume and temperature;

$$\begin{aligned} \alpha_0 dT &= \frac{1}{V} dV \\ \int_0^T \alpha_0 dT &= \left[\int_{V_0}^V \frac{1}{V} dV \right]_P \\ \langle \ln(V) \rangle_P &\approx \alpha_0 T + V_0 \end{aligned} \quad (1)$$

where V_0 is the equilibrium volume at 0 K. By running several MD simulations with different temperatures, α_0 can be calculated from the linear correlation between temperature and natural logarithm of volume expansion. The term $\langle \ln(V) \rangle_P$ can be obtained from the ensemble average of an MD simulation at the NPT ensemble.

Thermal pressure is given by:

$$\Delta P = \gamma \cdot \Delta T$$

where γ is the thermal pressure coefficient with:

$$\gamma = \alpha_0 \cdot \beta_0 \quad (2)$$

where β_0 is the bulk modulus and defined as:

$$\beta_0 = k_B \langle T \rangle \frac{\langle V \rangle}{\langle V^2 \rangle - \langle V \rangle^2} \quad (3)$$

where k_B is the Boltzmann coefficient. The volumetric term in the equation corresponds to the relative variance, the fluctuation of the volume throughout the MD simulation at T=300 K.

For the mechanical behavior of MOFs, the uniaxial compression systems were introduced by assigning the [001] plane to a vector and its gradual decrease. In line with literature studies, the $-3 \times 10^9 \text{ s}^{-1}$ strain rate was engineered as follows: During the MD simulations at the NVT ensemble, the box volume was reduced by decreasing the [001] plane by 0.03% after each ps simulation. A total 45% volume decrease in 150 ps total simulation time was achieved.

Adsorption properties of methane and hydrogen gases were studied using Grand Canonical Monte Carlo (GCMC) simulations with RASPA software code [51] using deformed/undeformed structures of IRMOF-1 and IRMOF-7 after extracting MD trajectories.

For both H₂ and CH₄ uptake, GCMC simulations for all deformed/undeformed structures were run for 1000 cycles with the first 500 being equilibration and the last 500 for the ensemble average. The methane adsorption calculations were carried out at 298 K and 35 bar while for hydrogen we used 77 K and 35 bar.

Using an N₂ probe with 1.86 Å radius, probe occupied accessible volume (POAV) and accessible surface area (ASA) values of deformed/undeformed MOFs were computed using Zeo++ software [52,53].

The $2 \times 2 \times 2$ supercell of MOFs was chosen for all studies except for the thermal expansion behaviors in which $1 \times 1 \times 1$ was used to reduce computational cost.

RESULTS AND DISCUSSIONS

This paper is organized in the following sections: First, we construct a single NNP for IRMOF-n series ($n=1,4,6,7,10$). Next, we test the NNP for both fragment and periodic bulk MOF structures. Finally, we study the application of this NNP for thermal and mechanistic behavior of selected MOFs.

Neural Network Potentials

Generation of training data set and construction of NNP for IRMOF-n ($n=1,4,6,7,10$)

Developing the NNP is an iterative process as outlined in **Figure 1**. Starting with an initial dataset of random geometries, primitive potentials, NNP1 and NNP2, differed by only in architecture, were constructed by DFT calculated energies and forces. Then, using MD simulations of fragments, new geometry configurations were added to the training datasets to better represent configurational space. This procedure is repeated until the sufficient quality of NNPs were produced (convergence).

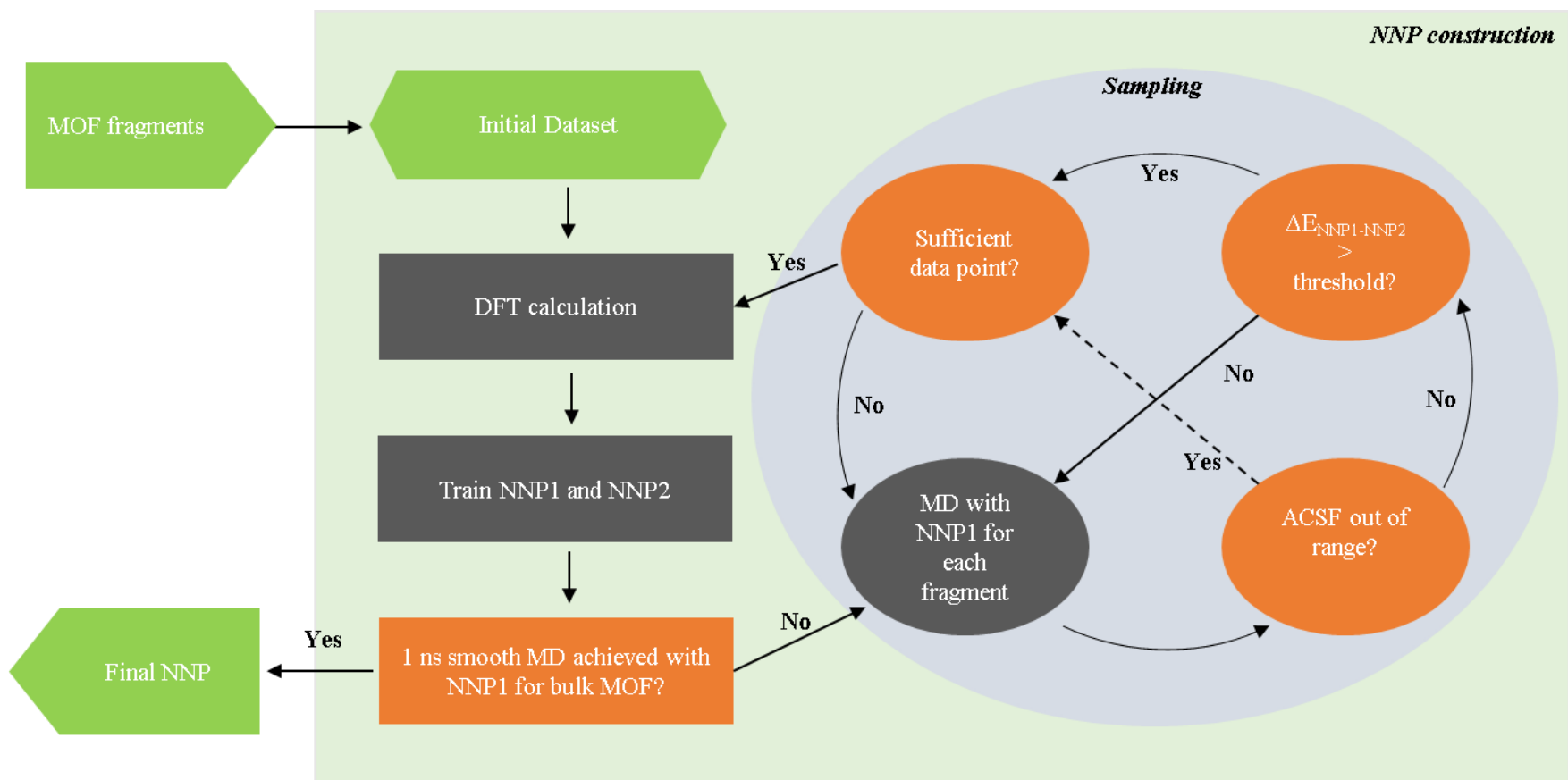


Figure 1. Iterative process for the construction of NNP from MOF fragments based on DFT calculations for IRMOF-n series (n=1,4,6,7, and 10)

For training data set and construction of NNP, calculation of energies and forces of bulk structures at the DFT level is computationally too costly. Therefore, to reduce cost, we used the idea of fragmentation as introduced by Behler et al. for construction and development of the dataset [34].

Since the NN uses atomic contributions, the environment of atoms must be clearly defined within a certain cutoff. Thus, the MOF structure has been divided into pieces (fragments) by finding chemically inequivalent atoms that form the bulk structure. Similar to Behler's work, we generated the different possible fragments so that each fragment includes all different atom types and their neighbor atoms within a sphere of 6 Å cutoff radius. This cutoff distance was also used by the interaction block. In creating MOF fragments, cutting from cyclic structures was avoided. In addition, the missing bonds were saturated with hydrogen to achieve charge balance.

We started with the fragments of IRMOF-1 as it is the base for all other MOFs and thus the fragments of IRMOF-1 are common in all other MOFs studied. We defined IRMOF-1 with 5 fragments. Other MOFs differ from IRMOF-1 by only an additional linker. The organic linker in IRMOF-1 is 1,4-benzenedicarboxylate (BDC) while BDC links with n-propoxy, cyclobutyl and fused benzene functional groups, in IRMOF-4, IRMOF-6 and IRMOF-7, respectively. The linker is 4,4'-biphenyldicarboxylate (BPDC) in IRMOF-10. Therefore, an additional fragment for each of the other IRMOF-n (n=4, 6, 7, and 10) was added to the five fragments of IRMOF-1 with a total of 9 fragments that cover all MOFs studied. The common five fragments for all MOFs and additional differing fragments belonging to each MOFs are present in **Figure S1** (Supplementary Information).

An initial set of 6750 reference fragments were generated by the following way: First, all the bonds of each fragment (total 9 fragments) were scaled from 0.95 to 1.10 with 0.0002 steps size

and 750 x 9 structures with non-equilibrium geometries were obtained. Then, the atomic positions were distorted randomly with a maximum displacement of 0.16 Å in the x, y, and z directions.

On the fly data sampling during the MD simulation

By performing single point energy DFT calculations for these fragments, the energies and forces were obtained in order to create the initial training dataset. The primitive NNP1 and NPP2 models were trained with this initial data set. The training is continued until ensuring that these two NNPs with different architectures produce nearly the same RMSE values on the same test dataset. To develop dataset, MD simulations of all fragments at different temperatures ranging from 50 to 600 K were carried out in the NVT ensemble using trial NNP1 as the energy and force calculator.

Since the configurational space is not adequately represented with initial dataset, missing structures from well-represented regions of configurational space must be added in order to develop the NNPs. Thus, wisely chosen new geometries were added to the training set as reference structures. We applied two different criteria to determine these new reference structures from the trajectories sampled from MD simulations. First, if a structure is not inside the training set, the prediction would be purely based on extrapolation and not reliable. Thus, the first criterion is that the structure must be included if the atom centered symmetry function (ACSF) values are outside the training set range. This criterion was carried out by adopting the idea of symmetry function values reported by Behler [41,54] although we used SchNet rather than HDNN. This is because the symmetry function vectors are translationally and rotationally invariant and it is much easier to compare different structures whether they are inside or outside the training set. Second criterion is that the energy produced by an NNP of different architecture (i.e. NNP2), which is trained by the same dataset, should yield similar result. Therefore, if the energy difference for two independent NNPs ($\text{Uncertainty} = \Delta E_{\text{NNP1-NNP2}}$) is above a certain threshold value, that structure

must be included in the training set. Starting with 0.5 eV, this threshold was gradually lowered to 0.05 eV, as the data set develops. The error will be larger in the NNP trained in the early data set, and this error rate decreases as the data set develops. Uncertainty improvement (i.e. RMSE/MAE) between two NNPs by addition of new configurations to the training set shows an exponential decay function (**Figure S2**). When the new structures based on these criteria have reached to a sufficient number (~5000 data points), these new structures were calculated by DFT and added to the reference set, which eventually created the next generation of NNPs. These processes were repeated several times until smooth NVT MD simulations for 1 ns could be run on the periodic bulk MOF structures.

Finally, ~170k reference data points were obtained when the NNPs were converged. The data were split into training (~126k, 80%) and validation (~17k, 10%) sets while the remainder were reserved for tests (~17k, 10%). The RMSE values on energies and forces were selected as metrics for the losses in training and validation. The loss consisted of weighted average RMSE of 99% forces and 1% energies ($\text{RMSE}_{\text{Loss}} = \text{RMSE}_{\text{energy}} \times 0.01 + \text{RMSE}_{\text{force}} \times 0.99$). Early stopping was carried out when the loss of the training and validation curves reached the plateau (**Figure S3**)

Validation and Test on Fragments

After NNP construction, energies and forces were tested on the fragments. The maximum errors for the energies are 0.036 eV/atom and 0.026 eV/atom for the training and test sets, respectively. **Figure 2** shows the plot of absolute errors between the NNP and DFT energies with respect to DFT energies for the training set and the test sets. The errors on most fragment structures have been estimated to lie within the narrow range of 0.005 eV/atom for both training and test sets. The vertical bar shows the counts of structures for the corresponding error range. Only 1.91% in the training set and 1.58% in the test set is estimated with error greater than 0.005 eV/atom. As the

structures approach to the equilibrium geometries, the errors shrink whereas large errors are concentrated in geometries that are too far from the equilibrium, higher-lying energies on PES. The distribution of the errors is random and fits to a normal Gaussian function (**Figure S4**). Reference DFT energies are in a perfect agreement with the those produced by NNP for all 9 fragments (**Figure S5**, $R^2 \approx 1.000$ for the training set and $R^2 \approx 0.999-1.000$ for the test set). Based on these energy data, the RMSE of the training and test sets is calculated to be 0.0018 and 0.0017 eV/atom, respectively.

In addition to the energies, the maximum NNP force components (F_{\max}) have been tested by means of deviations from the DFT values. **Figure 3** shows the absolute errors between the F_{\max} of NNP and DFT with respect to DFT forces for the training and test sets. The RMSE of these deviations is 0.14 eV/Å and 0.15 eV/Å for train and test datasets, respectively. Among the complete data set, only 1.00% of the training and 1.34% of the test sets have errors that are greater than 0.5 eV/Å. (For the distribution of the errors, refer to **Figure S6**). Reference DFT forces are in a good agreement with the those produced by NNP for all 9 fragments (**Figure S7**, $R^2 \approx 1.000$ for the training set and $R^2 \approx 0.999-1.000$ for the test set).

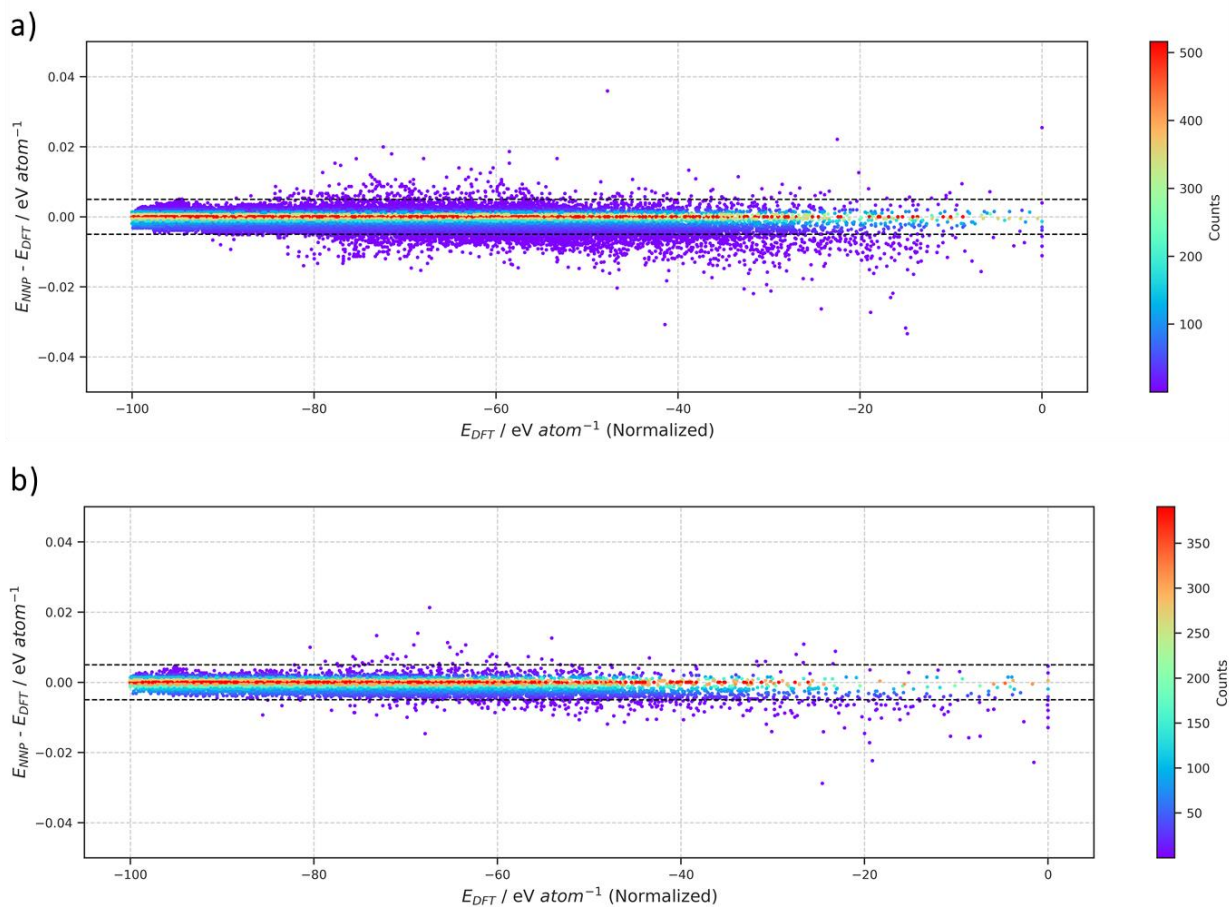


Figure 2. Errors in energies of fragments for a) training and b) test sets. Color bar shows the number of structures with the corresponding errors. Black dashed lines mark errors of 0.005 eV/atom. DFT energies (x-axis) are normalized for a clear representation.

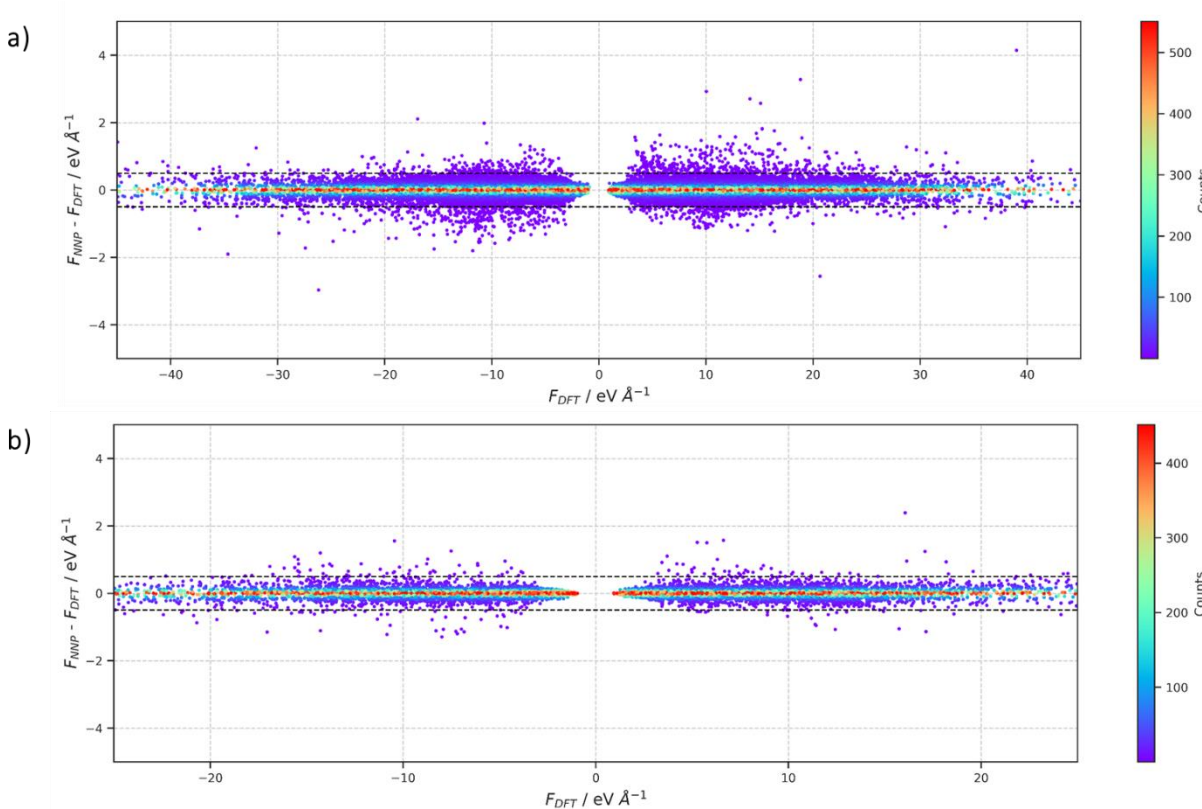


Figure 3. Errors in forces of fragments for a) training and b) test sets. Color bar shows the number of structures with the corresponding errors. Black dashed lines mark errors of $0.5 \text{ eV}/\text{\AA}$.

Performance on Bulk Properties

Apart from testing on fragments, 1 ns NVT MD simulations at 300 K on each periodic bulk IRMOF-1,4,6,7,10 structures were also run following by energy minimization in order to test the stability of NNP. All the configurations visited by the MD simulations were in the well-represented regions of NNP. No structural changes apart from subtle thermal fluctuations were observed during the simulations.

Lattice constants

The equilibrium geometries of fragments in bulk MOFs were not included in the development of NNP in order to avoid memorization of those values. Thus, reproducing the equilibrium lattice

constants from the final NNP is one of the most critical test parameters for the success of NNP. **Table 1** shows the comparison of equilibrium lattice constants produced by NNP of bulk IRMOF structures with experimental single crystal x-ray diffraction (SCXRD) values at 100 K [55] along with literature calculations. Our developed NNP was utilized for carrying out NPT MD simulations at 100 K and 300 K of 500 ps to produce values of equilibrium lattice constants and bulk modulus. The NNP calculated equilibrium lattice constants are in an excellent agreement with experiments and literature calculations, indicative of the reliability of the NNP. The difference between the values of NNP calculated and SCXRD ranges only from 0.04 Å to 0.64 Å for all MOFs trained. It should be stressed that NNP, which has been trained from non-equilibrium geometries of fragments, could successfully predicts the periodic equilibrium parameters.

Table 1. Equilibrium lattice constants and bulk modulus of IRMOF series calculated using NNP and compared with experimental SCXRD [55] and other calculations in literature

IRMOF-n	Equilibrium Lattice Constant (Å)				Bulk Modulus (GPa)		
	SCXRD (100 K)	NNP (100 K)	NNP (300 K)	Simulations	NNP (100 K)	NNP (300 K)	Simulations
1	25.89	26.03	25.95	26.05 ^a (0K); 26.08 ^a (0K); 26.88 ^b (0K); 25.26 ^c (100K); 26.09 ^d (100K); 25.21 ^c (300K); 26.08 ^d (300K)	13.17	8.97	16.0 ^a (0K); 19.0 ^a (0K); 11.95 ^f (0K); 15.76 ^d (0K); 8.70 ^b (0K); 19.37 ^c (100K); 16.66 ^e (300K); 13.3 ^d (300K); 12.3 (300K); 14.34 ^e (300K); 11 ^g (300K)
4	25.85	25.81	25.82	-	12.55	8.26	-
6	25.84	26.18	26.12	27.03 ^b (0K)	14.59	11.80	12.21 ^b (0K); 12.12 ^e (300K)
7	25.83	25.19	25.20	26.69 ^b (0K)	6.58	2.30	3.70 ^b (0K)
10	34.28	34.73	34.62	35.35 ^b (0K); 34.82 ^d (100K); 34.48 ^d (300K); 34.37 ^e (300K)	6.95	5.23	9.2 ^d (0K); 8.25 ^f (0K); 6.00 ^b (0K); 3.5 ^d (300K); 7.41 ^e (300K); 5 ^g (300K)

a: [28]; b:[56]; c:[57]; d:[58]; e:[59]; f:[60]; g:

Energy profiles on the organic linker's rotation in the bulk IRMOF-1

Orientation of organic linkers is a key feature for the structural stability of IRMOF series and accurate calculation of torsional energies could be an important indicator for the success of the NNP [28]. The phenylene rings are common in all MOFs studied here. Therefore, as a case study, we compared NNP and DFT produced rotational energy barrier of a phenylene ring in IRMOF-1 to assess the reliability of the NNP. In order to reduce cost of the DFT calculation, we truncated the IRMOF-1 structure and saturated it with hydrogens (**Figure S8**). NNP and DFT energy profiles were produced by scanning the dihedral angle of one of the phenylene rings. **Figure 4a** shows the energy profile along a 180° rotation. The DFT and NNP calculated values of 0.58 and 0.65 eV, respectively, diverges only by 0.07 eV. This difference might arise from a single NNP representation of all studied MOFs rather than developing an NNP for each of MOFs. In addition, both methods somewhat overestimate the experimental energy barrier obtained by ²H-NMR studies (0.490±0.087 eV) [61]. The overestimation could be due to the use of truncated structure rather than the periodic one.

Similar to phenylene ring torsions, the same truncated structure was also used to reproduce the aromatic C-H bond distance energy profile of the phenylene ring. **Figure 4b** shows the relative energies corresponding the C-H distance scan calculated by DFT and NNP. Although NNP was developed using non-equilibrium geometries, it can perfectly predict the DFT calculated equilibrium bond distance of C-H to be 1.09 Å along the scan coordinate. DFT calculated distance was totally reproduced by NNP. Both methods agree reasonably with the experimental equilibrium aromatic C-H distance of 1.10846 Å by neutron powder diffraction (NPD) measurements [62].

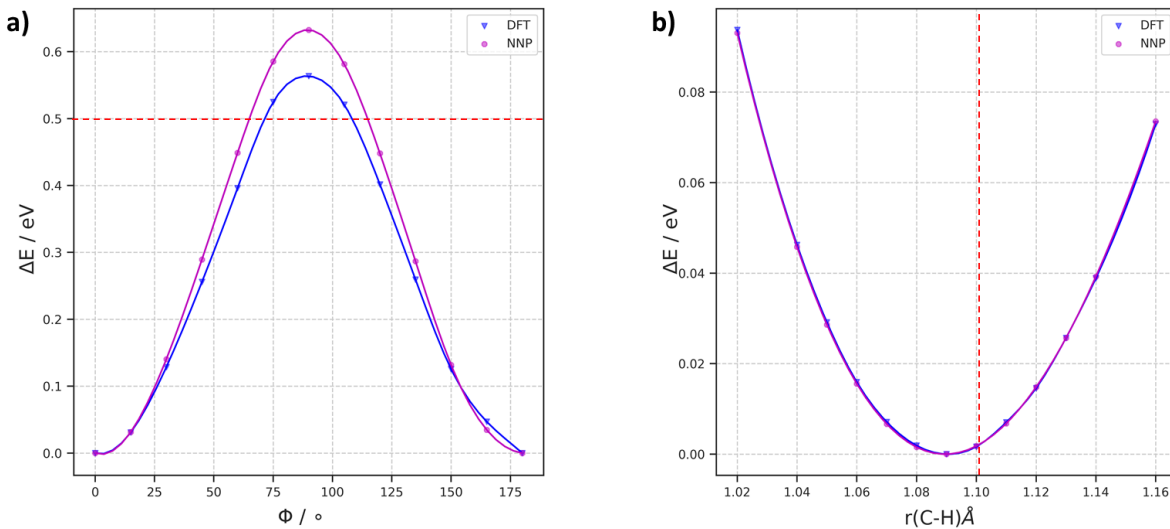


Figure 4. Potential Energy Surface in IRMOF-1 by scanning a): dihedral angle between carboxylate and phenylene planes from 0 to 180° and b) distance of aromatic C-H from 1.02Å to 1.16Å. Dashed lines shows experimental values as NMR for torsional barrier [61] and NPD for equilibrium distance [62].

Phonon modes

Reproducing phonon modes of a periodic bulk structure from fragment trained NNP is another critical test for the reliability of the potential. Calculating phonon modes directly from first principles can become quite costly for such large MOF systems. However, using the NNP, we utilized the calculation of vibrational modes available in ASE with much lower computational cost.

Experimental studies on IRMOF-1 vibrations showed that there are several spectral regions [63]. The dominant vibrations are C-H stretching (due to aromatic C-H on phenylene ring) at 2900-3100 cm^{-1} ; C=O stretching (due to carboxylate linker groups), C=C stretching (due to aromatic C-C on phenylene ring) at 1300-1600 cm^{-1} ; in-plane and out-plane deformation of aromatic ring (due to

the phenylene ring) at 600-1200 cm^{-1} ; the Zn involved bands (due to inorganic subunit Zn-O stretches and bends) at 200-600 cm^{-1} ; and collective vibrations and lattice modes below 200 cm^{-1} [63]. The latter (also reported as below 23 meV) are typical for most MOFs and play important role in negative thermal expansion properties.

IRMOF-1 shows aromatic C-H symmetric and antisymmetric stretching bands at the 3126.4 cm^{-1} (23) and 3138.9 cm^{-1} (46), respectively (**Figure 5** and **Figure S9**). These values are very consistent with previous DFT studies. Phonon modes of simplified model system of IRMOF-1 were reported in literature using two different functionals, B3LYP [64] and PBEsol-D3 [63] as 3230/3243 cm^{-1} and 3110/3122 cm^{-1} . Since our NNP was trained with PBE-D4, our values are in excellent agreement with PBEsol-D3 values. Not only the frequency values, but the intensities and splitting of 12 cm^{-1} between the symmetric and antisymmetric stretches were also totally reproduced by our NNP. The C-H stretches were also reported by HDNNP [28], which was also trained by PBE-D3, as 3104/3148 cm^{-1} . Our NNP calculated C-H stretching frequencies are also in agreement with previously reported experimental values.

IRMOF-10 shows very similar C-H stretching frequencies to IRMOF-1 since all C-H bonds are aromatic (Figure S8). However, due to unequal chemical environments, C-H bonds at the interface of the rings slightly differ from those that are towards carboxylate groups. The C-H stretches in IRMOF-4 are more complicated due to the existence of aliphatic C-H bonds of n-propoxy groups linked in BDC. Similarly, additional aliphatic C-H stretches are observed in IRMOF-6 due to cyclobutyl group linked in BDC. In IRMOF-7, the C-H stretches of fused benzene ring are observed at slightly different frequencies than those on BDC. This difference can be rationalized by the fact that H atoms on the fused benzene are closer to the carboxylate oxygen atoms whereas the H atoms on BDC are more distant from carboxylate groups. Indeed, two isomeric forms as

“twisted” and “planar” can conceivably exist in IRMOF-1 and IRMOF-7 and the stability of one of these forms were dominated by the steric repulsion of the fused benzene ring [65].

The second region of IRMOF-1 consists of the spectrum lying from 1514 cm^{-1} to 1625 cm^{-1} with several peaks including the one at 1612 cm^{-1} and 1595 cm^{-1} , corresponding to symmetric and antisymmetric stretches of aromatic carbon double bonds and carboxylate groups. This is in complete agreement with previously calculated DFT values of 1649 cm^{-1} and 1549 cm^{-1} and experimental IR peaks at 1612 cm^{-1} [55] and Raman peak at 1600 cm^{-1} . In IRMOF-7, these bands are calculated to be red shifted. For IRMOF-7, fused benzene rings’ hydrogens that are proximal to the carboxylate oxygen atoms show more intense C-H bending vibrations.

Experimentally observed Zn-O (the carboxylate oxygen) stretch in IRMOF-1 at 578 cm^{-1} were also reproduced by the peaks in the range 582-612 cm^{-1} . The same mode was reported as 579-606 cm^{-1} in DFT calculations previously. In addition, the Zn-O (inorganic oxygen coordinated by four Zn atoms) is calculated to be 518-523 cm^{-1} . Experimentally, this peak is assigned to be 524 cm^{-1} while the DFT calculations reported the value of 512 cm^{-1} . Finally, Zn involved collective motions were reported to be below 287 cm^{-1} . Our calculations show 293 cm^{-1} and below are the corresponding motions including Zn-Zn vibrations. Since the additional fragments on the BDC linker increase the degree of freedom, more accessible vibrational states exist in lower lying spectral regions.

Overall, our calculated vibrational spectra illustrate a good agreement with those reported experimentally and those predicted by DFT density of state calculations.

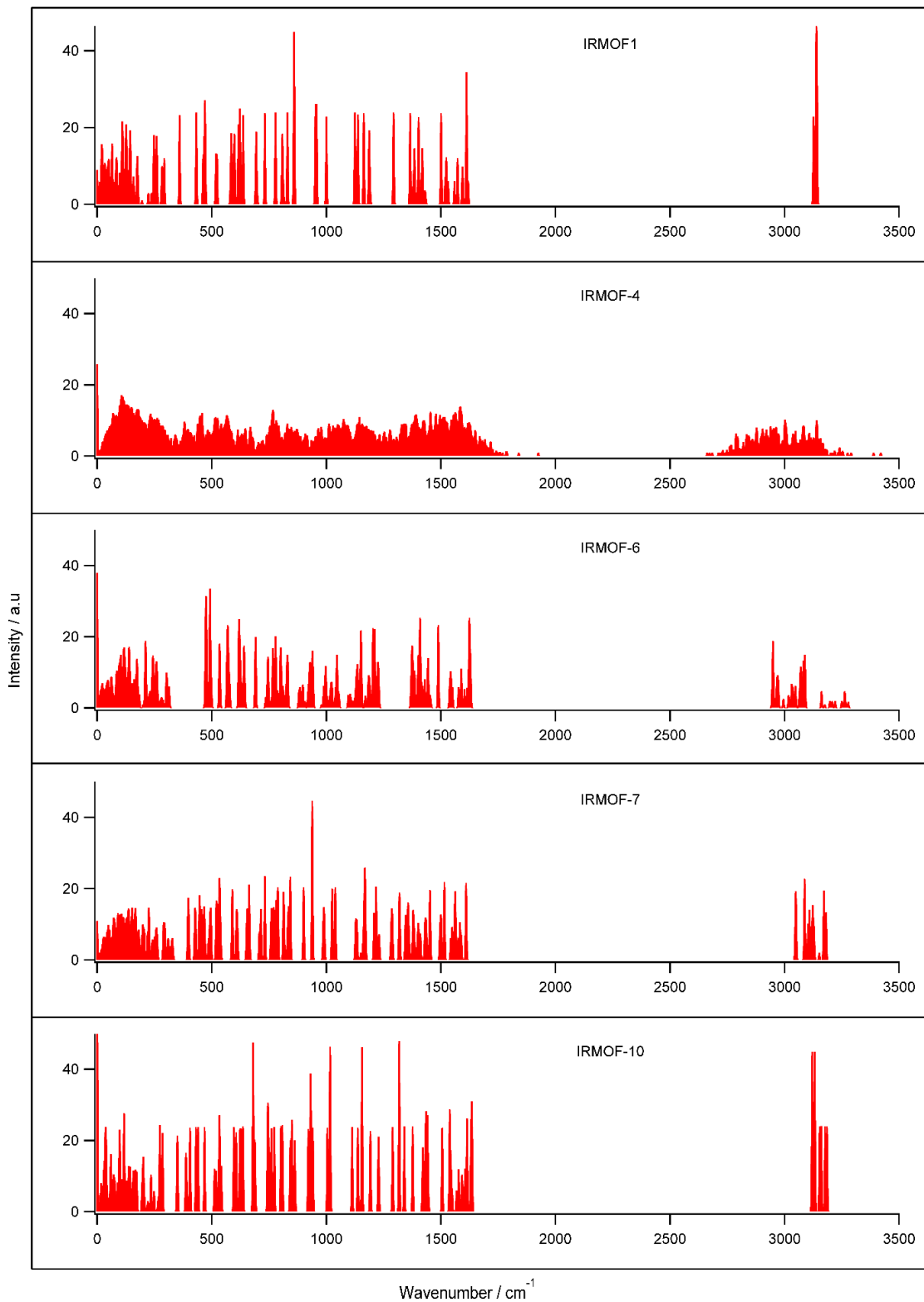


Figure 5. Calculated vibrational spectra of IRMOF-n series by NNP. For zoomed sections refer to **Figure S9**.

Thermal behavior

In terms of thermal behaviors of pure MOF crystals, we have calculated thermal expansion coefficient, thermal pressure coefficient and bulk modulus. Using NNP for energy and force calculations, several MD simulations with temperature ranging from 200 K to 400 K were performed at the NPT ensemble. The mechanical pressure was set to 1 atm, and the data of MD simulations with five different temperature values from 200 to 400 K in steps of 50 K were fitted the equation (1) to find thermal expansion coefficient values. Similarly, thermal pressure coefficients and coefficient factors have been calculated from equations (2) and (3). Table-3 shows the comparison of these parameters calculated using NNP with literature. Using NNP, we find the thermal expansion coefficient values of IRMOF-1 and IRMOF-10 to be 47 MK^{-1} and 54 MK^{-1} , respectively. In literature, there is no consensus about these values. Experiments suggest two values to be -48 [20] and -39 MK^{-1} [66] while predicted values from simulations in literature range from -9 to -79 MK^{-1} [18,57,59,60,67–70]. Our result for IRMOF-1 is very close to the one of the two experimentally reported data of -48 and -39 MK^{-1} . This might suggest the superiority of NNP over the classical simulations. Theoretical calculations on IRMOF-10 have also suggested different values ranging between -16 and -55 MK^{-1} [59,60,67,71], which is consistent with our prediction. To the best of our knowledge, the values of IRMOF-4, IRMOF-6 and IRMOF-7 have not been reported in literature. We predicted the value of IRMOF-6 as -35 MK^{-1} . However, we could not get a linear correlation for IRMOF-4 and IRMOF-7.

The origin of negative thermal expansion properties of MOFs is attributed to a few different scenarios. Shortening the bond lengths and phase changes, electronic effects, magnetorestriction and vibrational modes of bridging atoms. Bridging atoms can lead to high NTE with more than one contribution. First, the transverse vibrational modes on the bridging phenylene rings cause contraction of the Zn_4O nodes. As temperature increases, the contraction gets larger, resulting with

volume shrinking in MOF [28]. Second, the flexible carboxylate groups' vibrational modes lead shortening of the distance between the carboxylate groups and rigid O atoms of Zn₄O [20].

The torsional motion of the phenylene ring may also contribute higher NTE. In particular, IRMOF-7 has fused benzene and IRMOF-4 has n-propoxy linked in BDE. Independent torsion of BDC linker proximate these functional groups in IRMOF-4 and IRMOF-7, resulting with greater energy barrier when compared to IRMOF-1, IRMOF-6, and IRMOF-10. This might explain unusual thermal behavior of the former two MOFs. At low temperatures, due to the greater torsional barrier, IRMOF-4 and IRMOF-6 first expand with temperature, i.e. positive thermal expansion (PTE). After reaching to a thermal energy which can compensate this barrier, the MOFs start contracting with increasing temperature, i.e., negative thermal expansion (NTE). In order to justify this hypothesis, we have scanned the rotation of two of the phenylene rings that are facing to each other in IRMOF-1 and IRMOF-7.

IRMOF-1 and IRMOF-7 structures are similar with some distinct differences. In IRMOF-1 structure, carboxylate groups form two planes that are coplanar to each other (0°). Although these carboxylate planes can be twisted (90°) in principle, they cannot be interconverted without breaking bonds of the network. In addition, the phenylene ring is also coplanar (0°) with these carboxylate groups due to the conjugation between the aromatic system and the carboxylate groups. In addition, the attraction between the proximal H and carboxylate O atoms is dominant in the coplanar carboxylate and phenylene groups since the minimum distance of 2.4 Å is reached. [65]. On the other hand, IRMOF-7 has fused benzene ring attached (ring B) on the phenylene ring (ring A) of the BDC linker (**Figure 6**). The two carboxylate groups were experimentally observed in the twisted (90°) form in IRMOF-7 [55]. For the naphthalene being coplanar with either one of the carboxylate planes would cause steric clashes between the proximal H atoms on the fused ring

(ring B) with the O atoms of carboxylate groups. That is, the distance between proximal H atom on ring B and carboxylate O atom can be as low as 1.7Å, in which repulsions dominates. Thus, the equilibrium geometry of IRMOF-7 has the angle between naphthalene plane and carboxylate groups of 45° in which the distance between proximal H of ring B and carboxylate O is ~2.3Å, very close to the distance between H atom of the phenylene ring (ring A) and carboxylate O in the IRMOF-1 with 2.4 Å. In addition, similar to IRMOF-1, H atom of ring A in IRMOF-7 has a distance of 2.4Å with carboxylate O, which also contributes the stabilization. Steric hindrance is not limited only to the distance between the proximal H of ring B and carboxylate O but also exists between two neighboring rings. The distance between proximal H atoms between two neighboring B rings can be as low as 1.0 Å ($\Psi=135^\circ$) in the case of a torsion of the naphthalene group (**Figure 6**).

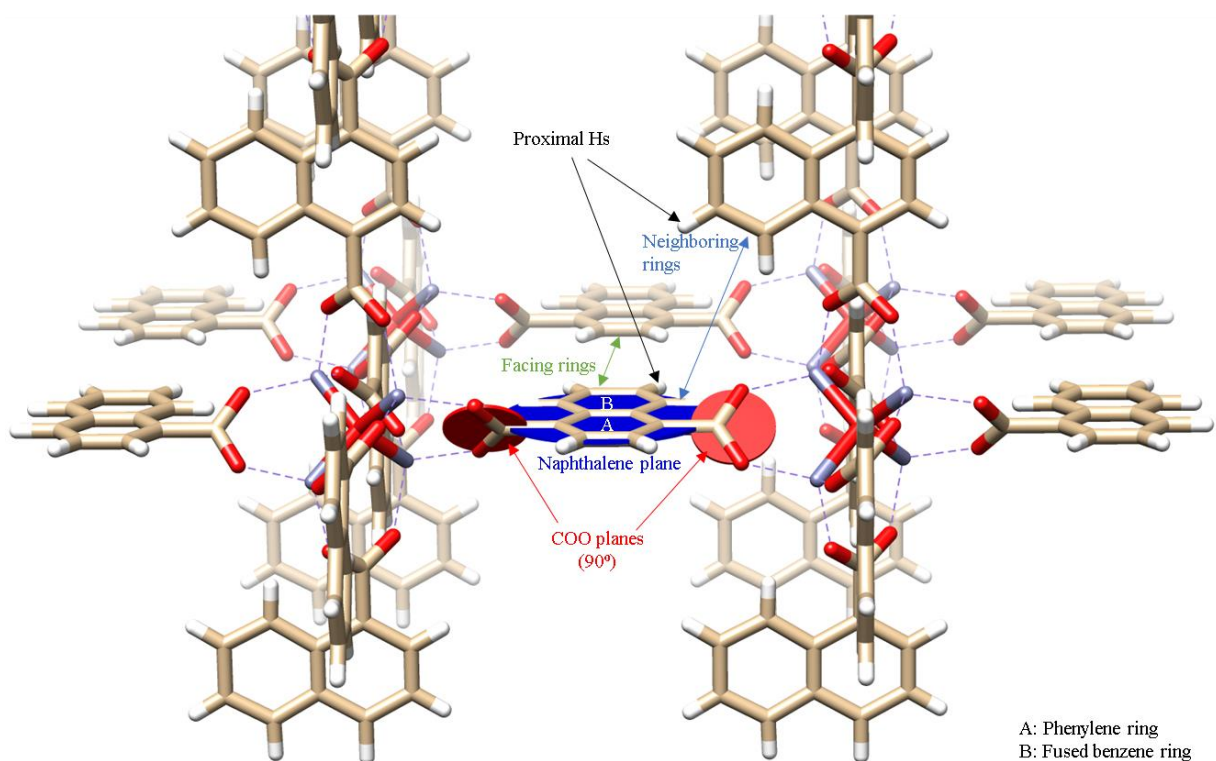


Figure 6. Representation of SCXRD structure of IRMOF-7 at 100 K. COO planes are orthogonal to each other and angled by 45° to the naphthalene plane to optimize interaction between H atoms of two neighboring B rings; and between H atoms of ring A/B and carboxylate O atoms.

In order to monitor the effect of steric hindrance from these two scenarios, a rigid scan has been carried out by changing the angle between the phenylene ring and carboxylate planes on a single unit cell of the MOF using NNP in a two dimensional (2D) scan of two facing rings. **Figure 7** shows the two dimensional potential energy surface (2D-PES) of IRMOF-1 and IRMOF-7.

In the case of IRMOF-1, an independent rotation around two facing phenylene rings could bring atoms to a distance that is no less than 2.4 \AA ($\Phi=\Psi=90^\circ$) with an energy barrier up to $\sim 1.2 \text{ eV}$. For a single dimensional scan, this barrier is almost halved to $\sim 0.6 \text{ eV}$ ($\Phi=0^\circ$, $\Psi=90^\circ$ or vice versa), which is in perfect agreement with experimental and DFT calculated values. On the other hand, due to the neighboring fused benzene rings, the atoms in IRMOF-7 can be as proximate as 1.0 \AA ,

which gives rise to the energy barrier as much as ~ 6 eV ($\Phi=\Psi=135^\circ$). For a single dimensional scan, this barrier is almost halved to ~ 3 eV ($\Phi=45^\circ$, $\Psi=135^\circ$ or vice versa). Additionally, another energy barrier (with double shoulders) of ~ 2.5 eV (translation to 1D is ~ 0.6 eV) has been observed from the fused benzene H and carboxylate repulsions with 1.7 Å distance ($\Phi=45^\circ$, $\Psi=0^\circ$ and 90°).

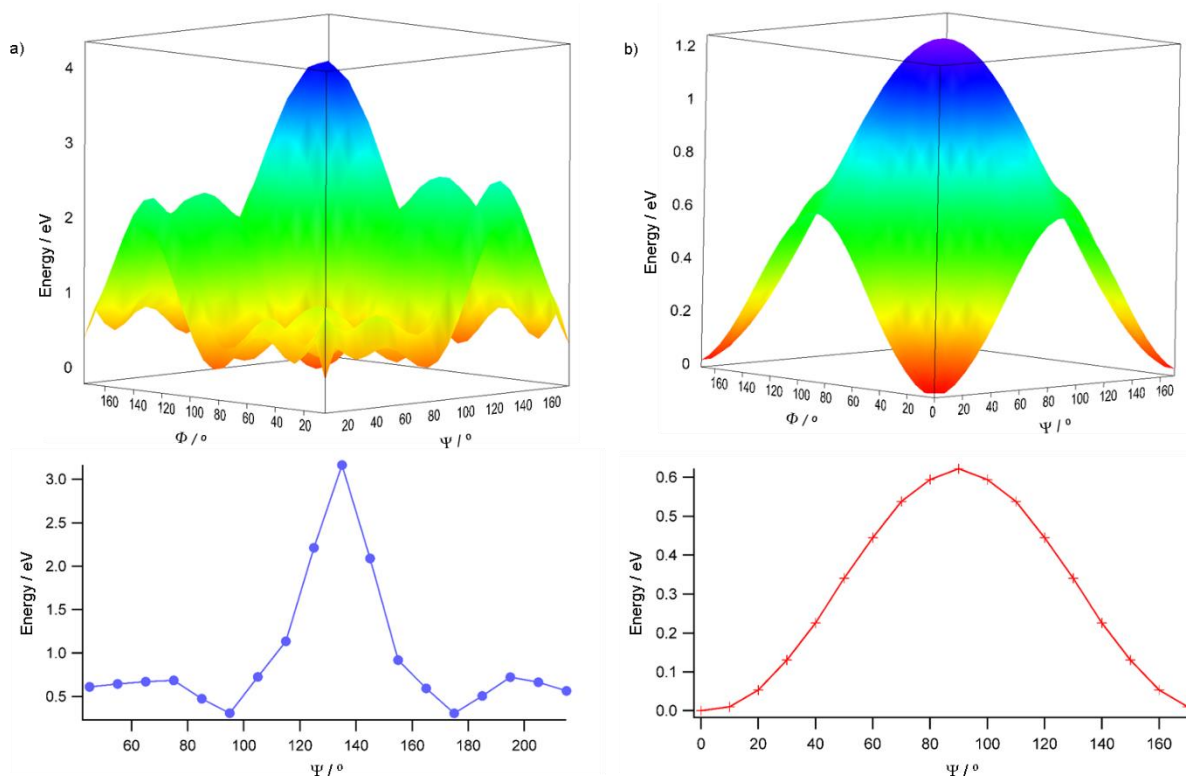


Figure 7. NNP-2D scan (with 10° steps) of rotations of two facing aromatic rings in a) IRMOF-1 and b) IRMOF-7. Φ and Ψ angles correspond to the angle between carboxylate and aromatic planes on each facing rings. The images were produced the data by fitting 12th order polynomial function. Bottom graphs show 1D scan (a portion in the top images) by fixing Φ angles to the initial angles ($\Phi=0^\circ$ for IRMOF-1 and $\Phi=45^\circ$ for IRMOF-7).

Switching from PTE to NTE were first reported in literature and this “switching temperature” were determined as 230 K for DUT-49 [19]. Although experimental observation of this phenomenon required solvated systems, underlying molecular basis can still be investigated for

unsolvated systems by theoretical calculations. It should be noted that the thermal expansion coefficients of MOFs with dual thermal expansion characteristics cannot be simply produced by linear $\ln(\langle V \rangle)$ vs. T plot in equation (2) since α_0 is not constant and thus should be reported as a function of temperature (For the details, refer to the discussion in Supplementary Information). **Table 2** lists thermal expansion coefficient values of IRMOF-n series calculated at different temperatures in comparison with other experimental and computational studies.

Table 2. Thermal expansion coefficient values of IRMOF-n series calculated at different temperatures. For the calculation at different temperatures, refer to Supplementary Information.

IRMOF-n	$\alpha_0 \times 10^{-6}$ (T=200 K)	$\alpha_0 \times 10^{-6}$ (T=300 K)	$\alpha_0 \times 10^{-6}$ (Constant T)	Experimental	Simulations
1	-51.392	-46.876	-46.900	-48[20]; -39 [66]	(-9) – (-79) [18,57,59,60,67–70]
4	159.92	-91.500	-91.506		
6	-31.623	-34.566	-34.566		
7	433.75	-79.030	-79.032		
10	-40.195	-50.233	-50.233		(-16) – (-55) [59,60,67,71]

IRMOF-1,6 and 10 show only NTE behavior while IRMOF-4 and 7 shows PTE to NTE transition. Although the magnitude of the NTE coefficient of IRMOF-10 was not reported, it was shown to be larger than that of IRMOF-1 in literature both experimentally [68] and from simulations [18]. Our calculations agree well with experiments and computations in terms of the magnitude and the decrease of α_0 from IRMOF-10 to IRMOF-1.

After reaching to the NTE region, IRMOF-4 and IRMOF-7 contract more than IRMOF-1 with increasing temperature. These higher NTE values can be explained by greater repulsion of the phenylene rings due to greater energy barrier discussed earlier.

We also calculated thermal pressure coefficients from MD simulations using NNP (**Table S3**). Thermal pressure is an indicator for the rigidity and stability of MOFs. Thermal pressure coefficient, on the other hand, indicates how the pressure would change when a material in a closed system is subjected to temperature increase. Among the IRMOF series studied, IRMOF-10 shows the lower thermal pressure coefficient than IRMOF-1, which is qualitatively in agreement with simulations in literature [18].

Mechanical behavior

In addition to thermal behaviors, after developing NNP and testing it with fragments and bulk materials, we have also investigated elastic response of IRMOFs to applied mechanical pressure. Pressure can also be utilized to control gas storage and release from MOFs. Thus, we have explored methane and hydrogen gas storage properties of IRMOF-1 and IRMOF-7 under applied external pressure. To do that, we have performed MD simulations using our NNP of uniaxial compression by reducing one of the direction vector in several steps. At each step, MD frames were extracted for the calculation of density, accessible surface area (ASA), probe occupied pore volume (POAV) and POAV fraction using 1.86 Å probe (N₂) by Zeo++ [52,53]. As expected, the stress applied increases the density of both MOFs smoothly while reducing POAV fraction. Reducing volume also causes the compression of the organic linkers (**Figure 8**). Due to the steric hindrance of fused benzene rings, the decrease of POAV fraction in IRMOF-7 is predicted to be less than IRMOF1.

The CH₄ uptakes of IRMOF-1 under mechanical compression has been reported by Zheng et al [27]. Using similar protocol, we have also studied the effect of mechanical compression on gas storage ability of IRMOF-1 and IRMOF-7. We have computed the CH₄ and H₂ uptakes of both compressed and uncompressed structures that are retrieved from strain applied MD simulations

using NNP. Our calculations for loading of CH₄ on IRMOF-1 is in a good agreement with literature.

Figure 9 shows the CH₄ and H₂ uptakes of IRMOF-1 and IRMOF-7 against the strain applied on MOFs. Gravimetric methane uptake starts decreasing once the MOF starts deforming. Since the accessible surface area (ASA) and probe occupied accessible volume (POAV) decreases by the bending of organic linkers upon the stress applied, the total gravimetric uptake decreases (Figure S10). The fused benzene rings on IRMOF-7 results in less ASA and POAV.

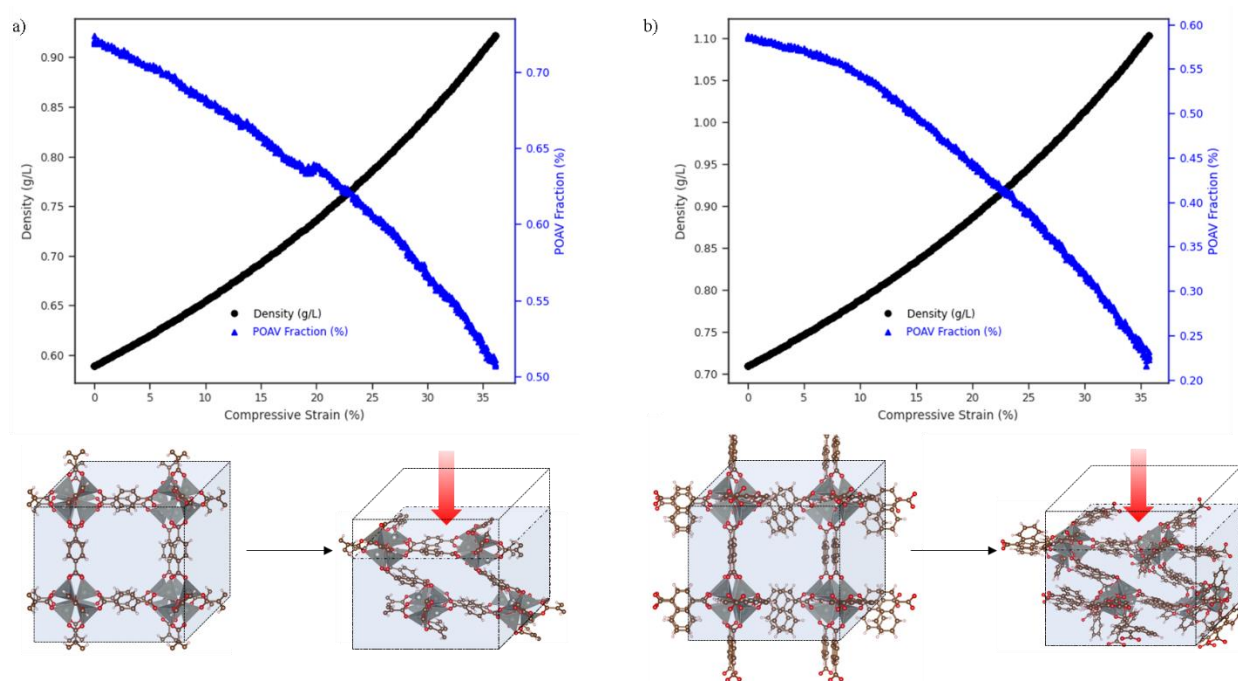


Figure 8. Changes in density and POAV fraction (top) upon compression along [001] direction for the fragments produced by MD simulations using NNP for a) IRMOF-1 and b) IRMOF-7. MOF structures are shown in cartoon (bottom) before (0%) and after (35%) strain applied.

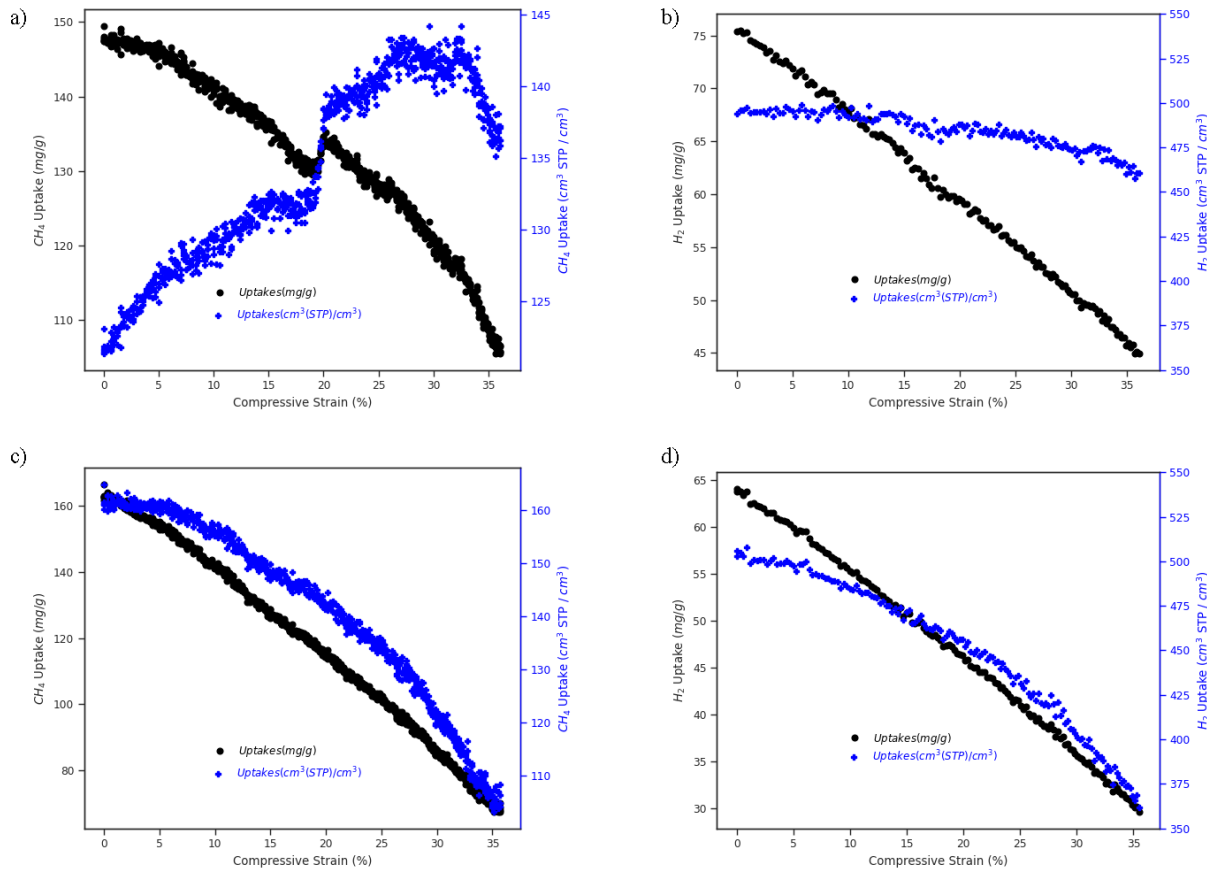


Figure 9 Gravimetric (black) and volumetric (blue) gas uptakes upon compression along [001] direction at each fragments obtained by MD simulations using NNP for a) CH₄ on IRMOF-1; b) H₂ on IRMOF-1; c) CH₄ on IRMOF-7; and d) H₂ on IRMOF-7.

Conclusions

In this work, we have developed neural network potential to describe rotationally and translationally invariant energies and forces of IRMOF-n (n=1,4,6,7,10) in the DFT level using fragmentation technique. We have validated our developed NNP on MOF fragments and bulk MOF structures. Results showed that a single NNP can successfully predict equilibrium bulk properties of IRMOF-n although only five fragments of IRMOF-1 and additional one fragment from each of the other MOFs were used in training. NNP allows us to study bulk properties of

MOFs such as thermal expansion, bulk modulus, mechanical deformation at the DFT accuracy with much reduced cost and extremely acceleration of computation. Thermal expansion studies showed that IRMOF-4 and IRMOF-7 have unusual thermal expansion behavior (as both PTE and NTE) whereas the other MOFs have only negative thermal expansion (NTE) in 200 K – 400 K temperature range. Mechanical strain applied on IRMOFs results with an increase in volumetric CH₄ uptake in IRMOF-1 and decrease in IRMOF-7.

ASSOCIATED CONTENT

Supplementary figures of IRMOF-n (n=1, 4, 6, 7, 10) fragments, uncertainty between NNPs, training and loss curves, histogram of energy errors, DFT and NNP produced energies of fragments, IRMOF-1 truncated structure used in torsional barrier, IRMOF-1 and IRMOF-7 vibrations' zoomed sections, ln(<V>)-T plots of IRMOF-n series, ASA and POAV change by stress applied structures of IRMOF-1 and IRMOF-7.

Supplementary tables of thermal expansion coefficients calculations, equilibrium lattice parameters and thermal pressure coefficients.

AUTHOR INFORMATION

Corresponding Author

Omer Tayfuroglu – Department of Chemistry, Gebze Technical University, 41400 Gebze, Kocaeli, Turkey; E-Mail: otayfuroglu@gtu.edu.tr; orcid.org/0000-0001-7834-3132

Authors

Abdulkadir Kocak – Department of Chemistry, Gebze Technical University, 41400 Gebze, Kocaeli, Turkey; E-Mail: kocak@gtu.edu.tr; orcid.org/0000-0001-6891-6929

Yunus Zorlu – Department of Chemistry, Gebze Technical University, 41400 Gebze, Kocaeli, Turkey; E-Mail: yzorlu@gtu.edu.tr; orcid.org/0000-0003-2811-1872

Notes

The authors declare no competing financial interest.

ACKNOWLEDGEMENTS

O.T. was supported by the Turkish Higher Education Council's 100/2000 Ph.D. fellowship program and TUBITAK BİDEB National Scholarship Program for Ph.D. students (2211-C). The numerical calculations reported in this paper were fully performed at TUBITAK ULAKBİM, High Performance and Grid Computing Center (TRUBA resources)

REFERENCES

- [1] R.E. Morris, P.S. Wheatley, Gas Storage in Nanoporous Materials, *Angew. Chem., Int. Ed.* 47 (2008) 4966.
- [2] O.K. Farha, I. Eryazici, N.C. Jeong, B.G. Hauser, C.E. Wilmer, A.A. Sarjeant, R.Q. Snurr, S.T. Nguyen, A.Ö. Yazaydin, J.T. Hupp, Metal–Organic Framework Materials with Ultrahigh Surface Areas: Is the Sky the Limit?, *J. Am. Chem. Soc.* 134 (2012) 15016.
- [3] D.A. Gómez-Gualdrón, P.Z. Moghadam, J.T. Hupp, O.K. Farha, R.Q. Snurr, Application of Consistency Criteria To Calculate BET Areas of Micro- And Mesoporous Metal–Organic Frameworks, *J. Am. Chem. Soc.* 138 (2016) 215.
- [4] D.A. Gómez-Gualdrón, Y.J. Colon, X. Zhang, T.C. Wang, Y.S. Chen, J.T. Hupp, T. Yildirim, O.K. Farha, J. Zhang, R.Q. Snurr, Evaluating Topologically Diverse Metal–Organic Frameworks for Cryo-Adsorbed Hydrogen Storage, *Energy Environ. Sci.* 9 (2016) 3279.
- [5] J.A. Mason, J. Oktawiec, M.K. Taylor, M.R. Hudson, J. Rodriguez, J.E. Bachman, M.I. Gonzalez, A. Cervellino, A. Guagliardi, C.M. Brown, P.L. Llewellyn, N. Masciocchi, J.R. Long, Methane storage in flexible metal–organic frameworks with intrinsic thermal management, *Nature.* 527 (2015) 357.
- [6] P. Horcajada, T. Chalati, C. Serre, B. Gillet, C. Sebrie, T. Baati, J.F. Eubank, D. Heurtaux, P. Clayette, C. Kreuz, J.-S. Chang, Y.K. Hwang, V. Marsaud, P.-N. Bories, L. Cynober, S. Gil, G. Férey, P. Couvreur, R. Gref, Porous metal–organic-framework nanoscale carriers as a potential platform for drug delivery and imaging, *Nat. Mater.* 9 (2010) 172.

- [7] P. Li, J.A. Modica, A.J. Howarth, E. Vargas L, P.Z. Moghadam, R.Q. Snurr, M. Mrksich, J.T. Hupp, O.K. Farha, Toward Design Rules for Enzyme Immobilization in Hierarchical Mesoporous Metal-Organic Frameworks, *Chem.* 1 (2016) 154.
- [8] T. Faust, Nanomedicine: MOFs deliver, *Nat. Chem.* 7 (2015) 270.
- [9] M.C. Bernini, D. Fairen-Jimenez, M. Pasinetti, A.J. Ramirez-Pastor, R.Q. Snurr, Screening of bio-compatible metal-organic frameworks as potential drug carriers using Monte Carlo simulations, *J. Mater. Chem. B.* 2 (2014) 766.
- [10] C. Orellana-Tavra, E.F. Baxter, T. Tian, T.D. Bennett, N.K.H. Slater, A.K. Cheetham, D. Fairen-Jimenez, Amorphous metal-organic frameworks for drug delivery, *Chem. Commun.* 51 (2015) 13878.
- [11] J.-R. Li, R.J. Kuppler, H.-C. Zhou, Selective gas adsorption and separation in metal-organic frameworks, *Chem. Soc. Rev.* 38 (2009) 1477.
- [12] J.M. Holcroft, K.J. Hartlieb, P.Z. Moghadam, J.G. Bell, G. Barin, D.P. Ferris, E.D. Bloch, M.M. Algaradah, M.S. Nassar, Y.Y. Botros, K.M. Thomas, J.R. Long, R.Q. Snurr, J.F. Stoddart, Carbohydrate-Mediated Purification of Petrochemicals, *J. Am. Chem. Soc.* 137 (2015) 5706.
- [13] J.B. DeCoste, G.W. Peterson, Metal-Organic Frameworks for Air Purification of Toxic Chemicals, *Chem. Rev.* 114 (2014) 5695.
- [14] C. Altintas, G. Avci, H. Daglar, A. Nematı Vesali Azar, S. Velioglu, I. Erucar, S. Keskin, Database for CO₂ Separation Performances of MOFs Based on Computational Materials Screening, *ACS Appl. Mater. Interfaces.* 10 (2018) 17257.

- [15] A. Corma, From Microporous to Mesoporous Molecular Sieve Materials and Their Use in Catalysis, *Chem. Rev.* 97 (1997) 2373.
- [16] J. Liu, L. Chen, H. Cui, J. Zhang, L. Zhang, C.Y. Su, Applications of metal-organic frameworks in heterogeneous supramolecular catalysis, *Chem. Soc. Rev.* 43 (2014) 6011.
- [17] A. Corma, H. García, F.X. Llabrés i Xamena, Engineering Metal Organic Frameworks for Heterogeneous Catalysis, *Chem. Rev.* 110 (2010) 4606.
- [18] J. Wieme, V. Van Speybroeck, Unravelling thermal stress due to thermal expansion mismatch in metal–organic frameworks for methane storage, *J. Mater. Chem. A.* 9 (2021) 4898–4906. <https://doi.org/10.1039/D0TA09462E>.
- [19] B. Garai, V. Bon, A. Efimova, M. Gerlach, I. Senkovska, S. Kaskel, Reversible switching between positive and negative thermal expansion in a metal–organic framework DUT-49, *J. Mater. Chem. A.* 8 (2020) 20420–20428. <https://doi.org/10.1039/D0TA06830F>.
- [20] W. Zhou, H. Wu, T. Yildirim, J.R. Simpson, A.R.H. Walker, Origin of the exceptional negative thermal expansion in metal-organic framework-5 $\text{Zn}_4\text{O}(\text{1,4-benzenedicarboxylate})_3$, *Phys. Rev. B.* 78 (2008) 54114. <https://doi.org/10.1103/PhysRevB.78.054114>.
- [21] C. Schneider, D. Bodesheim, M.G. Ehrenreich, V. Crocellà, J. Mink, R.A. Fischer, K.T. Butler, G. Kieslich, Tuning the Negative Thermal Expansion Behavior of the Metal-Organic Framework $\text{Cu}_3\text{BTC}(2)$ by Retrofitting., *J. Am. Chem. Soc.* 141 (2019) 10504–10509. <https://doi.org/10.1021/jacs.9b04755>.
- [22] I. Grobler, V.J. Smith, P.M. Bhatt, S.A. Herbert, L.J. Barbour, Tunable anisotropic thermal

- expansion of a porous zinc(II) metal-organic framework., *J. Am. Chem. Soc.* 135 (2013) 6411–6414. <https://doi.org/10.1021/ja401671p>.
- [23] T.A. Makal, J.-R. Li, W. Lu, H.-C. Zhou, Methane Storage in Advanced Porous Materials, *Chem. Soc. Rev.* 41 (2012) 7761.
- [24] D. Alezi, Y. Belmabkhout, M. Suyetin, P.M. Bhatt, Ł.J. Weseliński, V. Solovyeva, K. Adil, I. Spanopoulos, P.N. Trikalitis, A.-H. Emwas, MOF Crystal Chemistry Paving the Way to Gas Storage Needs: Aluminum-Based Soc-MOF for CH₄, O₂, and CO₂ Storage, *J. Am. Chem. Soc.* 137 (2015) 13308.
- [25] H. Li, M. Eddaoudi, M. O’Keeffe, O.M. Yaghi, Design and synthesis of an exceptionally stable and highly porous metal-organic framework, *Nature.* 402 (1999) 276.
- [26] H.K. Chae, D.Y. Siberio-Pérez, J. Kim, Y. Go, M. Eddaoudi, A.J. Matzger, M. O’Keeffe, O.M. Yaghi, D. Materials, G. Discovery, A Route to High Surface Area, Porosity and Inclusion of Large Molecules in Crystals, *Nature.* 427 (2004) 523.
- [27] B. Zheng, D. Tian, L. Zhang, X. Zheng, C. Liu, H. Lu, J. Chen, G. Maurin, Q. Shi, Investigation of Methane Adsorption in Strained IRMOF-1, *J. Phys. Chem. C.* 123 (2019) 24592–24597. <https://doi.org/10.1021/acs.jpcc.9b06960>.
- [28] M. Eckhoff, J. Behler, From Molecular Fragments to the Bulk: Development of a Neural Network Potential for MOF-5, *J. Chem. Theory Comput.* 15 (2019) 3793–3809. <https://doi.org/10.1021/acs.jctc.8b01288>.
- [29] M. Rupp, A. Tkatchenko, K.-R. Müller, O.A. von Lilienfeld, Fast and Accurate Modeling of Molecular Atomization Energies with Machine Learning, *Phys. Rev. Lett.* 108 (2012)

58301.

- [30] A.P. Bartók, G. Csányi, Gaussian approximation potentials: A brief tutorial introduction, *Int. J. Quantum Chem.* 115 (2015) 1051.
- [31] A.P. Bartók, M.C. Payne, R. Kondor, G. Csányi, Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons, *Phys. Rev. Lett.* 104 (2010) 136403.
- [32] K.T. Schuett, H.E. Sauceda, P.-J.J. Kindermans, A. Tkatchenko, K.-R. Mueller, K.T. Schütt, H.E. Sauceda, P.-J.J. Kindermans, A. Tkatchenko, K.R. Müller, K.T. Schuett, H.E. Sauceda, P.-J.J. Kindermans, A. Tkatchenko, K.-R. Mueller, K.T. Schütt, H.E. Sauceda, P.-J.J. Kindermans, A. Tkatchenko, K.R. Müller, SchNet - A deep learning architecture for molecules and materials, *J. Chem. Phys.* 148 (2018) 241722. <https://doi.org/10.1063/1.5019779>.
- [33] S. Faraji, S.A. Ghasemi, S. Rostami, R. Rasoulkhani, B. Schaefer, S. Goedecker, M. Amsler, High accuracy and transferability of a neural network potential through charge equilibration for calcium fluoride, *Phys. Rev. B Condens. Matter Mater. Phys.* 95 (2017) 104105.
- [34] J. Behler, Neural network potential-energy surfaces in chemistry: A tool for large-scale simulations, *Phys. Chem. Chem. Phys.* 13 (2011) 17930–17955. <https://doi.org/10.1039/c1cp21668f>.
- [35] J. Behler, First Principles Neural Network Potentials for Reactive Simulations of Large Molecular and Condensed Systems, *Angew. Chemie - Int. Ed.* 56 (2017) 12828–12840.

<https://doi.org/10.1002/anie.201703114>.

- [36] J. Behler, Representing potential energy surfaces by high-dimensional neural network potentials, *J. Phys. Condens. Matter.* 26 (2014) 183001. <https://doi.org/10.1088/0953-8984/26/18/183001>.
- [37] J.S. Smith, O. Isayev, A.E. Roitberg, ANI-1: an extensible neural network potential with DFT accuracy at force field computational cost, *Chem. Sci.* 8 (2017) 3192–3203. <https://doi.org/10.1039/C6SC05720A>.
- [38] N. Artrith, A. Urban, An implementation of artificial neural-network potentials for atomistic materials simulations: Performance for TiO₂, *Comput. Mater. Sci.* 114 (2016) 135–150. <https://doi.org/10.1016/j.commatsci.2015.11.047>.
- [39] E. Kocer, J.K. Mason, H. Erturk, A novel approach to describe chemical environments in high-dimensional neural network potentials, *J. Chem. Phys.* 150 (2019) 1–9. <https://doi.org/10.1063/1.5086167>.
- [40] Y. Zuo, C. Chen, X. Li, Z. Deng, Y. Chen, J. Behler, G. Csányi, A. V. Shapeev, A.P. Thompson, M.A. Wood, S.P. Ong, Performance and Cost Assessment of Machine Learning Interatomic Potentials, *J. Phys. Chem. A.* 124 (2020) 731–745. <https://doi.org/10.1021/acs.jpca.9b08723>.
- [41] J. Behler, Atom-centered symmetry functions for constructing high-dimensional neural network potentials, *J. Chem. Phys.* 134 (2011) 1–13. <https://doi.org/10.1063/1.3553717>.
- [42] N. Artrith, J. Behler, High-dimensional neural network potentials for metal surfaces: A prototype study for copper, *Phys. Rev. B - Condens. Matter Mater. Phys.* 85 (2012) 1–13.

<https://doi.org/10.1103/PhysRevB.85.045439>.

- [43] K.T. Schütt, P. Kessel, M. Gastegger, K.A. Nicoli, A. Tkatchenko, K.R. Müller, SchNetPack: A Deep Learning Toolbox for Atomistic Systems, *J. Chem. Theory Comput.* 15 (2019) 448–455. <https://doi.org/10.1021/acs.jctc.8b00908>.
- [44] K.T. Schütt, F. Arbabzadah, S. Chmiela, K.R. Müller, A. Tkatchenko, Quantum-chemical insights from deep tensor neural networks, *Nat. Commun.* 8 (2017) 13890. <https://doi.org/10.1038/ncomms13890>.
- [45] F. Neese, F. Wennmohs, U. Becker, C. Riplinger, The ORCA quantum chemistry program package, *J. Chem. Phys.* 152 (2020) 224108. <https://doi.org/10.1063/5.0004608>.
- [46] F. Neese, An improvement of the resolution of the identity approximation for the formation of the Coulomb matrix, *J. Comput. Chem.* 24 (2003) 1740–1747. <https://doi.org/https://doi.org/10.1002/jcc.10318>.
- [47] E. Caldeweyher, S. Ehlert, A. Hansen, H. Neugebauer, S. Spicher, C. Bannwarth, S. Grimme, A generally applicable atomic-charge dependent London dispersion correction, *J. Chem. Phys.* 150 (2019) 154122. <https://doi.org/10.1063/1.5090222>.
- [48] A. Hjorth Larsen, J. Jørgen Mortensen, J. Blomqvist, I.E. Castelli, R. Christensen, M. Dułak, J. Friis, M.N. Groves, B. Hammer, C. Hargus, E.D. Hermes, P.C. Jennings, P. Bjerre Jensen, J. Kermode, J.R. Kitchin, E. Leonhard Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. Bergmann Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K.S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng, K.W. Jacobsen, The atomic simulation environment—a Python library for

- working with atoms, *J. Phys. Condens. Matter.* 29 (2017) 273002.
<https://doi.org/10.1088/1361-648x/aa680e>.
- [49] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera--a visualization system for exploratory research and analysis., *J. Comput. Chem.* 25 (2004) 1605–1612. <https://doi.org/10.1002/jcc.20084>.
- [50] K. Momma, F. Izumi, VESTA 3 for three-dimensional visualization of crystal, volumetric and morphology data, *J. Appl. Crystallogr.* 44 (2011) 1272.
- [51] D. Dubbeldam, S. Calero, D.E. Ellis, R.Q. Snurr, RASPA: Molecular Simulation Software for Adsorption and Diffusion in Flexible Nanoporous Materials, *Mol. Simul.* 42 (2016) 81.
- [52] D. Ongari, P.G. Boyd, S. Barthel, M. Witman, M. Haranczyk, B. Smit, Accurate Characterization of the Pore Volume in Microporous Crystalline Materials, *Langmuir.* 33 (2017) 14529.
- [53] T.F. Willems, C.H. Rycroft, M. Kazi, J.C. Meza, M. Haranczyk, Algorithms and tools for high-throughput geometry-based analysis of crystalline porous materials, *Microporous Mesoporous Mater.* 149 (2012) 134.
- [54] J. Behler, Constructing high-dimensional neural network potentials: A tutorial review, *Int. J. Quantum Chem.* 115 (2015) 1032–1050. <https://doi.org/10.1002/qua.24890>.
- [55] M. Eddaoudi, J. Kim, N. Rosi, D. Vodak, J. Wachter, M. O’Keeffe, O.M. Yaghi, Systematic design of pore size and functionality in isorecticular MOFs and their application in methane storage, *Science (80-.).* 295 (2002) 469–472. <https://doi.org/10.1126/science.1067208>.
- [56] A. Kuc, A. Enyashin, G. Seifert, Metal–Organic Frameworks: Structural, Energetic,

- Electronic, and Mechanical Properties, *J. Phys. Chem. B.* 111 (2007) 8179–8186.
<https://doi.org/10.1021/jp072085x>.
- [57] S.S. Han, W.A. Goddard, .Metal–Organic Frameworks Provide Large Negative Thermal Expansion Behavior, *J. Phys. Chem. C.* 111 (2007) 15185.
- [58] K. Banlusan, A. Strachan, First-principles study of elastic mechanical responses to applied deformation of metal–organic frameworks, *J. Chem. Phys.* 146 (2017) 184705.
- [59] Y. Sun, H. Sun, An all-atom force field developed for Zn₄O(RCO₂)₆ metal organic frameworks, *J. Mol. Model.* 20 (2014) 2146. <https://doi.org/10.1007/s00894-014-2146-3>.
- [60] J.K. Bristow, D. Tiana, A. Walsh, Transferable Force Field for Metal–Organic Frameworks from First-Principles: BTW-FF, *J. Chem. Theory Comput.* 10 (2014) 4644–4652.
<https://doi.org/10.1021/ct500515h>.
- [61] S.L. Gould, D. Tranchemontagne, O.M. Yaghi, M.A. Garcia-Garibay, Amphidynamic Character of Crystalline MOF-5: Rotational Dynamics of Terephthalate Phenylenes in a Free-Volume, Sterically Unhindered Environment, *J. Am. Chem. Soc.* 130 (2008) 3246.
- [62] T. Yildirim, M.R. Hartman, Direct Observation of Hydrogen Adsorption Sites and Nanocage Formation in Metal-Organic Frameworks, *Phys. Rev. Lett.* 95 (2005) 215504.
<https://doi.org/10.1103/PhysRevLett.95.215504>.
- [63] B. Civalleri, F. Napoli, Y. Noel, C. Roetti, R. Dovesi, Y. Noël, C. Roetti, R. Dovesi, Ab-initio prediction of materials properties with CRYSTAL: MOF-5 as a case study, *CrystEngComm.* 8 (2006) 364. <https://doi.org/10.1039/b603150c>.
- [64] J.K. Bristow, J.M. Skelton, K.L. Svane, A. Walsh, J.D. Gale, A general forcefield for

- accurate phonon properties of metal-organic frameworks, *Phys. Chem. Chem. Phys.* 18 (2016) 29316–29329. <https://doi.org/10.1039/c6cp05106e>.
- [65] S. Amirjalayer, R. Schmid, Conformational isomerism in the isorecticular metal organic framework family: A force field investigation, *J. Phys. Chem. C.* 112 (2008) 14980–14987. <https://doi.org/10.1021/jp8061948>.
- [66] N. Lock, Y. Wu, M. Christensen, L.J. Cameron, V.K. Peterson, A.J. Bridgeman, C.J. Kepert, B.B. Iversen, Elucidating Negative Thermal Expansion in MOF-5, *J. Phys. Chem. C.* 114 (2010) 16181.
- [67] J. Wieme, S. Vandenbrande, A. Lamaire, V. Kapil, L. Vanduyfhuys, V. Van Speybroeck, Thermal Engineering of Metal–Organic Frameworks for Adsorption Applications: A Molecular Simulation Perspective, *ACS Appl. Mater. Interfaces.* 11 (2019) 38697–38707. <https://doi.org/10.1021/acsami.9b12533>.
- [68] D. Dubbeldam, K.S. Walton, D.E. Ellis, R.Q. Snurr, Exceptional Negative Thermal Expansion in Isorecticular Metal–Organic Frameworks, *Angew. Chemie Int. Ed.* 46 (2007) 4496–4499. <https://doi.org/https://doi.org/10.1002/anie.200700218>.
- [69] J.A. Greathouse, M.D. Allendorf, Force Field Validation for Molecular Dynamics Simulations of IRMOF-1 and Other Isorecticular Zinc Carboxylate Coordination Polymers, *J. Phys. Chem. C.* 112 (2008) 5795.
- [70] M.R. Ryder, J. Maul, B. Civalleri, A. Erba, Quasi-Harmonic Lattice Dynamics of a Prototypical Metal–Organic Framework, *Adv. Theory Simulations.* 2 (2019) 1900093. <https://doi.org/https://doi.org/10.1002/adts.201900093>.

- [71] P.G. Boyd, S.M. Moosavi, M. Witman, B. Smit, Force-Field Prediction of Materials Properties in Metal–Organic Frameworks, *J. Phys. Chem. Lett.* 8 (2017) 357.