Carbon Capture Phenomena in Metal-Organic Frameworks with Neural Network Potentials

Yusuf Shaidu,[1,2,*] Alex Smith,[1,2] Eric Taw,[2,3] and Jeffrey B. Neaton[1,2,4,†]

[1] *Department of Physics, University of California Berkeley, Berkeley, California 94720, United States*
[2] *Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States*
[3] *Department of Chemical and Biomolecular Engineering, University of California Berkeley, Berkeley, California 94720, United States*
[4] *Kavli Energy NanoSciences Institute at Berkeley, Berkeley, California 94720, United States*

(Received 30 January 2023; revised 23 March 2023; accepted 7 April 2023; published 10 May 2023)

Diamine-appended metal-organic frameworks are ultraporous materials exhibiting selective and cooperative CO_2 adsorption mechanisms, leading to chemically tunable step-shaped isotherms and isobars that enable a large fraction of their full CO_2 capacity to be captured or released with a modest change in temperature or pressure. Although progress has been made in elucidating carbon capture phenomena in this system, its thermal properties are poorly understood. Here, we develop density-functional-theory-derived neural network potentials for amine-appended Mg_2(dobpdc), metal-organic frameworks made up of Mg cations and dobpdc [dobpdc^{4-} = 4,4'-dihydroxy-(1,1'-biphenyl)-3,3'-dicarboxylic acid] linkers. These potentials are constructed with an active learning approach where their accuracy and transferability are improved through iterative generation of training datasets based on molecular dynamics. Our potentials predict adsorption energy, mechanical properties, vibrational and thermal properties with and without CO_2, reaching *ab initio* accuracy at a fraction of the computational cost. We compute the temperature-dependent heat capacity and lattice thermal expansion of Mg_2(dobpdc), with and without amine appendages and with and without CO_2, quantitatively capturing measured trends where available and explaining their molecular origins. Furthermore, we show that these potentials can be incorporated into a simulated annealing approach to identify starting structures for energy minimization needed to compute accurate binding enthalpies of these and other complex systems. Our density-functional-theory-derived neural network potentials explain the temperature evolution of the structure and properties of an important class of metal-organic frameworks at molecular length scales, while providing a necessary foundation for future studies of the chemical dynamics of carbon capture.

DOI: 10.1103/PRXEnergy.2.023005

I. INTRODUCTION

Energy combustion and industrial processes contribute close to 89% of global carbon dioxide (CO_2) emissions, totaling more than 40 Gt of CO_2 in 2021, the highest levels recorded [1,2]. The continued rise in CO_2 emissions calls for an urgent intervention through the development and implementation of mitigation strategies. One strategy is CO_2 capture from coal-based power plants, natural and synthetic gas processing plants, and cement plants, and subsequent sequestration. This carbon capture and sequestration (CCS) strategy is classified as a postcombustion approach in which the CO_2 is captured from the point source, compressed, transported, and stored beneath the earth's surface [3]. Another broader strategy is direct air capture that involves the removal and subsequent storage of CO_2 from air [4].

In the state-of-the-art postcombustion CCS approach, an aqueous amine solution is used to capture CO_2 from flue gas [5,6]. Similarly, aqueous basic solutions [7] are used to capture CO_2 from air. However, both amine solution adsorbents and aqueous basic solutions possess high adsorbent regeneration energies, low CO_2 cycling capacities, and exhibit thermal and oxidative solvent degradation [4,8], limiting their efficiency. Solid adsorbents, such as carbonaceous materials, silicas, zeolites, and metal-organic frameworks, have been proposed as alternatives to improve the performance of the CO_2 capture process [3].

*yshaidu@berkeley.edu
†jbneaton@berkeley.edu

While a challenge for these solid sorbent materials continues to be synthesizing them at sufficient scales to replace amine solutions, they are highly promising, owing to their high surface areas available for CO_2 adsorption, faster adsorption-desorption kinetics, lower regeneration energies, lower cost, higher stabilities, and tunable chemistry [3,9,10].

Among all solid sorbents explored so far, metal-organic frameworks (MOFs) exhibit the greatest structural and chemical diversity and tunability. MOFs are a class of nanoporous materials with ultrahigh internal surface areas that consist of a periodic array of metal ions connected together by organic molecules, or ligands. In particular, the Mg_2(dobpdc) variant studied here is composed of Mg metal ions connected via a dobpdc^{4-} [dobpdc^{4-} = 4,4'-dihydroxy-(1,1'-biphenyl)-3,3'-dicarboxylic acid] organic molecule. Crucially, the Mg atom is exposed to the interior of the pore and undercoordinated, a so-called open-metal site to which small molecules can readily bind. Mg_2(dobpdc) has a unidirectional hexagonal pore oriented along the *c*-crystallographic axis.

Through a postsynthetic process, diamine molecules can be introduced into the pores and bind at the Mg open-metal sites, and the so-called diamine-appended variants of Mg_2(dobpdc) have been demonstrated as a transformative material for carbon capture applications [11–18]. These diamine-functionalized materials form carbamate species upon exposure to CO_2 gas and exhibit cooperative adsorption mechanisms resulting in step-shaped isotherms that enable nearly the full CO_2 capacity to be accessed with minimal temperature swings [12]. In this reversible cooperative phenomenon, the CO_2 displaces the diamine bound to Mg and reacts with it to form carbamate. This specific reaction leads to high selectivity for CO_2 over H_2O [19] and other small molecules found in flue gas or the atmosphere [20]. Additionally, the diamine-functionalized MOFs have low regeneration heat and are postsynthetically tunable, which enables them to be engineered to capture CO_2 from different CO_2 point sources [11–16,19]. In addition, their mechanical strength [17] and thermal transport properties [18] are predicted to be enhanced compared to the unfunctionalized Mg_2(dobpdc) systems.

Up to now, a few dozen diamines have been considered [13,14,16], and it has been shown that diamines with lower molecular weight have higher CO_2 adsorption capacity but lower thermal stability, leading to deleterious volatilization during temperature-swing cycling. Bulkier diamines with a larger number of degrees of freedom and higher molecular weights have been reported to exhibit more resistance towards volatilization [14], but half of the theoretical CO_2 capacity per diamine can be attained due to steric hindrances along the plane perpendicular to the pore channel that gives rise to spaced double-stepped CO_2 adsorption profiles [13,14].

The realization of new amine-appended MOFs that exhibit both thermal stability and high CO_2 adsorption capacity is at the forefront of research, and recent efforts have been devoted to studying the properties of CO_2 adsorption in tetraamine-functionalized Mg_2(dobpdc) due to their tendency to bridge nearest-neighbor metals across the pore [21]. This binding to multiple-metal sites is expected to yield an adsorbent with enhanced thermal stability over their diamine counterparts. Indeed, tetraamine-functionalized MOFs have been reported to exhibit high thermal stability during temperature-swing cycling with step temperatures up to 160 °C. The step temperature correlates with the adsorption step at low pressures and is sufficient to enable 90% CO_2 capture from natural flue gas streams [22].

Density functional theory (DFT) calculations using van der Waals–corrected functionals have been used to elucidate the equilibrium thermodynamics, the chemical environments before and after CO_2 insertions [11–17,19,21], as well as the mechanical properties of amine-appended MOF materials [17], with and without CO_2. Recently, classical molecular dynamics simulations with the DREIDING force field [23] have been used to compute and study trends in thermal conductivity for the Zn_2(dobpdc) variant, including the bare framework crystal, the diamine-appended MOF, and the MOF with interior carbamate structures [18]. Enhanced thermal transport was found upon carbamate formation relative to bare frameworks, although trends with diamine appendages and the origin of the enhancement have yet to be understood and quantified. Thermal conductivity is critically important as it determines the rate of heat dissipation (and therefore the overall amount of heat dissipated) during the adsorption-desorption process. Moreover, lattice thermal expansion, or volume change, in MOFs upon CO_2 insertion, relevant for the long-term durability of these materials, has not been extensively studied or understood at a molecular level. Likewise, studies of CO_2 adsorption kinetics are only just emerging [24] and mostly underexplored in these systems, despite their importance for carbon capture.

While DFT calculations have been used to study the thermodynamic properties of these systems, they are in general too costly to be employed to study the thermal properties and the kinetics of CO_2 adsorption. Force fields that offer fast evaluation of energy and forces are needed to simulate finite-temperature effects of the adsorption of molecules to surfaces [25,26] and to porous solids [27,28]. However, existing force fields are not yet optimized for these classes of materials. A recent study of the dynamics of diamines in Mg_2(dobpdc) [29] was performed with reactive force fields (ReaxFFs) using parameters optimized for other systems [30,31]. The accuracy and reliability of these force fields have not been extensively tested on diamine-appended systems, although experimental activation energies of some of the diamine appendages were

reproduced [29]. Linear-scaling DFT-derived force fields, with their ability to enable calculations of systems with about 200 atoms per unit cell that would otherwise take several hours to a few days to be performed in minutes, have yet to be developed or applied to this important class of MOFs.

In this work, we present an accurate and validated reactive interatomic neural network potential (NNP), constructed using machine learning approaches, for functionalized Mg_2(dobpdc) systems with the accuracy of state-of-the-art DFT. These potentials scale linearly with the number of atoms, making them significantly more efficient than DFT methods and suitable for thermal and dynamical properties of materials as complex as diamine-functionalized Mg_2(dobpdc). Our DFT-derived NNP belongs to the class of reactive force fields that allow one to simulate bond breaking and bond formation that are prevalent in these systems during finite-temperature dynamics. The potentials developed here are constructed using an active learning approach combining artificial neural networks and van der Waals (vdW) dispersion-corrected DFT [32]. Our potentials reproduce DFT adsorption energies and zero-temperature mechanical properties, as well as phonon modes and dispersion as a function of temperature, enabling identification of specific temperature-dependent vibrations that govern the thermal properties of these MOFs. Moreover, our potentials can predict the heat capacity and thermal expansion of bare and amine-appended MOF accounting for quantum zero-point and anharmonic effects, in excellent agreement with experiments where available. Finally, we demonstrate that these potentials can be combined with a simulated annealing approach to find global minima for better generating structures for DFT binding energy calculations, particularly for amine-appended MOFs with many degrees of freedom.

II. DENSITY-FUNCTIONAL-THEORY-DERIVED NEURAL NETWORK POTENTIALS

In this work, we construct state-of-the-art machine-learning-based interatomic potentials from a series of DFT calculations of amine-appended MOFs. We refer to these potentials as DFT-derived neural network potentials or simply NNPs, in what follows. These NNPs scale linearly with the number of atoms, making them far more efficient than DFT calculations, and they capture bond breaking and formation, lattice vibrations, and lattice thermal expansion with the accuracy of DFT calculations. DFT-derived force fields have been constructed for MOFs before [33–36], but prior approaches lack sufficient accuracy due to their fixed functional form; and no existing potential is capable of treating amine-appended MOFs and their interaction with CO_2 with the needed accuracy to capture experimental trends, quantitatively or qualitatively. Here, we use modified Behler-Parrinello symmetry functions (BPSFs) and an iterative active learning scheme with an extensive DFT-derived training set to train neural network interatomic potentials for amine-appended MOFs. In what follows, we first describe the BP framework, and then detail the active learning approach and training set, and finally summarize the results of our numerical validation calculations. We use Quantum ESPRESSO [37,38] for all DFT calculations, the PANNA package [39] to train our potentials, and molecular dynamics (MD) simulations are performed with the large-scale atomic/molecular massively parallel simulator (LAMMPS) [40]. As we discuss below, we obtain a high-quality potential from this process, with a root-mean-square error below 5 meV per atom and 0.2 eV/Å relative to DFT for total energies and forces, respectively, on the validation set.

A. Modified Behler-Parrinello symmetry functions

In the neural-network-based interatomic potential framework used here, the total energy E of an N-atom system is expressed as a sum of chemical environment-dependent atomic contributions in the following manner [41]:

$$E = \sum_{i=1}^{N} E_i(G_i(\{\mathbf{R}\})). \tag{1}$$

Here G_i is a function that represents the description of the local chemical environment of atom i and $\{\mathbf{R}\}$ is the set of Cartesian coordinates of all N atoms in the periodic solid.

We use modified BPSFs [39,41,42] for G to describe the local atomic chemical environment. These symmetry functions are atom centered and, for each atom, have both radial and angular components. Both components are described up to a spherical cutoff, R_c, imposed by a function that ensures that the energy and its derivatives smoothly approach zero at the cutoff. The BPSF cutoff function takes the form

$$f_c(R_{ij}) = \begin{cases} \frac{1}{2}\left[\cos\left(\frac{\pi R_{ij}}{R_c}\right)+1\right], & R_{ij} \le R_c, \\ 0 & R_{ij} > R_c, \end{cases} \tag{2}$$

where R_{ij} is the distance between atom i and j.

The radial part of the BPSFs is a function of only the interatomic distances from a central atom i to all its neighbors j within R_c, and is expressed as

$$G_i^{\mathrm{Rad}}[s] = \sum_{j\neq i} e^{-\eta(R_{ij}-R_s)^2} f_c(R_{ij}), \tag{3}$$

where index s runs over the Gaussian centers, R_s, with a fixed Gaussian width, η. By $G_i^{\mathrm{Rad}}[s]$ we denote the sth component of the input vector of atom i. The angular part of the BPSFs depends on all pairs of distances between atom

i and neighbours j and k within R_c that make an angle θ_{ijk}. The angular part of the BPSFs is defined as

$$G_i^{\mathrm{Ang}}[s] = 2^{1-\zeta} \sum_{j,k\neq i} [1 + \cos(\theta_{ijk} - \theta_s)]^{\zeta} \\ e^{-\eta[(R_{ij}+R_{ik})/2 - R_s]^2} f_c(R_{ij}) f_c(R_{ik}), \qquad (4)$$

where s runs over the pair (R_s, θ_s) with R_s being the Gaussian centers of fixed width, η, θ_s is used to represent the three-body angular environments, and ζ controls the resolution of the angular representation.

The final input vector G_i is a concatenation of the radial and angular descriptors resolved along pair and triplets of atomic species, respectively. For instance, for n_s atomic species, $N_{R_s}^r$ Gaussian centers for the radial part, and $N_{R_s}^a$ and N_{θ_s} angular Gaussian centers, the size of G for each atom is $[n_s \times N_{R_s}^r + n_s(n_s+1)/2 \times N_{R_s}^a \times N_{\theta_s}]$.

The BPSF G_i's must contain a minimal but sufficient number of parameters such that the different chemical environments arising in our system for each chemical species are uniquely described. Here, we represent the chemical environment of each central atom i with 24 evenly spaced Gaussian centers, R_s, between 0.75 Å and $R_c = 6$ Å with the value of η fixed at 21 Å^{-2} for the radial part. For the angular part, η is set to zero while the value of θ_s is evenly sampled between 0 and π at an interval of $\pi/16$ with $\zeta = 72$ with a cutoff $R_c = 4$ Å. With this choice of parameters, for five atomic species (H, C, N, O, and Mg), the size of the input descriptor per atom is $5 \times 24 + (5 \times 6)/2 \times 16 = 360$ elements.

B. Active learning algorithm

Our DFT-derived NNPs are constructed with an active learning approach whereby the neural network model self-improves through iterative data generation. DFT calculations on a set of crystalline polymorphs or atomic configurations are performed with Quantum ESPRESSO [37,38] and comprise an initial DFT-based dataset. The dataset is used to train an ensemble of NNPs using PANNA. The NNPs are used to compute classical MD trajectories using LAMMPS, and an ensemble root-mean-square deviation between the NNPs are computed for each subsampled MD trajectory snapshot. If the root-mean-square deviation for a configuration is larger than a predefined threshold then a single-point DFT calculation is performed on it, and it is incorporated into the training set of the next iteration of NNPs. This approach allows us to limit additional single-point DFT calculations to configurations with the greatest disagreement between the ensemble of NNP, i.e., configurations consisting of atomic environments that are less familiar to the NNPs. The ensemble of NNPs is iterated until no newly generated structure has an ensemble root-mean-square deviation greater than the threshold. This active learning approach, summarized in Fig. 1, exploits the inability of NNPs to extrapolate to yet unseen chemical environments to sample relevant configurations and improves the transferability of the NNP.

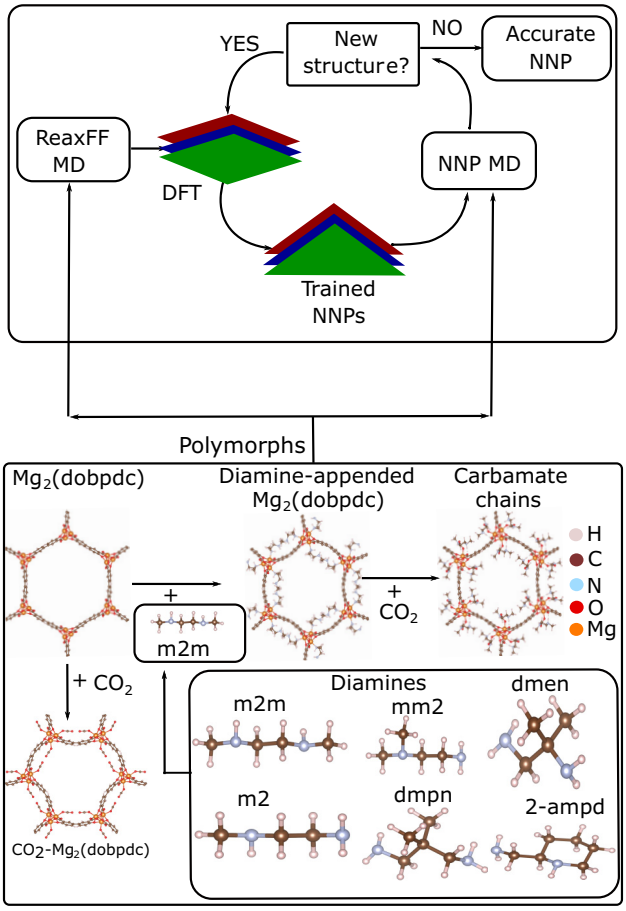

FIG. 1. Summary of the active learning approach and examples of systems considered. Upper panel: the active learning procedure. We initiate the approach using structures generated using an existing reactive force field (ReaxFF) with molecular dynamics (MD), perform a single-point vdW-corrected DFT calculations on subsampled trajectories, train an ensemble of NNPs, perform MD with one of the NNPs to generate structures, perform single-point total energy calculations on structures from a subsampled MD trajectory with another NNP, and compute the ensemble root-mean-square deviation (RMSD). If the RMSD is greater than 5 meV/atom from a structure, it is considered new. New structures are accumulated until all the polymorphs are considered, and then a single-point DFT calculation is performed on the new structures. The newly generated structures are added to the previous ones to train a new ensemble of NNPs. The loop is continued until no new structures are found. Lower panel: the structures used for training the NNP. This includes the bare Mg_2(dobpdc), CO_2 bound Mg_2(dobpdc), diamine-functionalized Mg_2(dobpdc), and the CO_2 bound on the functionalized structures (carbamates). For each diamine (shown in the box titled "Diamine"), the structure with and without CO_2 are added to the training set, which gives rise to 12 polymorphs for the six diamines.

C. Mg_2(dobpdc)-diamine structure and computational details

Here, we briefly summarize the structure of Mg_2(dobpdc), with and without diamine appendages, and with and without CO_2. Figure 1 depicts the crystal structure of Mg_2(dobpdc) in the $P3_221$ space group, as well as with diamine molecules grafted onto the Mg sites and with CO_2 inserted. These MOFs have experimentally well-defined structures [11], although amine appendages and CO_2 inserted structures are known to take up multiple stable configurations at room temperature [13]. This is a result of the strong electrostatic interactions between the framework and the ammonium carbamate chain and between neighboring ammonium carbamate units along the pore direction, as has been reported previously [17], as well as the large degrees of freedom of amines.

In order to generate a diverse training dataset for Mg_2(dobpdc) with different diamine appendages and with CO_2, we start with initial polymorph geometries from a combination of prior experiment and theory work as tabulated in Ref. [16]. The chemical names and abbreviations of the diamines considered are shown in Fig. 2. For each diamine, the corresponding carbamate chain structure formed upon CO_2 insertion is also included in the initial training set of polymorphs. Additionally, the dmpn diamine has the tendency to form carbamic acid pairs and mixed phases comprising a 1:1 ratio of carbamates and carbamic acid pairs [14,16]. Both of these CO_2 bound phases are also included in the training set to increase the diversity of the chemical environments. Finally, the bare Mg_2(dobpdc) and CO_2 bound to the unfunctionalized Mg_2(dobpdc) structures are also considered. In total, 17 structures are included in the training set; the list of experimentally observed CO_2 bound species is illustrated with the dmpn diamine in Fig. 3.

The initial training dataset is first generated using ReaxFF-based molecular dynamics using LAMMPS with parameters previously employed to study the dynamics of diamines in Mg_2(dobpdc) [29]. This initial set of calculations is expected to sample reasonable starting geometries, even as ReaxFF has limitations, as mentioned above (ReaxFF is used only at the initial step). Our MD calculations are done in the *N-P-T* ensemble at a temperature of 300 K and at a pressure of 1 kbar. In the subsequent iterations, we sample the potential energy surface (PES) at four different temperatures, 50, 200, 300, and 500 K, in order to increase the accessible region of PES within the MD simulations and three pressures of −1, 0, and 1 kbar. The accumulated numbers of configurations in the first, second, and third iterations are 21 000, 37 300, and 45 700, respectively. At each iteration, 90% of the dataset is used to train the ensemble of NNPs while the remaining 10% is set aside for validation.

All DFT calculations are performed with the plane-wave projector-augmented-wave [44] framework using the Quantum ESPRESSO package [37,38] using revised-Perdew-Wang [45] exchange and correlation paired with the revised-Vydrov–Van Voorhis (rVV10) [32,46] functional to account for van der Waals dispersion interactions, critical for these materials [47]. This

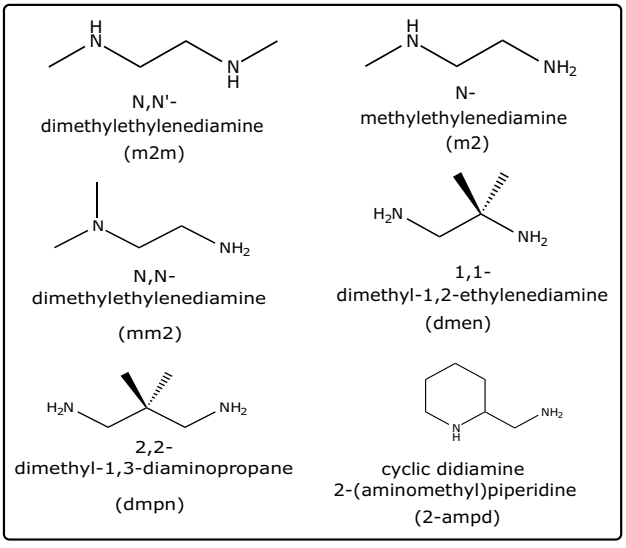

FIG. 2. Lewis structures of diamines and their molecular names used to generated the training dataset during the active learning procedure. Abbreviations are shown in parentheses. Lewis structures of other amine considered are shown in the Supplementary Material [43].

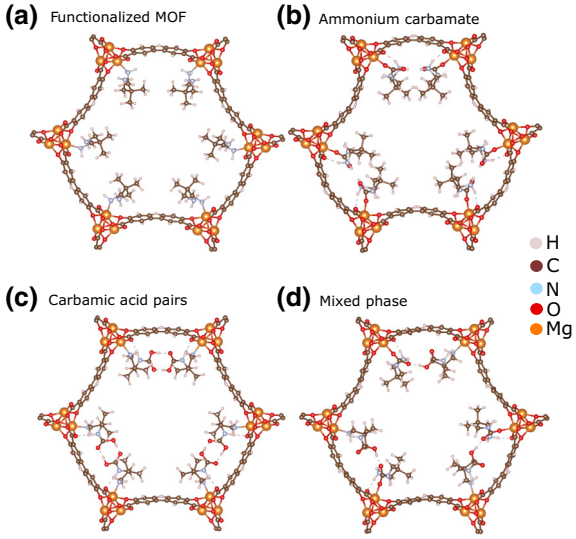

FIG. 3. Examples of diamine-appended MOFs with the three different CO_2 binding species observed experimentally [16], illustrated with the dmpn diamine. (a) The dmpn-functionalized Mg_2(dobpdc), (b) ammonium carbamate chain along the perpendicular direction, (c) carbamic acid pairs, and (d) a mixed phase of ammonium carbamate and carbamic acid pair.

functional was recently shown to perform well describing the structural properties as well as energetics of Zn-based amine-appended MOFs [47]. A plane-wave basis set is employed, and kinetic energy cutoffs of 80 and 480 Ry are used for the expansions of wave functions and charge density, respectively. The valence electrons treated explicitly for each element are H($1s^1$), C($2s^2 2p^2$), N($2s^2 2p^3$), O($2s^2 2p^4$), and Mg($2s^2 2p^6 3s^2$). We sampled the Brillouin zone uniformly according to the Monkhorst-Pack [48] procedure with a Marzari-Vanderbilt [49] smearing of 0.02 Ry. A **k**-point mesh of $1 \times 1 \times 3$ centered around Γ is found to be sufficient to converge the ground-state energy to 1 meV per atom for the bare Mg_2-dobpdc unit cell. Diamine-appended and CO_2-bound M2-dobpdc are sampled with the same **k**-point density. Our calculations of isolated CO_2 and diamine molecules are carried out in a 15 Å cubic cell with a Γ point sampling of the Brillouin zone. Similar **k**-point meshes have been used in a recent prior study of amine-appended MOFs [47].

We use the PANNA software package [39] to construct a neural network consisting of two hidden layers with 20 nodes each and an output layer with a single node for the atomic energy that enters Eq. (1). The total number of parameters per species is $((360 \times 20 + 20) + (20 \times 20 + 20) + (20 \times 1 + 1)) = 7661$, where 360 is the size of the input vector. After the active learning approach is converged, we increased the number of parameters by increasing the size of the hidden layers to 64 to obtained a more accurate NNP. In this case, the total number of parameters is $((360 \times 64 + 64) + (64 \times 64 + 64) + (64 \times 1 + 1)) = 27\,329$.

D. Neural network potential training

As mentioned above, we take an active learning approach to training our NNP. As a first step, we generate configurations using molecular dynamics with ReaxFF for all systems in the initial dataset, and perform single-point DFT calculations on a subsample of these ReaxFF trajectories. Typical ReaxFF trajectories consist of 40 000 snapshots, and within these trajectories we perform DFT calculations on a subsample of 1000 structures. Including all systems and subsampled trajectories considered in the initial dataset, we perform a total of 17 000 single-point DFT calculations on unit cells containing between 84 and 228 atoms. Out of the 17 000 DFT calculations, 15 000 calculations converged within 50 self-consistent cycles of the Kohn-Sham equations. In addition, we included configurations of isolated diamine molecules that are required, for instance, to compute the binding energy of the diamine itself in amine-appended MOFs. Figures 4(a) and 4(b) show total energies and forces obtained with DFT and ReaxFF on the initial dataset. Notably, for ReaxFF, the R^2 scores for the energy and forces are 0.68 and −0.06, respectively, indicating that it does not yield reliable energetics and does not accurately sample the PES, motivating the need for DFT-derived NNPs for this system.

Following the ReaxFF calculations, we train an ensemble of NNPs with the DFT single-point calculations, and perform longer MD simulations of 200 000 time steps with one of the NNPs to generate new trajectories and structures; we sample these new trajectories by performing single-point calculations with the other NNPs in the ensemble, and compute the ensemble root-mean-square deviation (RMSD) for the sampled snapshots. If the RMSD is greater than 5 meV/atom for a given snapshot, we compute its energy and forces with DFT, and add the structure to the dataset, updating it. These new DFT data are added to the previous training set and used to train a new ensemble of NNPs, and the procedure is iterated until no new structures need to be added to the dataset. See Fig. 1 for a summary of the procedure.

The performance of our DFT-derived NNP is summarized in Fig. 4. Training the NNP with DFT energy and forces, we find that the R^2 scores on the validation set (i.e., structural snapshots that are not included in the training set) for the total energies and forces become 1.00 and 0.99 [see Figs. 4(c) and 4(d)], respectively, indicating essentially perfect correlation between the NNP and DFT results. The root-mean-square errors (RMSEs) for energies and forces for the test set are 3.93 meV/atom and 116 meV/Å, respectively.

In Table I, we summarize the RMSEs between DFT and the NNP for the total energies and atomic forces for the training and validation sets at each of the three active learning steps. We denote the NNP associated with the first step as NNP_1, the second as NNP_2, and the third as NNP_3. The corresponding validation sets are also denoted VS_1, VS_2, and VS_3. Each VS_k corresponds to the accumulation of all the validation sets up to iteration k. TS is used to denote an average RMSE on the training dataset obtained by averaging the RMSE of the last 10% of the training steps. The RMSE in energy obtained with NNP_3 on the VS_1 validation set is 5.1 meV/atom, comparable to the RMSEs obtained with NNP_2 and NNP_1 of 3.1 meV/atom and 3.9 meV/atom, respectively. The RMSE obtained with NNP_1 on VS_2 increases to 44.9 meV/atom and the corresponding RMSEs for NNP_2 and NNP_3 are much lower and are 5.6 and 6.4 meV/atom, respectively. The increase in the RMSE of NNP_1 is a result of a large region of the potential energy surface that is sampled in subsequent iterations (see Fig. S2 in the Supplemental Material [43]). Finally, using NNP_3 with VS_3 yields an RMSE in energy of 7.288 meV/atom, lower than 17.76 and 49.06 meV/atom obtained with NNP_2 and NNP_1, respectively. Figure 5 shows the deviation of the NNP predictions from DFT energies and forces on VS_3 dataset, showing the improvement in accuracy of NNP_3 over NNP_1 and NNP_2, and NNP_2 over NNP_1 in both energy and force predictions. From this analysis, we can see that NNP_3 is

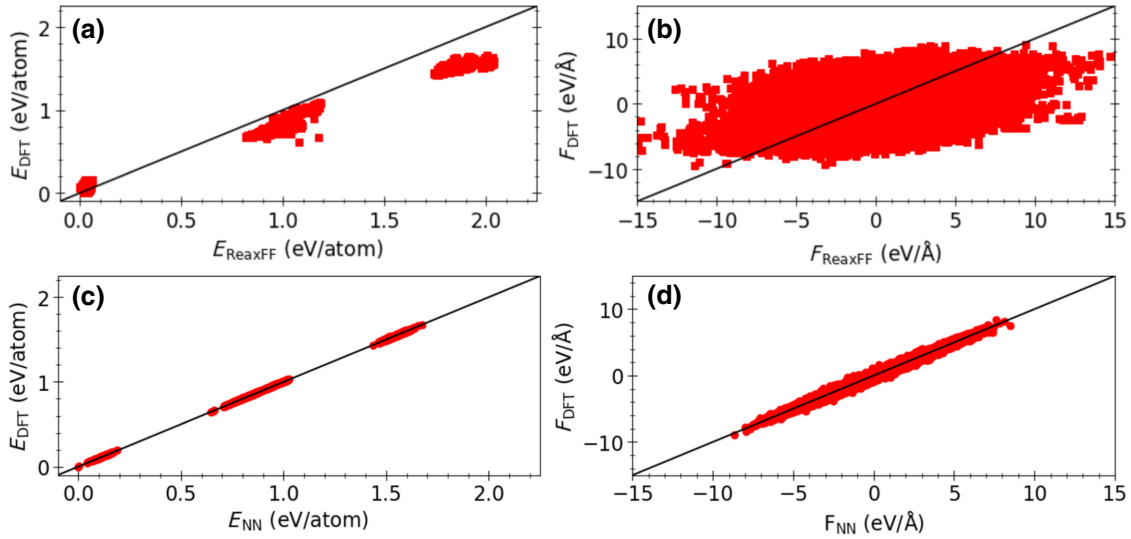

FIG. 4. Comparison of cohesive energies (relative to the respective lowest energy) and atomic forces with DFT and (a),(b) ReaxFF and (c),(d) our NNP at the first active learning step. The R^2 scores for the energy and forces with ReaxFF are 0.68 and −0.06, respectively, while our NNP yields R^2 scores of 1.000 and 0.993 for the energy and forces, respectively.

more transferable to a diverse chemical environment than NNP_2 and NNP_1, while still having similar accuracy to NNP_1 and NNP_2 on VS_1, a less diverse validation dataset (see Fig. S3 in the Supplemental Material [43]).

The increase in the training RMSE in both energy and forces going from NN_1 to NN_3 reflects the increasing diversity of the training dataset as the active learning procedure progresses, sampling a wider range of energy and forces (see Fig. S2 in the Supplemental Material [43]) while the number of parameters in the neural network model remains unchanged. For instance, the training RMSE in energy (forces) at the first step of the active learning is 3.96 meV/atom (113 meV/Å), increases to 6.12 meV/atoms (140 meV/Å) at the second step, and then increases to 9.31 meV/atom (219 meV/Å) at the third step although the prediction error on a diverse dataset, VS_3, decreases as discussed above. The dependence of the RMSE on the diversity of the training and validation dataset has been discussed recently for carbon systems in Ref. [51], where decreasing diversity of the dataset was found to lead to a smaller RMSE validation dataset. However, when the model trained on less diverse data is used to evaluate a dataset of greater diversity, the model performed much worse than that trained on a more diverse dataset. In order to minimize the effect of parameters to diversity ratio, we increase the number of nodes in the hidden layers from 20 to 64 nodes, leading to a total of 27 329 parameters per species, and then retrain the NNP model. It is worthwhile to mention that this final step is merely a postprocessing of the dataset generated via the active learning procedure.

In Table I, we also show the RMSEs for the total energy and atomic forces obtained with 64 nodes in the hidden layers, comparing its performance to the 20-node case. We find that the final RMSE in energy is 5.9 meV/atom, which is better than the RMSE of 7.3 meV/atom obtained using the model with fewer parameters. The RMSE in forces is greatly improved and decreased from 222 to 173 meV/Å. Figures S4(a) and S4(b) in the Supplemental Material [43] compare the prediction errors of the two models. It can be seen that the latter model shows an improvement on VS_3 over the former model. The RMSEs on VS_1 and VS_2 are also summarized in Table I. They follow the same trends as in the NNP with hidden layers of 20 nodes, showing improvement in the prediction of energy and forces.

E. Performance of DFT-derived NNP on equilibrium properties of diamine-appended Mg(dobpdc)

1. Lattice parameters and elastic moduli

Having developed our DFT-derived NNP, we analyze its performance by first computing the zero-temperature structural and mechanical properties of the bare Mg(dobpdc) MOF, multiple amine-appended Mg(dobpdc) MOFs, and CO_2 inserted amine-appended Mg(dobpdc). The latter structure involves CO_2 molecules cooperatively inserting into metal-diamine bonds followed by formation of a well-ordered ammonium carbamate chain structure. For all systems, we minimize the total energy with respect to volume and atomic coordinates using our NNP with LAMMPS.

As mentioned above, Mg_2(dobpdc) consists of a periodic arrangement of MgO_5 square pyramidal units connected by dobpdc linkers, crystallizing in the $P3_221$ space group. Using the NNP, our computed lattice parameters, a and c, are 21.851 and 6.926 Å; the DFT-rVV10 results

TABLE I. RMSEs in the total energy and atomic forces for the validation set for each step of the active learning process. VS_i denotes the accumulated validation dataset at iteration i. TS denotes the RMSE computed by averaging over the RMSE obtained at the last 10% of the training steps. The error bar indicates the deviation of the energy and atomic forces during training due to the minibatch stochastic optimization used for parameters optimization [50]. Throughout the active learning approach, we use a 20:20:1 neural network architecture (neural network with two hidden layers each having 20 nodes and an output layer with one node); in a final postprocessing step, we use a 64:64:1 architecture.

	RMSE							
	Energy (meV/atom)				Force (meV/Å)			
	TS	VS_1	VS_2	VS_3	TS	VS_1	VS_2	VS_3
				20:20:1				
NNP_1	3.962 ± 1.9	3.933	44.94	49.06	113.1 ± 11.7	116.8	663.2	916.9
NNP_2	6.119 ± 5.1	3.090	5.593	17.76	139.8 ± 41.1	127.9	167.1	690.1
NNP_3	9.306 ± 6.1	5.070	6.359	7.288	218.8 ± 80.9	156.5	201.1	221.6
				64:64:1				
NNP_3	6.263 ± 4.1	3.565	5.178	5.858	169.3 ± 78.9	115.4	164.2	172.7

are 21.972 and 6.885 Å, respectively, yielding excellent agreement, with deviations of only −0.55% and 0.60% respectively. Similarly, the NNP lattice parameters, a and c, of the CO_2-bound Mg_2(dobpdc) are 20.840 and 6.762 Å, compared with 21.089 and 6.853 Å from DFT, equivalent to 1.19% and −1.32% deviations. The NNPs are highly accurate at just a fraction of the computational cost of a DFT calculation.

To further assess the performance of our NNP, we additionally compute the equilibrium lattice parameters of eleven diamine-appended systems and 343 tetraamine. The amines can be classified into two categories: those that are used to generate the dataset to train the NNP model, which we henceforth refer to as "in-sample" amines, and those that not used to train the NNP model, or "out-of-sample" amines. There are six in-sample amines, namely, m2m, m2, mm2, dmen, dmpn, and ampd; and five out-of-sample amines, e2, e2e, ee2,i2, mpn, and 343. The out-of-sample amines allow us to directly assess the transferability of the NNP model on unseen chemical environments.

As a measure of NNP accuracy in what follows, we define the percentage deviation of our NNP calculations from DFT rVV10 as $\delta l = [(l_{\mathrm{NNP}} - l_{\mathrm{DFT}})/l_{\mathrm{DFT}}] \times 100$, where l represents the norm of the lattice vectors. Figure 6 summarizes our results for all systems considered. The values plotted are tabulated in Table S1 of the Supplemental Material [43]. The a lattice parameters for the diamine-appended MOFs in the absence of CO_2 are

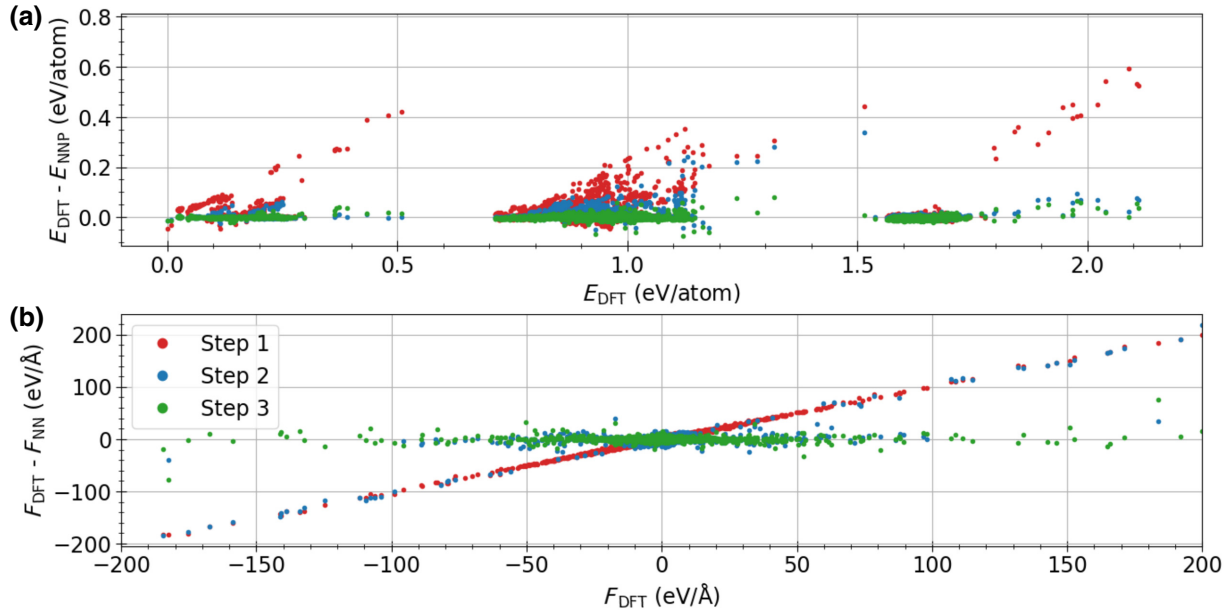

FIG. 5. Errors on the final validation set (VS_3) relative to DFT using NNP at step 1 (red), step 2 (blue), and step 3 (green) of the active learning process on (a) cohesive energy per atom and (b) force components.

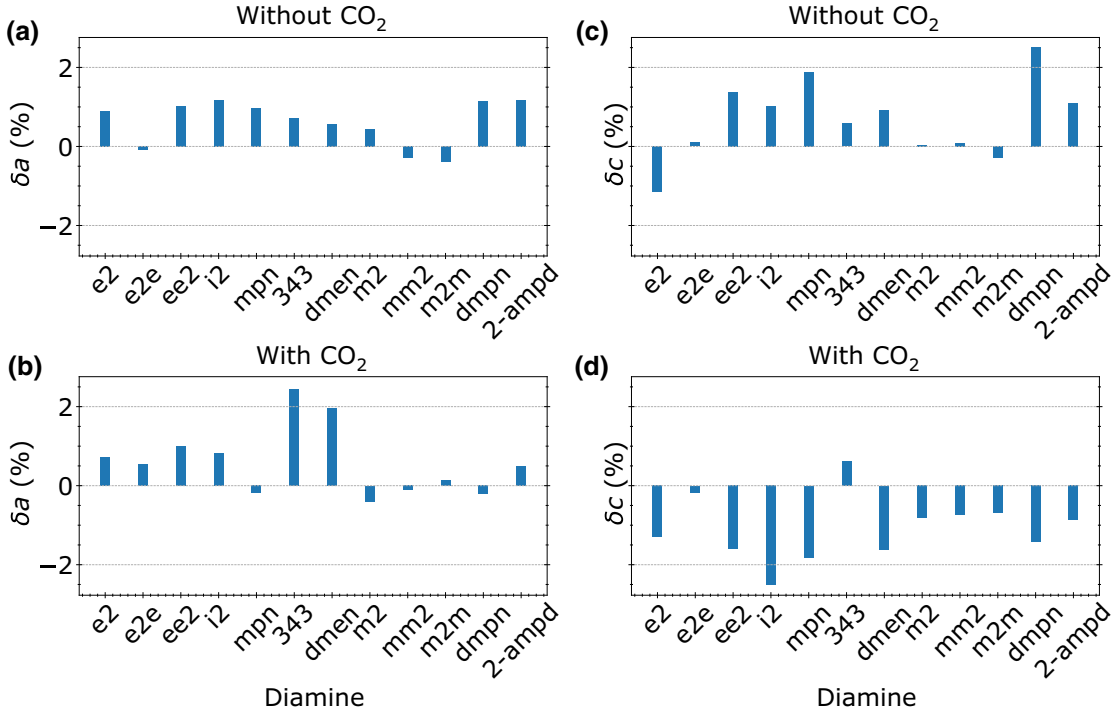

FIG. 6. Comparison of lattice parameters predicted by our NNP and by DFT rVV10. Percent differences of *a* lattice parameters for (a) amine-appended MOFs without CO_2 and (b) amine-appended MOFs with CO_2. Percent differences of *c* lattice parameters for (c) amine-appended MOFs without CO_2 and (d) amine-appended MOFs with CO_2.

reproduced accurately, all within 1.5% of the DFT-rVV10 results. Good accuracy can be seen for *c* lattice parameters for both in-sample and out-of-sample diamines, including the 343 tetraamine that differs more substantially from the diamines. Among the in-sample diamines, dmpn is an outlier, and even that system exhibits a modest deviation, just slightly above 2%. For the CO_2-bound systems, the *a* lattice parameters are on average predicted to deviate by 1% from the DFT results, with two outliers, namely, dmen and 343, for which we compute deviations of 1.95% and 2.44%, respectively. For the *c* lattice parameters, the NNP predicts values consistently smaller than the DFT results, by about 2% except the 343 tetraamine, which is overestimated by 1%. Notably, this deviation is of a similar magnitude to the deviation between rVV10 and other nonlocal vdW functionals, as reported by a study on Zn_2(dobpdc) systems [47].

The overall accuracy of our DFT-derived NNP Mg_2(dobpdc) lattice parameters, within less than 1% of DFT for a diverse set of diamine appendages, is an affirmation of the robustness of our active learning procedure and the transferability of our NNP in capturing structural properties in Mg_2(dobpdc) and its amine-appended and CO_2-bound variants.

We also compute the elastic constants for select systems with our NNP and compare with DFT (see Table S2 of the Supplemental Material [43]). We report the computed moduli using Voigt-Reuss-Hill averages [52], defined as in Ref. [17]. We consider bare Mg_2(dobpdc) (denoted MOF), and m2m, m2, e2, and dmen diamine-appended Mg_2(dobpdc) with and without CO_2. Figure 7 compares the bulk modulus, shear modulus, and Young's modulus obtained with the NNP and with DFT. Consistent with previous DFT calculations, we find that the diamine appendages enhance the mechanical properties, both with and without CO_2. For all systems considered, the NNP predicts bulk moduli in excellent agreement with DFT, both quantitatively and qualitatively. For shear and Young's moduli, the NNP captures DFT trends in which the mechanical properties of diamine-appended MOFs with and without CO_2 is enhanced over the unfunctionalized MOF but underestimates the moduli. The challenges in comparing the DFT and NNP shear and Young's moduli is a result of their small values and, more physically, the overall compliant or soft elastic character of these systems, which challenge the numerical calculation of second derivatives and thus comparison of the DFT and NNP stress tensors.

2. CO_2 adsorption energies

An important quantity determining the thermodynamic stability of CO_2 adsorption in amine-appended MOF-274 is the CO_2 binding enthalpy [13,14,21]. A closely related but distinct quantity, the binding energy—which neglects

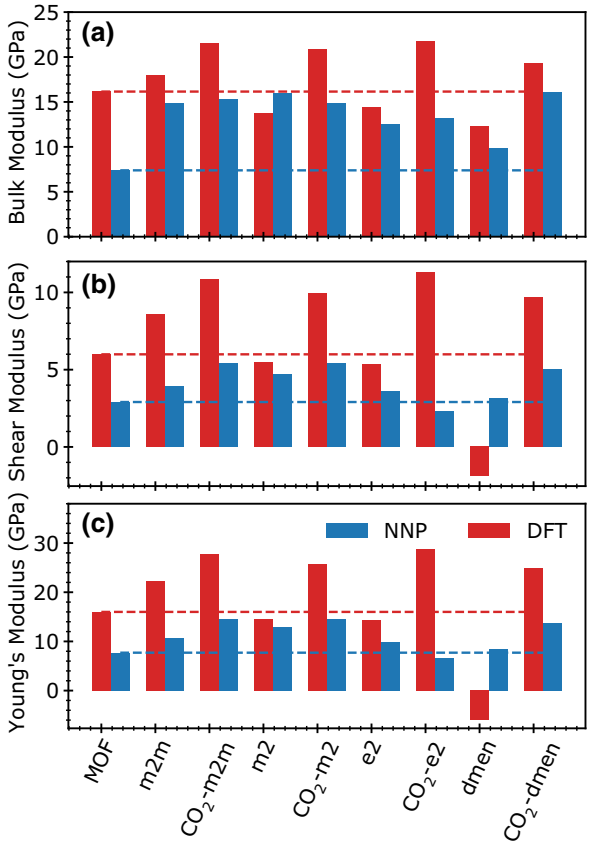

FIG. 7. Mechanical properties of diamine-appended Mg_2(dobpdc) systems: (a) bulk modulus, (b) shear modulus, and (c) Young's modulus for bare and several diamine-appended Mg_2(dobpdc). DFT here refers to the rVV10 functional; see the text.

the contributions from the finite pressures and temperatures, typically a few to 10 kJ/mol—can be accurately computed with DFT using van der Waals–corrected functionals [16,17,21,47]. Finite-temperature contributions to the binding enthalpy from lattice vibrations can also be accurately computed with DFT, as has also been shown elsewhere [53]. Here, we define the binding energy, E_b, as

$$E_b = E_{\mathrm{CO_2+diamine}} - E_{\mathrm{diamine}} - E^{\mathrm{iso}}_{\mathrm{CO_2}}, \tag{5}$$

where $E_{\mathrm{CO_2+diamine}}$ is the total energy of CO_2-bound diamine-appended Mg_2(dobpdc), E_{diamine} is the total energy without CO_2, and $E^{\mathrm{iso}}_{\mathrm{CO_2}}$ is the total energy of an isolated CO_2 in a cubic supercell of size 15 Å.

For CO_2 adsorption at open metal sites in unfunctionalized Mg_2(dobpdc), our NNP predicts CO_2 adsorption energy of −48.3 kJ/mol, in excellent agreement with the DFT-rVV10 value of −50 kJ/mol. Additionally, the NNP accurately reproduces the DFT adsorption energy as a function of the Mg-O (O in CO_2) distance between 1.5 and 5 Å (see Fig. S5 in the Supplemental Material [43]).

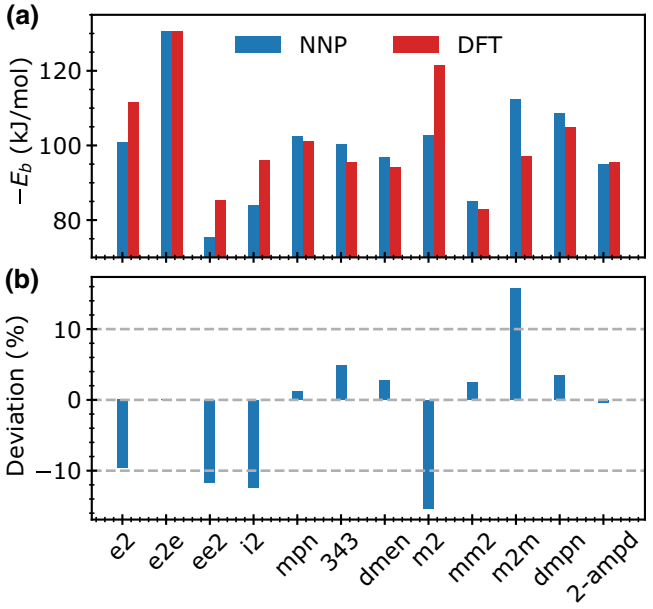

FIG. 8. Comparison of the CO_2 adsorption enthalpies predicted by the NNP and DFT rVV10. (a) Adsorption enthalpies with red bars correspond to DFT results and blue bars represent the NNP. (b) Percent difference between the NNP and DFT, defined as $(E_b^{\mathrm{NNP}} - E_b^{\mathrm{DFT}})/E_b^{\mathrm{DFT}} \times 100$.

These results suggest that the NNP can be used to simulate the dynamics of CO_2 adsorption in unfunctionalized Mg_2(dobpdc).

Figure 8(a) compares the CO_2 adsorption energy predicted with our NNP and with DFT using the rVV10 functional. The NNP E_b is within 5% for six of the diamines (e2e, mpn, dmen, mm2, dmpn, and 2-ampd) and the 343 tetraamine. The remaining amines are well reproduced to within about 10% of the respective DFT values [see Fig. 8(b)]. To understand the origin of these somewhat larger deviations in the adsorption energy of some of the amines, we express the NNP adsorption energy in terms of DFT adsorption energy and the difference δ as

$$E_b^{\mathrm{NNP}} = E_b^{\mathrm{DFT}} + \delta, \tag{6}$$

where

$$\begin{aligned}\delta &= (E^{\mathrm{NNP}}_{\mathrm{CO_2+diamine}} - E^{DFT}_{\mathrm{CO_2+diamine}}) - (E^{\mathrm{NNP}}_{\mathrm{diamine}} - E^{\mathrm{DFT}}_{\mathrm{diamine}}) \\ &= \Delta E_{\mathrm{CO_2+diamine}} - \Delta E_{\mathrm{diamine}},\end{aligned}$$

and where ΔE is the energy difference between results from NNP and DFT. Here we omit energy differences from DFT and NPP total energy calculations of the isolated CO_2 molecule.

As defined above, δ is thus the size of error cancelation in the energy differences. Figures 9(a) and 9(b) show the values of ΔE for the constituents and δ, respectively. The deviations of the NNP total energy from DFT is within

5 meV/atom for dmen-, m2-, m2m-, mm2-, dmpn- and 2-ampd-appended Mg_2(dobpdc) with and without CO_2; since these systems were part of the in-sample amines used to construct the NNP, this agreement reflects the usual bias towards data trained upon. As expected, the out-of-sample diamine-appended Mg_2(dobpdc) systems, namely, MOFs with e2, e2e, ee2, i2, and mpn diamines, have larger deviations but still exhibit good accuracy relative to their DFT-rVV10 values. However, δ is notably larger for m2m and m2 diamine appendages. These larger values of δ reflect a difference in the sign of the two separate errors it is composed of, suggesting that the NNP errors in binding energy can be minimized via two independent routes. The first route, the one taken here and the one which is arguably most important and controllable, is that the NNP is constructed so that its total energy errors are as small as possible, thereby reducing ΔE. However, minimizing the error in total energies against the in-sample set must be done with care so as not to overfit, leading to poor transferability to a wider range of chemical environments. The second route is to bias the potential so that the error has very similar magnitude (with the same sign) for the two errors making up δ, maximizing error cancelation. This can be achieved by training on binding energy rather than the total energy. However, training on the binding energy would be expected to lead to poorer accuracy in atomic forces, leading to a NNP that is inaccurate for molecular dynamics simulations. For these reasons, we take the first (former) route here, accepting the errors in binding energies to ensure more accurate atomic forces and dynamics.

III. VIBRATIONAL AND FINITE-TEMPERATURE PROPERTIES OF AMINE-APPENDED MOFS WITH AND WITHOUT CO_2

In the previous section, we established the performance of our NNP in reproducing DFT structural and mechanical properties as well and the adsorption energy of a wide range of amine-appended systems. In this section, we demonstrate the utility of the NNP by applying it to compute amine-appended MOF phonon spectra and heat capacity, and lattice thermal expansion via finite-temperature molecular dynamics, underexplored and computationally demanding properties of amine-appended MOFs relevant to their ability to efficiently capture CO_2. We also discuss the use of our NNP for simulated annealing, and illustrate its promise with a specific example.

A. Phonon spectrum and phonon density of states of amine-appended MOFs

Accurate and inexpensive calculations of the vibrational spectra of amine-appended Mg_2(dobpdc) crystals is essential to understanding their thermal properties. For instance, the heat capacity that dictates a significant part of the

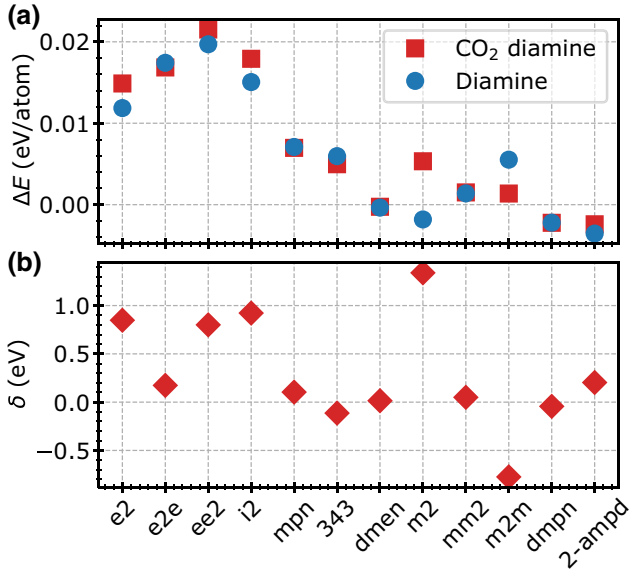

FIG. 9. Error analysis of CO_2 adsorption energies for a number of diamine-appended MOFs, with and without CO_2. (a) Total energy differences between NNP and DFT rVV10 for diamine-appended Mg_2(dobpdc) and carbamate structures. (b) Differences in errors (see the text for a definition of δ).

CO_2 regeneration energy in carbon capture applications [54] can be computed from the vibrational frequencies; this approach has been used in conjunction with machine learning to predict the regeneration energy of unfunctionalized MOFs [55] in which the harmonic heat capacity was learned from the geometric features and the identities of the metal nodes of MOFs. Furthermore, the thermal conductivity, an essential quantity understanding heat dissipation during adsorbent regeneration, can be computed from the harmonic and higher-order force constants, which, although beyond the scope of the present work, requires accurate and low cost computation of vibrational properties for these complex systems.

All vibrational frequencies of Mg_2(dobpdc) systems are calculated with our NNP via a finite-difference approach as implemented in phonopy [56] and phonoLAMMPS [57] using a $1 \times 1 \times 3$ supercell. Figure 10(a) shows the phonon dispersion of unfunctionalized Mg_2(dobpdc) along a high-symmetry path of its hexagonal Brillouin zone. Only low-frequency modes are shown and we depict the total phonon density of states at all frequencies in Fig. 10. Interestingly, a gap appears between 47.8 and 92.6 THz. All phonon frequencies are real, implying that the NNP potential predicts a mechanically stable structure. The first nonzero mode at Γ corresponds to the motion of adjacent linkers in opposite directions (see Fig. S6 in the Supplemental Material [43]) and possesses a frequency of 0.7 THz. The next vibration mode at 1.52 THz corresponds to a rotational vibration of atoms in the linkers around

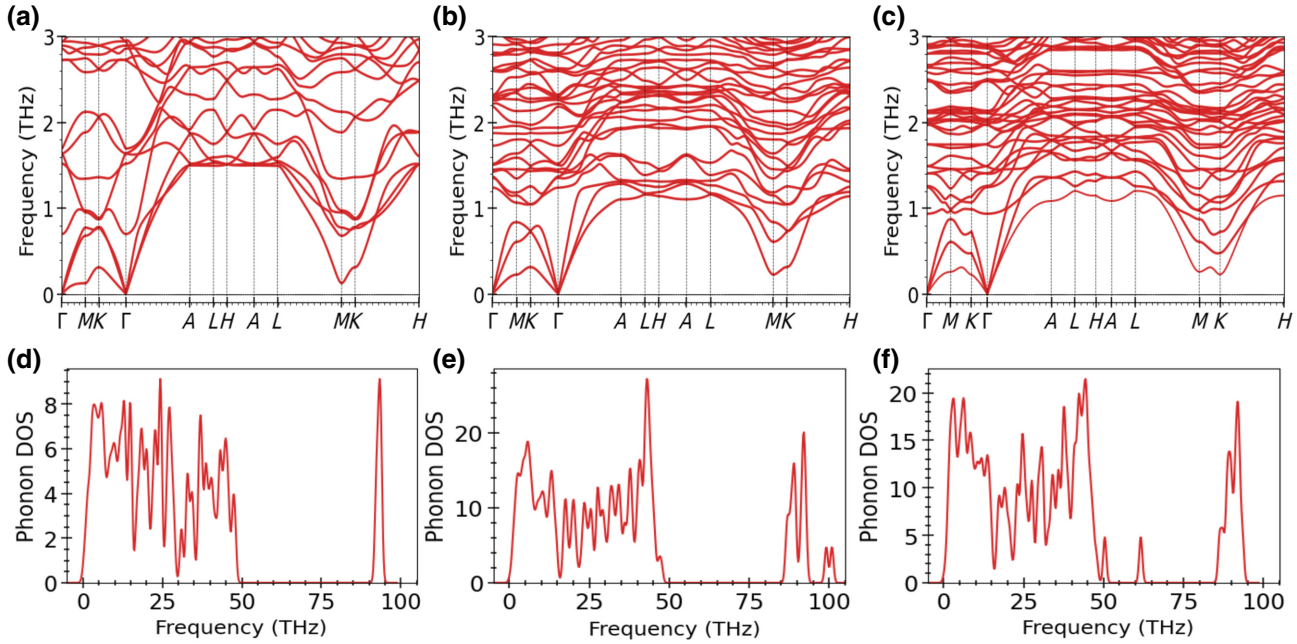

FIG. 10. NNP-computed phonon dispersion of (a) Mg_2(dobpdc), (b) m2m diamine-appended Mg_2(dobpdc), and (c) CO_2 bound to m2m diamine-appended Mg_2(dobpdc). NNP-computed phonon density of states for (d) Mg_2(dobpdc), (e) m2m diamine-appended Mg_2(dobpdc), and (f) CO_2 bound to m2m diamine-appended Mg_2(dobpdc).

the Mg-O clusters, with atomic motion around neighboring nodes in opposite directions. At the top of the gap [see Fig. 10(d)], the mode at 47.8 THz corresponds to C—C bond stretching. Finally, the highest energy vibrational modes, above 92 THz, involve C—H bond stretching.

Next, we analyze the phonon modes of diamine-appended Mg_2(dobpdc), with and without CO_2; in particular, we focus on the m2m system. In Figs. 10(b) and 10(c), we show the phonon dispersion of m2m-Mg_2(dobpdc) and CO_2-bound m2m-Mg_2(dobpdc)(right), respectively, with accompanying phonon densities of states in Figs. 10(e) and 10(f). For m2m-Mg_2(dobpdc) without CO_2, the lowest nonzero frequency modes at Γ, occurring at 1.24 THz compared to 0.7 THz for the bare Mg_2(dobpdc), corresponds to a motion of two adjacent organic linkers and their diamines towards different pores. The vibrational mode at 1.34 THz corresponds to a rotational motion of atoms in the linkers and diamine about the Mg-O clusters, while atoms adjacent to the cluster rotate in the opposite direction; the mode at 1.35 THz corresponds to out-of-phase vibration of adjacent nodes along the pore direction. The highest energy mode, just at the gap at 47.5 THz, corresponds to C—C stretching as in bare Mg_2(dobpdc). The modes above the gap, starting at 86 THz, correspond to C—H and N—H bond stretches, with the secondary-amine N—H bonds dominating at the maximum computed frequency at 100.9 THz. Unlike Mg_2(dobpdc) without diamine appendages, which exhibits a narrow peak above the gap, the diamine-appended MOF (also with CO_2; see below) leads to a broader band above the gap and introduces a different type of C—H bond stretching mode unique to the diamines. For instance, the C-H vibrational modes in the diamine have lower energy than those in the dobpdc organic linkers.

For CO_2-bound m2m-Mg_2(dobpdc), the lowest energy optical vibrational mode at 0.95 THz corresponds to a rotation of atoms on linkers and CO_2-bound diamines about the Mg-O cluster nodes with neighboring nodes vibrating in opposite directions. The rest of the modes are similar to m2m-Mg_2(dobpdc) without CO_2. One noticeable difference between the vibrational properties of m2m and CO_2-bound m2m-Mg_2(dobpdc) is the appearance of spectral weight in the gap of the phonon density of states as seen in Fig. 10(f). This band corresponds to N—H bond stretching on the protonated amine terminal of the CO_2-bound diamine. The protonation weakens the N—H bond, leading to a softening of the phonon mode.

B. Heat capacity

Despite the importance the heat capacity in determining the regeneration energy for a carbon capture adsorbent, limited information about the heat capacity of amine-appended MOFs is available in literature, as computing phonon spectra for structures as complex as MOFs with DFT is computationally challenging. Only recently a machine learning model was developed for the computation of the heat capacity of a diverse class of unfunctionalized MOFs from geometric features [58]. The results demonstrate the importance of using the actual heat capacity of MOFs to estimate their regeneration energy rather

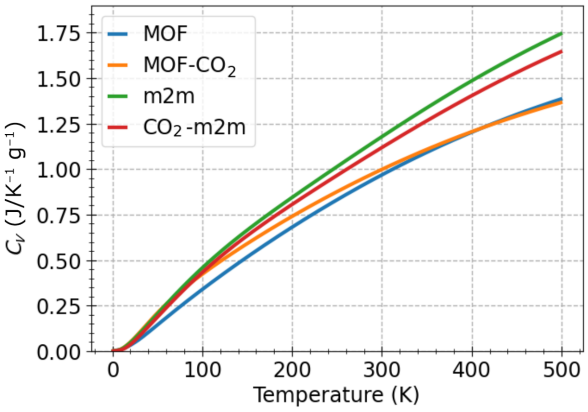

FIG. 11. Heat capacity as a function of temperature, determined from NNP-calculated phonons at zero temperature.

than a constant reference value [58]. Here, we compute and analyze the heat capacity of Mg_2(dobpdc) and its amine-functionalized variants via a finite-difference approach using our DFT-derived NNP.

The heat capacity at constant volume of a system of harmonic phonons can be computed as

$$C_v(T) = \frac{1}{N_q}\sum_{\mathbf{q},\nu} k_B \left(\frac{\hbar\omega(\mathbf{q},\nu)}{2k_BT}\right)^2 \frac{1}{\sinh^2(\hbar\omega(\mathbf{q},\nu)/2k_BT)}, \tag{7}$$

where $\omega(\mathbf{q},\nu)$ is the vibrational frequency at wave vector $\mathbf{q}$ with a branch index ν and N_q is the number of $\mathbf{q}$ points used to sample the Brillouin zone.

In Fig. 11, we show the calculated heat capacity per gram (also referred to as the "specific heat capacity") as a function of temperature of the bare Mg_2(dobpdc), CO_2-Mg_2(dobpdc), and m2m-Mg_2(dobpdc) systems. The specific heat capacity of Mg_2(dobpdc) is 0.97 $JK^{-1}\,g^{-1}$ at 300 K, and we further find that it increases by 3% upon adsorption of CO_2, a trend consistent with experimental measurements for the related Zn_2(dobdc) system [59]. Adding amine appendages increases the heat capacity by 18% for ampd-Mg_2(dobpdc) and 25% for dmpn-Mg_2(dobpdc) without CO_2 adsorption. When CO_2 is adsorbed, the heat capacity decreases by 5% and 6% for ampd-Mg_2(dobpdc) and dmpn-Mg_2(dobpdc), respectively.

To understand the origin of the increase with amine appendages, we decompose the heat capacity into individual phonon mode contributions as a function of temperature. The result is shown in Fig. 12. At 50 K only modes up to 4 THz contribute to the heat capacity for all three systems. As the temperature increases, higher frequency modes become activated. In particular, at 300 K, the heat capacity saturates at about 40 THz. A steeper slope is observed for the amine-functionalized MOF with and without CO_2 compared to the bare MOF, reflecting the presence of additional low-frequency modes with the diamine appendages. These modes correspond to the collective motion of the diamines in the pore. The decrease in the heat capacity of CO_2-bound m2m-Mg_2 is due to a reduction of the flexibility of the diamine when CO_2 is adsorbed. For all temperatures considered, the modes above the gap, corresponding to H—C and H—N bond stretches, do not contribute to the heat capacity.

From the foregoing analysis, it is expected that, generally speaking, diamines with a larger number of degrees of freedom will have higher heat capacities, as they should exhibit additional low-frequency modes. To verify this, we compute the heat capacity of ampd-, e2e-, and dmpn-Mg_2(dobpdc). ampd-Mg_2(dobpdc) has less degrees of freedom than m2m, while e2e and dmpn have more degrees of freedom that m2m. Table II shows the computed heat capacities and the change in heat capacity with respect to the unfunctionalized Mg_2(dobpdc) at 300 K. The heat capacity per gram increases with increasing degrees of freedom of the diamine with ampd having the lowest value and dmpn the highest, verifying intuition and an important consideration for tuning the heat capacity in these systems.

C. Lattice thermal expansion

The lattice thermal expansion of M_2(dobpdc) and its amine functionalized variants has yet to be theoretically

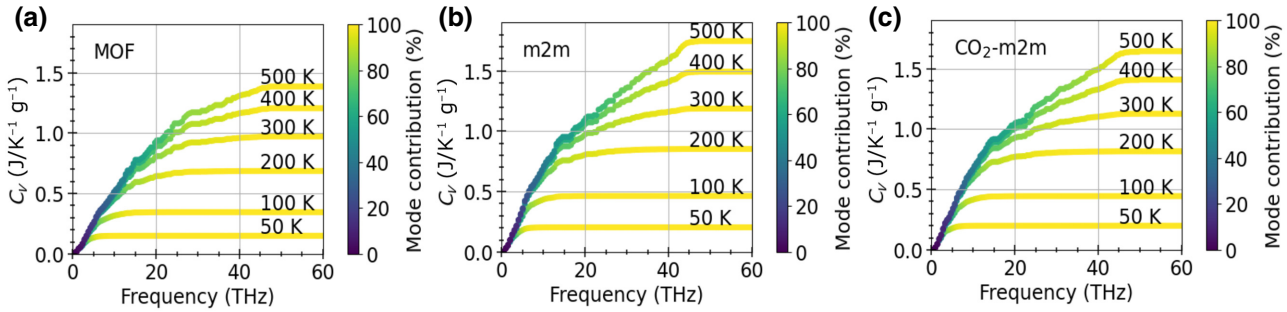

FIG. 12. Heat capacity decomposed into phonon modes, highlighting the dominant phonon modes at each temperature. (a)Mg_2(dobpdc) (b) m2m diamine-appended Mg_2(dobpdc), and (c) CO_2 bound to m2m diamine-appended Mg_2(dobpdc).

TABLE II. Heat capacity at constant volume computed from the NNP for different diamines at 300 K within the harmonic approximation using zero-temperature phonons. Here ΔC_v is the change in C_v with respect to unfunctionalized Mg_2(dobpdc).

Diamine	C_v (J/K^{-1} g^{-1})	ΔC_v (%)
MOF	0.968	0
CO_2	0.998	3.09
ampd	1.137	17.5
343	1.165	20.4
m2m	1.178	21.7
dmpn	1.212	25.2
e2e	1.200	23.9
CO_2-ampd	1.087	12.3
CO_2-343	1.098	13.4
CO_2-m2m	1.119	15.6
CO_2-dmpn	1.150	18.8
CO_2-e2e	1.146	18.4

calculated, and experimental data on the temperature dependence of the lattice constants of amine-appended MOFs are limited in the literature [12]. As the lattice thermal expansion is a measure of the mechanical change of the adsorbents to applied heat, it is relevant to the performance of these systems during the temperature-swing experiments and its overall durability over multiple cycles, and thus understanding thermal expansion is vital to the application of this technology for large-scale extraction of CO_2 from flue gas and beyond. For example, materials with large thermal expansion may not be suitable for adsorption-desorption processes when the regeneration of the adsorbent is temperature driven.

Experimental measurements of the thermal expansion of Mg_2(dobdc) [60], a reduced pore diameter variant of the material under study here, are notable in that they indicate a very small but anisotropic change in lattice parameters up to 500 K. The *a*-*b* plane lattice parameters decrease with temperature while the lattice parameter along the pore direction (*c* direction) increases; the changes in the lattice parameter along the *a*-*b* plane and the *c* direction between 0 to 500 K are -0.096% and 0.248%, respectively. The changes in this quantity are smaller than typical error bars between DFT-predicted lattice parameters and experiment, and therefore capturing and explaining these trends represents a challenge for theory.

Recently, the temperature dependence of the lattice parameters of Zn_2(dobdc) was reported experimentally and theoretically [61]. A statistical analysis of the experimental data using Bayesian methods showed higher uncertainties in the measured lattice parameters [61]. Within this large uncertainty, similar trends as in Mg_2(dobdc) were found experimentally, namely, small negative thermal expansion along the *a*-*b* plane of $(-2 \pm 1) \times 10^{-6}$ K^{-1} and a positive expansion along the *c* direction of $(4 \pm 2) \times 10^{-6}$ K^{-1}.

Prior DFT calculations showed that the thermal expansion is very sensitive to the choice of exchange correlation functional and the approach to calculating the thermal expansion [61]. Using the quasiharmonic Grüneisen theory of thermal expansion [62], using the PBEsol functional and a finite-difference approach with a strain size of 10^{-3}, the measured trends in lattice parameters with temperature were reproduced; however, the same strain size leads to an opposite trend with the original Perdew-Berke-Ernzerhof (PBE) functional. Furthermore, with PBEsol, doubling the strain size to 2×10^{-3} results in an opposite trend. In addition, the thermal expansion was also found to be sensitive to the order of the numerical differentiation. In this case, a second-order differentiation techniques fail to capture the experimental trends, requiring the adoption of a forth-order scheme.

The sensitivity of the thermal expansion to the details of the strain size and the order of the numerical differentiation method may be related to anharmonic effects that are not well captured using small uniform displacements within the quasiharmonic approximation. In order to minimize this sources of error, we compute the Grüneisen parameter tensor from third-order force constants [63] as implemented in the Phono3py package [64] using our NNP. We note that such calculations for a system of this complexity are currently prohibitive at the DFT level and require our NNP. The third-order force constants are extracted from temperature-dependent interatomic force constants [65,66] using the HIPHIVE code [67] and we use Grüneisen theory to compute the thermal expansion. The HIPHIVE software implements a method of computing higher-order interatomic force constants in which the energy is written as a Taylor expansion in atomic displacements whose coefficients correspond to the interatomic force constants. The interatomic force constants are extracted from a linear fit, given a set of atomic displacements in supercells. Temperature effects can be incorporated in to the interatomic force constants by generating atomic displacements with finite-temperature MD [68] displacements along phonon modes at finite temperatures [69] or via random displacements [70]. This approach incorporates both the quantum effects required to describe low-temperature thermal properties and anharmonic effects that are dominant at higher temperatures. High-temperature anharmonic effects are captured through temperature-dependent displacements, which could be missed when uniform displacements of the order of 0.01 Å are used. This value of 0.01 Å can be thought of as corresponding to a root-mean-square displacement at low temperatures.

In what follows, we summarize the expressions used to compute the lattice thermal expansion coefficients for M_2(dobpdc) and its amine-functionalized variants using our approach. Further details can be found in Ref. [61]. The thermal expansion coefficients for hexagonal lattices

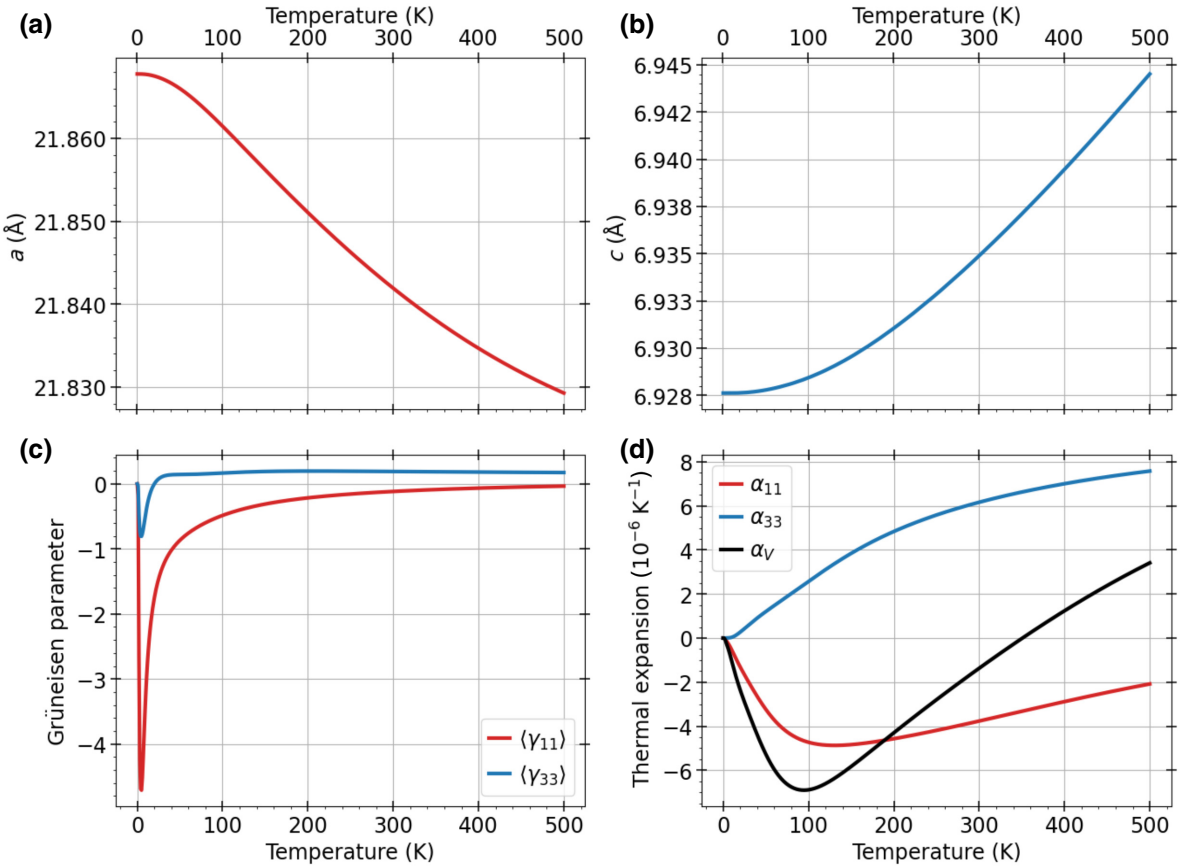

FIG. 13. Thermal properties of Mg_2(dobpdc) computed with the NNP: (a) lattice parameter a (b) lattice parameter c, (c) Grüneisen parameter as a function of temperature (with γ_{11} and γ_{33} the Grüneisen parameters along the a and c directions, respectively), and (d) thermal expansion. The corresponding thermal expansion coefficients are respectively α_{11} and α_{33}. Here $\alpha_V = 2\alpha_{11} + \alpha_{33}$ is the volume thermal expansion coefficient.

[61] along a and c lattice directions are

$$\alpha_{11}(T) = \frac{C_V}{V}[(S_{1111} + S_{1122})\langle\gamma_{11}\rangle(T) + S_{1133})\langle\gamma_{33}\rangle(T)] \tag{8}$$

and

$$\alpha_{33}(T) = \frac{C_V}{V}[2S_{1133}\langle\gamma_{11}\rangle(T) + S_{3333})\langle\gamma_{33}\rangle(T)], \tag{9}$$

where $\alpha_{11} = \alpha_{22}$, S_{ijkl} are the compliance, and $\langle\gamma_{ij}\rangle$ are the average Grüneisen parameters given by

$$\langle\gamma_{ij}\rangle(T) = \frac{\sum_{\mathbf{q},\nu} c_\nu(\mathbf{q},\nu,T)\gamma_{ij}(\mathbf{q},\nu)}{\sum_{\mathbf{q},\nu} c_\nu(\mathbf{q},\nu,T)}, \tag{10}$$

where $c_\nu(\mathbf{q},\nu,T)$ is the heat capacity at wave vector $\mathbf{q}$ and phonon branch ν and $C_V(T) = (1/N_q)\sum_{\mathbf{q},\nu} c_\nu(\mathbf{q},\nu,T)$.

In the following, we present our results on the thermal expansion of Mg_2(dobpdc), m2m-Mg_2(dobpdc), and CO2-m2m-Mg_2(dobpdc). All displacements are generated according to the harmonic phonon eigenmodes obtained at finite temperatures [69]. We use supercells with these displacements to construct temperature-dependent interatomic force constants. As with the vibrational properties, we use $1 \times 1 \times 3$ supercells for all systems.

In Figs. 13(a)–13(d), we show the lattice parameters, Grüneisen parameters, and thermal expansion coefficients as a function of temperature computed with our NNP for Mg_2(dobpdc). We find that the value of a decreases with temperature below 100 K but increases for higher temperatures; and we find that c monotonically increases with temperature. At 300 K, thermal expansion coefficients of -3.766×10^{-6} and 6.162×10^{-6} K^{-1} along a and c, respectively, are obtained from our NNP, corresponding to an overall volume thermal expansion of -1.37×10^{-6} K^{-1}. Similarly for Mg_2(dobdc), we obtain a thermal expansion coefficient of 0.106×10^{-6} K^{-1} along a and 5.124×10^{-6} K^{-1} along c (see Fig. S11 in the Supplemental Material [43]). The results for Mg_2(dobpdc) and Mg_2(dobdc) follow the experimental trends observed for Mg_2(dobdc) [60] and Zn_2(dobdc) [61]. Notably, these trends are not captured from our calculations when we

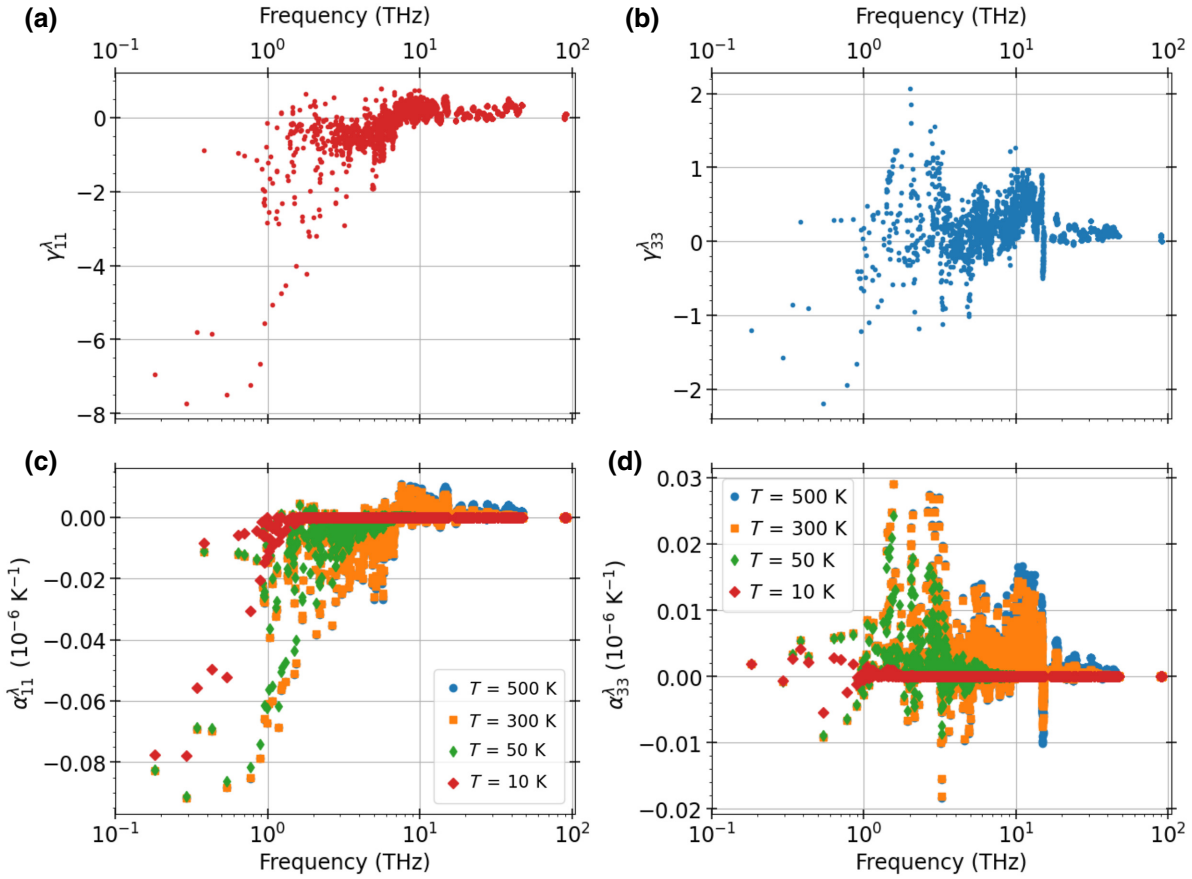

FIG. 14. Mode-resolved Grüneisen parameters and thermal expansion computed from the NNP. (a) Grüneisen parameter along a, (b) Grüneisen parameter along c, (c) mode-resolved thermal expansion along a, and (d) mode-resolved thermal expansion along c.

construct the interatomic force constants (IFCs) from displacements at 1 K, as shown in Fig. S9 in the Supplemental Material [43]. The mean square displacement at this temperature is consistent with about 0.01 Å, which is typically used to extract IFCs using finite-difference-based methods for computing Grüneisen parameters. However, the trends are consistently captured when IFCs generated at temperatures above 50 K are used (see Fig. S10 in the Supplemental Material [43]). These observations stress again the importance of constructing IFCs from finite-temperature displacements, which requires an accurate and computationally efficient model such as our NNP.

We can understand the microscopic origin of the small thermal expansion in Mg_2(dobpdc) by examining the mode-resolved Grüneisen parameter and thermal expansion as a function of temperature. Figures 14(a)–14(d) show the mode-resolved Gruneisen parameters and thermal expansion coefficients along a- and c-crystallographic axes. At 10 K, only vibrational modes below 1 THz contribute to the thermal expansion. Because of the presence of low-frequency modes corresponding to the organic ligands, the magnitude of $\langle\gamma_{11}\rangle$ is larger than $\langle\gamma_{33}\rangle$ and hence the corresponding thermal expansion α_{11} is larger at low temperatures. At 50 K, vibrational modes up to 10 THz are activated. Subsequently, contributions to the in-plane Grüneisen parameter go to zero and the corresponding thermal expansion is nearly constant. On the other hand, α_{33} continues to increase as a function of temperature due to the presence of nonzero contributions from modes above 10 THz to $\langle\gamma_{33}\rangle$ that become activated at higher temperatures. The combined effects of negative thermal expansion in-plane and positive thermal expansion along the c axis results in a negligibly small volume thermal expansion. The macroscopic origin of the small thermal expansion coefficients s is consistent with what has previously been discussed for Zn_2(dobdc) systems [61]. Although the individual compliances S_{ijkl} are larger than those of Zn_2(dobdc) systems, the magnitudes of $S_{1111} + S_{1122} = 0.052$ GPa^{-1}, $2S_{1133} = -0.010$ GPa^{-1}, and $S_{3333} = 0.055$ GPa^{-1} are similar.

Building on the success of the approach in reproducing the experimental trends in Mg_2(dobdc) and Mg_2(dobpdc), we now compute the thermal expansion coefficients for m2m-Mg_2(dobpdc) and CO_2-m2m-Mg_2(dobpdc). Figures 15(a)–15(d) show relevant quantities for m2m-Mg_2(dobpdc). The thermal expansion of

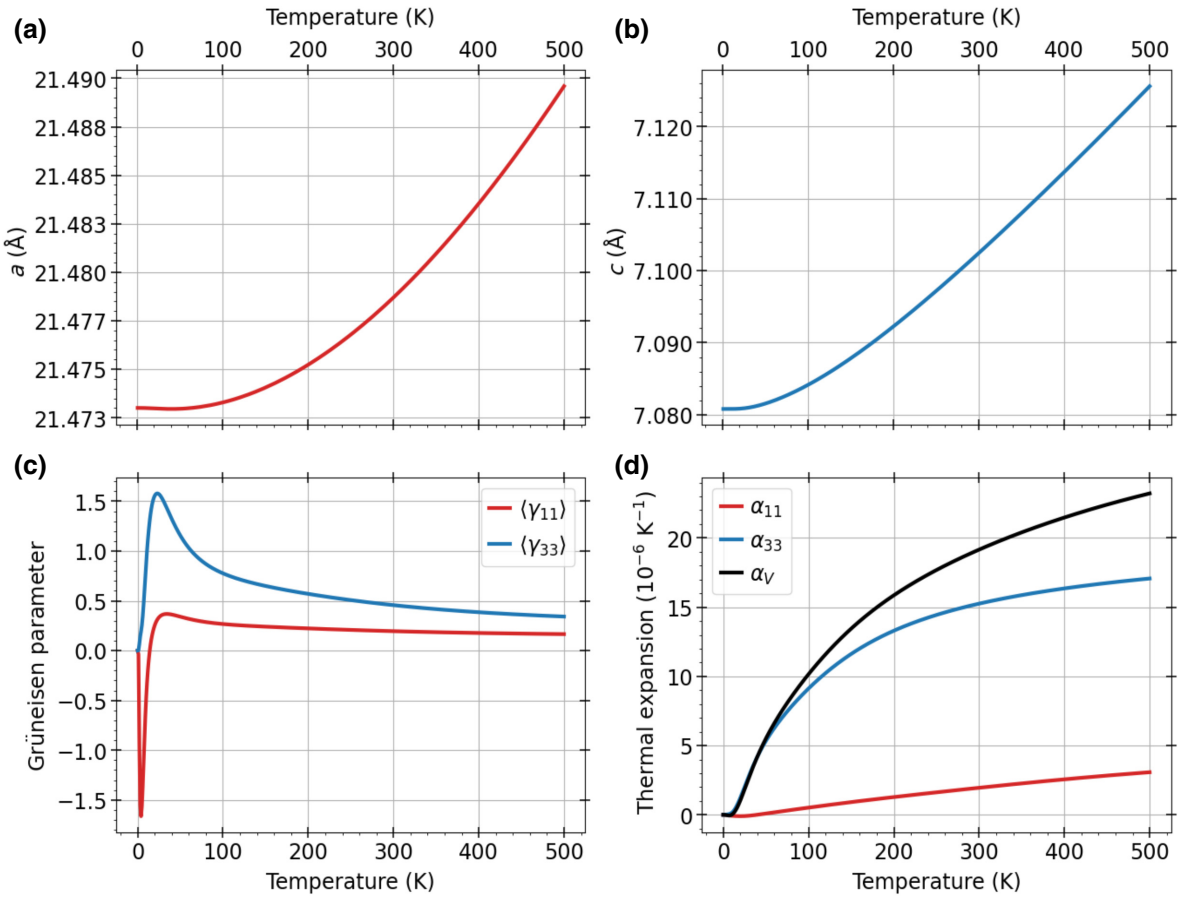

FIG. 15. Thermal properties of m2m-Mg_2(dobpdc) computed with our NNP: (a) lattice parameter a, (b) lattice parameter c, (c) Grüneisen parameter as a function of temperature (with γ_{11} and γ_{33} the Grüneisen parameters along the a and c directions, respectively), and (d) thermal expansion. The corresponding thermal expansion coefficients are respectively α_{11} and α_{33}. Here $\alpha_V = 2\alpha_{11} + \alpha_{33}$ is the volume thermal expansion coefficient.

m2m-Mg_2(dobpdc) is positive for both a and c, leading to a large volume expansion for the system with an amine appendage. The coefficients α_{11} and α_{33} at 300 K are respectively 1.959×10^{-6} and 15.25×10^{-6} K^{-1} along c. Our computed thermal expansion coefficient along c for m2m-Mg_2(dobpdc) has the same positive sign as for the unfunctionalized MOF, and thus overall m2m-Mg_2(dobpdc) expands upon heating, in contrast to the bare Mg_2(dobpdc). In the a crystallographic direction, the thermal expansion coefficient of m2m-Mg_2(dobpdc) is smaller than that of the unfunctionalized system, while along the c direction we find larger and positive expansion coefficients.

In Figs. 16(a)–16(d) we show the thermal properties of CO_2-bound m2m-Mg_2(dobpdc). The c lattice parameter increases with temperature similar to the trends observed in Mg_2(dobpdc) and m2m-Mg_2(dobpdc) systems; however, here the change of a with temperature is truly negligible. The thermal expansion coefficients of CO_2-bound m2m-Mg_2(dobpdc), α_{11} and α_{33}, at 300 K are -1.314×10^{-6} and 8.556×10^{-6} K^{-1}, respectively. These results indicate that the formation of ammonium carbamate chains upon CO_2 insertion leads to increased thermal stability by constraining the vibration of the organic linkers. The enhancement in the thermal stability of CO_2-bound m2m-Mg_2(dobpdc) over Mg_2(dobpdc) and m2m-Mg_2(dobpdc) originates from small elastic compliances and the reduced contribution of the vibrational modes between 1 and 10 THz to the average Grüneisen parameters and thermal expansion (see the Supplemental Material [43]) in both a and c axes compared to Mg_2(dobpdc) and m2m-Mg_2(dobpdc).

D. Simulated annealing and CO_2 binding energies

As has already been discussed in the previous sections, the binding energy is a key quantity for assessing the CO_2 capture performance of a given MOF. These calculations commonly proceed by performing zero-temperature DFT optimizations of crystal structures for MOFs with and without CO_2. However, because of the complexity of the potential energy surface of amine-appended MOFs, calculations of the adsorption energy can be influenced by the

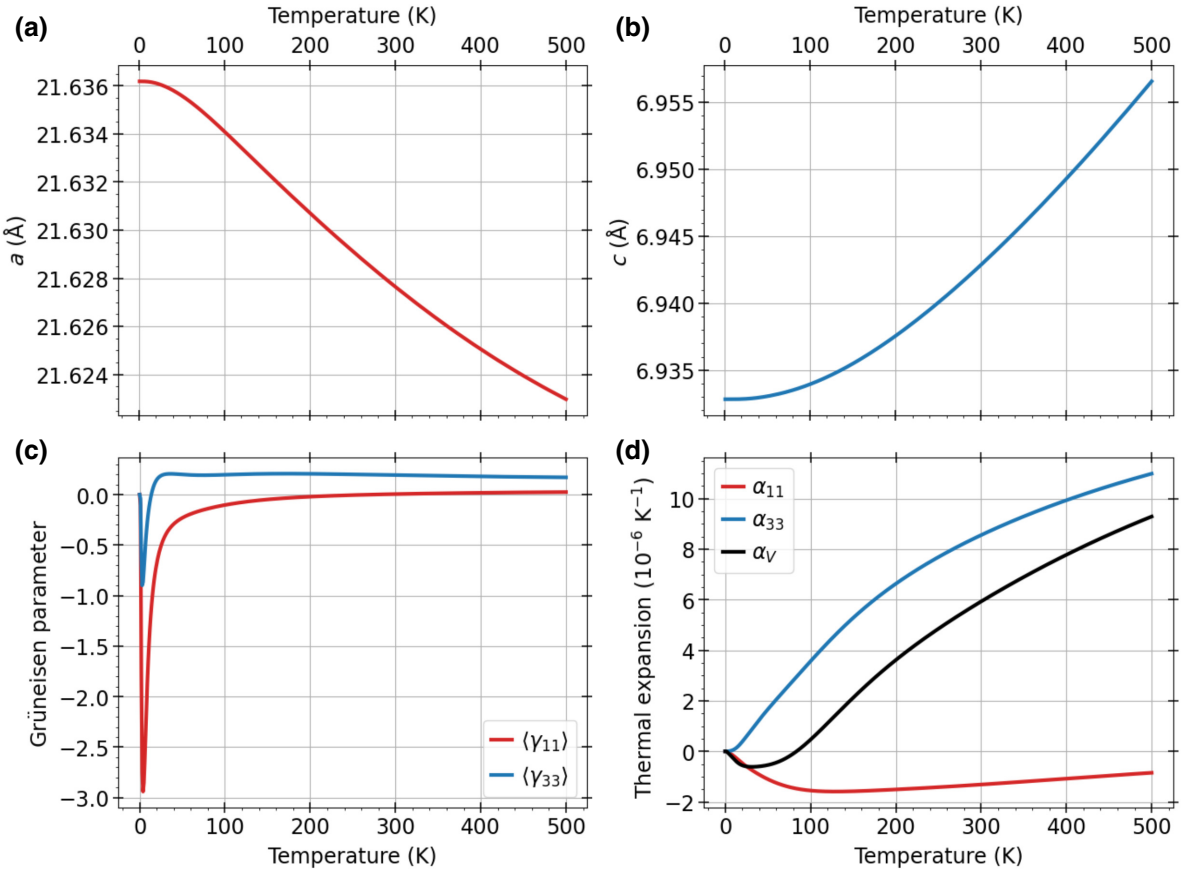

FIG. 16. Thermal properties of CO_2-bound m2m-Mg_2(dobpdc) from our NNP: (a) lattice parameter a (b) lattice parameter c, (c) Grüneisen parameter as a function of temperature (with γ_{11} and γ_{33} the Grüneisen parameters along the a and c directions, respectively), and (d) thermal expansion. The corresponding thermal expansion coefficients are respectively α_{11} and α_{33}. Here $\alpha_V = 2\alpha_{11} + \alpha_{33}$ is the volume thermal expansion coefficient.

initial geometry of the amine appendages, which may relax into local minima with higher energy using conventional optimization algorithms, resulting in an inaccurate prediction for the CO_2 binding energy. In prior DFT calculations, the initial structures are either obtained from experimental single crystal structures, which represent a thermally averaged structure, or by manually inserting the amine and adjusting it until it is as close as possible to the measured powder diffraction patterns. This approach has been used to resolve several experimental structures for these systems [13–15,17,19,47], and accurate adsorption energies have been predicted using this approach. The weakness of this technique is that it relies on the availability of structural information from experiment, limiting its predictive power; additionally, it is not applicable in the absence of experimental information. A theoretical approach capable of predicting high CO_2 capacity amine would assist experimentalists in reducing the number of exploratory measurements.

Here, we demonstrate a simulated annealing procedure with our NNP to generate an ensemble of the diamine conformations in Mg_2(dobpdc) starting from a manually inserted diamine in Mg_2(dobpdc), obviating the need for a starting point biased by experiment or an intuitive guess. In simulated annealing, a molecular dynamics simulation is performed in which the temperature of the initial structure is raised above room temperature and allowed to equilibrate, followed by a slow cooling of the temperature towards 0 K; a structural relaxation is then performed on the final equilibrated structure. While this procedure would be prohibitively expensive with DFT, it can be readily performed with our NNP. As a proof of concept, we demonstrate our approach on dmpn-Mg_2(dobpdc). An experimental crystal structure is lacking for dmpn-Mg_2(dobpdc), and it has a more complicated CO_2 binding geometry than other common diamine-appended MOFs [16]: dmpn-Mg_2(dobpdc) can bind CO_2 via the formation of ammonium carbamate chains, carbamic acid pairs and a mixed species comprises of 1:1 ratio of ammonium carbamate and carbamic acid. The mixed phase structure is observed for Mg_2(dobpdc) while the carbamic acid pair is the thermodynamically stable phase. It has also shown promise in realistic flue gas conditions where SO_2 and water are present in the gas stream [20],

performing better than other diamines for CO_2 capture in this case.

We start with the dmpn-Mg_2(dobpdc) system without CO_2 and perform simulated annealing using NNP. The initial dmpn-Mg_2(dobpdc) was obtained from Ref. [16]; it was generated in analogy with others in which experimentally derived diamine-appended MOFs where x-ray crystallographic data are available and optimized with DFT-vdw-DF2 [71]. We raise the temperature of the system to 400 K, perform a *N-P-T* MD for 10 ps, and then slowly anneal to 10 mK in a total time interval of 40 ps, and finally minimize the total energy of the annealed system.

Figures 17(a) and 17(b) show the initial and final (annealed) structures of dmpn-Mg_2(dobpdc) systems, obtained after a final DFT optimization step. Remarkably, the final structure is 105 kJ/mol (17.5 kJ/mol per diamine) lower in energy than the initial structure. While the initial structure optimized to a dimerization of diamines, the annealing leads to independent conformation of diamines in the pore. The annealed system has a diamine conformation that minimizes steric interactions in the *a-b* plane and maximizes hydrogen bonding between the primary and secondary amines.

In order to further quantify the effect of the starting geometry on the binding energy, we perform simulated annealing on the ammonium carbamate-carbamic acid mixed phase of the CO_2-bound dmpn-Mg_2(dobpdc), the experimental observed species, following the procedure outlined above. The initial and annealed structures are shown in Figs. 17(c) and 17(d). The annealed structure is 6 kJ/mol per CO_2 lower in energy than the starting geometry, and the binding energies, ΔE, before and after annealing are -92.2 and -80.5 kJ/mol, respectively, when evaluated with DFT (see Table III), a significant and experimentally relevant difference. Our *ab initio* NNP predicts -90.5 and -71.5 kJ/mol before and after annealing, respectively, in good agreement with the DFT calculations.

To compare with the measured binding enthalpy, we compute the zero-point and thermal contributions to the enthalpy within the harmonic approximation using vibrational frequencies calculated with our NNP. The zero-point and thermal contributions are calculated as described in Ref. [53]. However, here we relax the fixed framework approximation used in Ref. [53] that assumes that the phonon frequencies of the frameworks do not change upon gas adsorption. Experimentally, the CO_2 binding enthalpy is 70 kJ/mol in the CO_2 concentration range 0–2.8 mmol/g and 53 kJ/mol at 3.2 mmol/g [16]. The binding enthalpy at 1 mmol/g is measured to be -74 ± 1 [14]. As shown in Table III, the binding enthalpy, ΔH obtained with the NNP on the annealed structures is -62.0 kJ/mol (the corresponding DFT value is -71.2 kJ/mol using the NNP vibrational energies). Our predicted binding enthalpies are

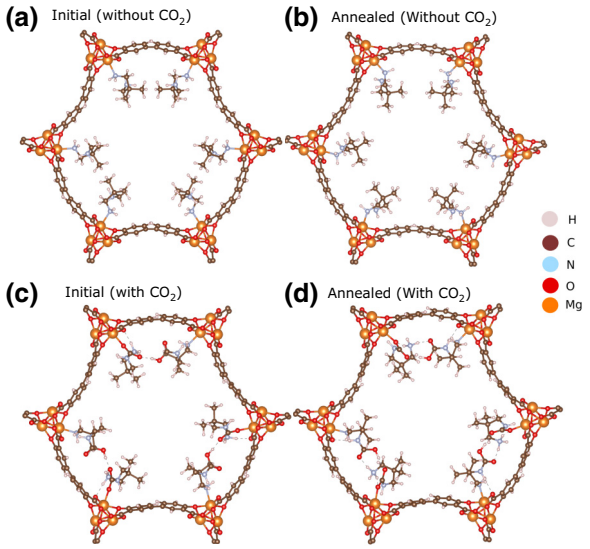

FIG. 17. Structures of dmpn-Mg_2(dobpdc) (a) before annealing and (b) after annealing. Atomic structures of the carbamic acid pair and ammonium mixed phase CO_2 adsorption dmpn-Mg_2(dobpdc) (c) before annealing and (d) after annealing. The initial structures are taken from Ref [47].

TABLE III. Lattice parameters before and after annealing of dmpn-Mg_2(dobpdc) with our NNP. Energies are in kJ/mol. Zero-point energy (ZPE) and thermal energy (TE) corrections are also indicated. Here ΔE is the CO_2 binding energy while $\Delta H = \Delta E + \text{ZPE} + \text{TE}$ is the CO_2 binding enthalpy. The different adsorption enthalpies from Ref. [16] correspond to CO_2 concentrations of 0–2.8 and 3.2 mmol/g.

		Without CO_2		With CO_2					
		a (Å)	c (Å)	a (Å)	c (Å)	ΔE	ZPE	TE	ΔH
DFT	Initial	21.605	6.803	21.340	7.194	−92.2			
	Annealed	21.643	6.864	21.354	6.927	−80.5			
NNP	Initial	21.855	6.974	21.568	7.115	−90.5	13.24	−2.22	−79.5
	Annealed	21.735	6.991	21.489	6.928	−71.3	10.52	−1.26	−62.0
Expt.		21.624[a]	6.930[a]	21.473[b]	6.916[b]	...	...	...	−74±1[a],−70[b], −53[b]

[a]Ref. [14].
[b]Ref. [16].

within adequate range of the experimental values, demonstrating the value of simulated annealing for generating structures of amine-appended MOFs as well as testing the accuracy of our NNP.

IV. CONCLUSIONS

In this work, we have developed a neural network interatomic potential for amine-appended Mg_2(dobpdc) with the accuracy of a state-of-the-art vdW-corrected periodic DFT calculation. We demonstrated that our NNP can accurately capture temperature-dependent vibrational properties, lattice thermal expansion, and heat capacities in amine-appended Mg_2(dobpdc), establishing it as suitable for future studies of the dynamical and thermal properties of CO_2 binding in amine-appended Mg_2(dobpdc) to guide further optimization of the MOF for carbon capture applications. To develop our NNP, we used a Behler-Parrinello framework and a training dataset generated via active learning, leading to an interatomic potential that accurately reproduces the energies and atomic forces. While DFT calculations are too computationally expensive to be used to study CO_2 dynamics in amine-appended MOFs, our DFT-derived NNPs have near DFT accuracy with a cost more akin to that of classical MD calculations, enabling accurate calculations of long-time-scale finite-temperature simulations beyond those that would be possible with purely DFT-based approaches.

We have shown that our NNP is accurate in reproducing the DFT structural and mechanical properties of a wide variety of amine-appended MOFs. The adsorption energies are also well reproduced with only a few amines having a deviation above 10%. For those amines, we found that the discrepancies between the DFT prediction and NNP were mainly due to a lack of error cancelation between the fragments that constitute the adsorption energy. To illustrate the nature of the error cancelation, we defined a measure of the degree of error cancelation, δ, and found a perfect correlation between δ and the NNP prediction error on the adsorption energy. Moreover, we showed that the NNP can be used to understand the roles of different key vibrational modes in the thermal properties of amine-appended systems. We found that the low-energy modes correspond to either a collective strain of the linkers or collective rotations around the MgO clusters while the very high energy vibrations correspond to C—H and N—H stretching. Our NNP accurately describes N—H bonding stretching arising from a protonated amine terminal upon CO_2 insertion. The lowering of the energy of this mode with CO_2 insertion is consistent with the fact that the N—H bond of a charged amine terminal is weaker. We computed the heat capacity of amine-appended Mg_2(dobpdc) MOFs, and quantified the manner in which more flexible amines give rise to higher heat capacity. At a given temperature however, the variation of the heat capacity across different amines is small. Furthermore, we used our NNP to compute finite-temperature-dependent third-order force constants and lattice thermal expansion in these materials, incorporating both quantum zero-point effects as well as the anharmonic effects that are prevalent at high temperature. Our results for Mg_2(dobpdc) follow the same trends observed for Mg_2(dobdc), validating the accuracy of the NNP. Moreover, we predicted an enhancement in the thermal properties of CO_2-bound diamine-appended Mg_2(dobpdc) systems. Finally, we demonstrated that our NNP can be used to obtained low-energy structures starting from a handcrafted diamine-appended system using the simulated annealing approach. This approach can be used to study CO_2 adsorption in amine-appended MOFs even in the absence of experimental crystal structures.

The accuracy and transferability of our NNP for finite-temperature properties of these systems is an indication that the potential can be reliably used to predict and understand dynamical properties of carbon capture adsorbents. The potential can be used for a high throughput screening of the properties of several yet to be synthesized amine to discover amine with improved carbon capture performance. This NPP will enable study of the thermal conductivity of these materials as well as the kinetics of CO_2 insertion and subsequent carbamate formation within amine-appended Mg_2(dobpdc). In addition, it may be used to search for yet to be synthesized amine and classify them according to their energetics and thermal properties.

ACKNOWLEDGMENTS

This work is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Award No. DE-SC0019992. E.T. is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Award No. DE-AC02-05CH11231. Computational resources were provided by the National Energy Research Scientific Computing Center, USA.

[1] Greenhouse gas emissions from energy: Overview, IEA, Paris (2021). https://www.iea.org/reports/greenhouse-gas-emissions-from-energy-overview.

[2] Global energy review: CO2 emissions in 2021, IEA, Paris (2022). https://www.iea.org/reports/global-energy-review-co2-emissions-in-2021-2.

[3] A. Samanta, A. Zhao, G. K. H. Shimizu, P. Sarkar, and R. Gupta, Post-combustion CO_2 capture using solid sorbents: A review, Ind. Eng. Chem. Res. **51**, 1438 (2012).

[4] M. N. Dods, S. C. Weston, and J. R. Long, Prospects for simultaneously capturing carbon dioxide and harvesting water from air, Adv. Mater. **34**, 2204277 (2022).

[5] G. T. Rochelle, Amine scrubbing for CO_2 capture, Science **325**, 1652 (2009).

[6] P. Markewitz, W. Kuckshinrichs, W. Leitner, J. Linssen, P. Zapp, R. Bongartz, A. Schreiber, and T. E. Müller, Worldwide innovations in the development of carbon capture technologies and the utilization of CO_2, Energy Environ. Sci. **5**, 7281 (2012).

[7] F. Zeman and K. Lackner, Capturing carbon dioxide directly from the atmosphere, World Res. Rev. **16**, 157 (2004).

[8] S. A. Mazari, B. Si Ali, B. M. Jan, I. M. Saeed, and S. Nizamuddin, An overview of solvent management and emissions of amine-based CO_2 capture technology, Int. J. Greenh. Gas Control **34**, 129 (2015).

[9] T. C. Drage, C. E. Snape, L. A. Stevens, J. Wood, J. Wang, A. I. Cooper, R. Dawson, X. Guo, C. Satterley, and R. Irons, Materials challenges for the development of solid sorbents for post-combustion carbon capture, J. Mater. Chem. **22**, 2815 (2012).

[10] H. Furukawa, K. E. Cordova, M. O'Keeffe, and O. M. Yaghi, The chemistry and applications of metal-organic frameworks, Science **341**, 1230444 (2013).

[11] T. M. McDonald, W. R. Lee, J. A. Mason, B. M. Wiers, C. S. Hong, and J. R. Long, Capture of carbon dioxide from air and flue gas in the alkylamine-appended metal-organic framework mmen-mg2(dobpdc), J. Am. Chem. Soc. **134**, 7056 (2012).

[12] T. M. McDonald, J. A. Mason, X. Kong, E. D. Bloch, D. Gygi, A. Dani, V. Crocellà, F. Giordanino, S. O. Odoh, and W. S. Drisdell, *et al.*, Cooperative insertion of CO_2 in diamine-appended metal-organic frameworks, Nature **519**, 303 (2015).

[13] R. L. Siegelman, T. M. McDonald, M. I. Gonzalez, J. D. Martell, P. J. Milner, J. A. Mason, A. H. Berger, A. S. Bhown, and J. R. Long, Controlling cooperative CO_2 adsorption in diamine-appended mg2(dobpdc) metal-organic frameworks, J. Am. Chem. Soc. **139**, 10526 (2017).

[14] P. J. Milner, R. L. Siegelman, A. C. Forse, M. I. Gonzalez, T. Runčevski, J. D. Martell, J. A. Reimer, and J. R. Long, A diaminopropane-appended metal-organic framework enabling efficient CO_2 capture from coal flue gas via a mixed adsorption mechanism, J. Am. Chem. Soc. **139**, 13541 (2017).

[15] P. J. Milner, J. D. Martell, R. L. Siegelman, D. Gygi, S. C. Weston, and J. R. Long, Overcoming double-step CO_2 adsorption and minimizing water co-adsorption in bulky diamine-appended variants of mg2(dobpdc), Chem. Sci. **9**, 160 (2018).

[16] A. C. Forse, P. J. Milner, J.-H. Lee, H. N. Redfearn, J. Oktawiec, R. L. Siegelman, J. D. Martell, B. Dinakar, L. B. Zasada, and M. I. Gonzalez, *et al.*, Elucidating CO_2 chemisorption in diamine-appended metal-organic frameworks, J. Am. Chem. Soc. **140**, 18016 (2018).

[17] J.-H. Lee, R. L. Siegelman, L. Maserati, T. Rangel, B. A. Helms, J. R. Long, and J. B. Neaton, Enhancement of CO_2 binding and mechanical properties upon diamine functionalization of m2(dobpdc) metal-organic frameworks, Chem. Sci. **9**, 5197 (2018).

[18] H. Babaei, J.-H. Lee, M. N. Dods, C. E. Wilmer, and J. R. Long, Enhanced thermal conductivity in a diamine-appended metal-organic framework as a result of cooperative CO_2 adsorption, ACS Appl. Mater. Interfaces **12**, 44617 (2020).

[19] R. L. Siegelman, P. J. Milner, A. C. Forse, J.-H. Lee, K. A. Colwell, J. B. Neaton, J. A. Reimer, S. C. Weston, and J. R. Long, Water enables efficient CO_2 capture from natural gas flue emissions in an oxidation-resistant diamine-appended metal-organic framework, J. Am. Chem. Soc. **141**, 13171 (2019).

[20] S. T. Parker, A. Smith, A. C. Forse, W.-C. Liao, F. Brown-Altvater, R. L. Siegelman, E. J. Kim, N. A. Zill, W. Zhang, and J. B. Neaton, *et al.*, Evaluation of the stability of diamine-appended mg2(dobpdc) frameworks to sulfur dioxide, J. Am. Chem. Soc. **144**, 19849 (2022).

[21] E. J. Kim, R. L. Siegelman, H. Z. H. Jiang, A. C. Forse, J.-H. Lee, J. D. Martell, P. J. Milner, J. M. Falkowski, J. B. Neaton, and J. A. Reimer, *et al.*, Cooperative carbon capture and steam regeneration with tetraamine-appended metal-organic frameworks, Science **369**, 392 (2020).

[22] A. Zoelle, D. Keairns, L. L. Pinkerton, M. J. Turner, M. Woods, N. Kuehn, V. Shah, and V. Chou, Cost and performance baseline for fossil energy plants volume 1: Bituminous coal and natural gas to electricity revision 3 (2015).

[23] S. L. Mayo, B. D. Olafson, and W. A. Goddard, Dreiding: A generic force field for molecular simulations, J. Phys. Chem. **94**, 8897 (1990).

[24] J. D. Martell, P. J. Milner, R. L. Siegelman, and J. R. Long, Kinetics of cooperative CO_2 adsorption in diamine-appended variants of the metal-organic framework mg2(dobpdc), Chem. Sci. **11**, 6457 (2020).

[25] M. Raju, P. Ganesh, P. R. C. Kent, and A. C. T. van Duin, Reactive force field study of Li/C systems for electrical energy storage, J. Chem. Theory Comput. **11**, 2156 (2015).

[26] Y. Shaidu, E. Küçükbenli, and S. de Gironcoli, Lithium adsorption on graphene at finite temperature, J. Phys. Chem. C **122**, 20800 (2018).

[27] D. Smykowski, B. Szyja, and J. Szczygieł, GCMC simulations of CO_2 adsorption on zeolite-supported Ir4 clusters, J. Mol. Graphics Modell. **50**, 35 (2014).

[28] R. Mercado, B. Vlaisavljevich, L.-C. Lin, K. Lee, Y. Lee, J. A. Mason, D. J. Xiao, M. I. Gonzalez, M. T. Kapelewski, and J. B. Neaton, *et al.*, Force field development from periodic density functional theory calculations for gas separation applications using metal-organic frameworks, J. Phys. Chem. C **120**, 12590 (2016).

[29] J. Xu, Y. M. Liu, A. S. Lipton, J. Ye, G. L. Hoatson, P. J. Milner, T. M. McDonald, R. L. Siegelman, A. C. Forse, B. Smit, J. R. Long, and J. A. Reimer, Amine dynamics in diamine-appended mg2(dobpdc) metal-organic frameworks, J. Phys. Chem. Lett. **10**, 7044 (2019).

[30] S. Cheung, W.-Q. Deng, A. C. T. van Duin, and W. A. Goddard, Reaxffmgh reactive force field for magnesium hydride systems, J. Phys. Chem. A **109**, 851 (2005).

[31] L. Huang, K. L. Joshi, A. C. T. van Duin, T. J. Bandosz, and K. E. Gubbins, ReaxFF molecular dynamics simulation of thermal stability of a $Cu_3(BTC)_2$ metal-organic framework, Phys. Chem. Chem. Phys. **14**, 11327 (2012).

[32] R. Sabatini, T. Gorni, and S. de Gironcoli, Nonlocal van der Waals density functional made simple and efficient, Phys. Rev. B **87**, 041108 (2013).

[33] S. Bureekaew, S. Amirjalayer, M. Tafipolsky, C. Spickermann, T. K. Roy, and R. Schmid, MOF-FF—a

flexible first-principles derived force field for metal-organic frameworks, phys. Status Solidi (b) **250**, 1128 (2013).
[34] J. K. Bristow, D. Tiana, and A. Walsh, Transferable force field for metal-organic frameworks from first-principles: BTW-FF, J. Chem. Theory Comput. **10**, 4644 (2014).
[35] L. Vanduyfhuys, S. Vandenbrande, T. Verstraelen, R. Schmid, M. Waroquier, and V. Van Speybroeck, Quickff: A program for a quick and easy derivation of force fields for metal-organic frameworks from ab initio input, J. Comput. Chem. **36**, 1015 (2015).
[36] J. K. Bristow, J. M. Skelton, K. L. Svane, A. Walsh, and J. D. Gale, A general forcefield for accurate phonon properties of metal-organic frameworks, Phys. Chem. Chem. Phys. **18**, 29316 (2016).
[37] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, and I. Dabo, *et al.*, QUANTUM ESPRESSO: A modular and open-source software project for quantum simulations of materials, J. Phys.: Condens. Matter **21**, 395502 (2009).
[38] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, and M. Cococcioni, *et al.*, Advanced capabilities for materials modelling with Quantum ESPRESSO, J. Phys.: Condens. Matter **29**, 465901 (2017).
[39] R. Lot, F. Pellegrini, Y. Shaidu, and E. Küçükbenli, PANNA: Properties from artificial neural network architectures, Comput. Phys. Commun. **256**, 107402 (2020).
[40] S. Plimpton, Fast parallel algorithms for short-range molecular dynamics, J. Comput. Phys. **117**, 1 (1995).
[41] J. Behler and M. Parrinello, Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces, Phys. Rev. Lett. **98**, 146401 (2007).
[42] J. S. Smith, O. Isayev, and A. E. Roitberg, ANI-1: An extensible neural network potential with dft accuracy at force field computational cost, Chem. Sci. **8**, 3192 (2017).
[43] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PRXEnergy.2.023005 for Lewis structures of amines not used to generate the training dataset; an energy-volume plot of the training dataset as a function of active learning iterations; performance of NNP with architecture 20:20:1 at first, second, and third active learning iterations on the VS_1 dataset; comparison of the accuracy of NNP with architectures 20:20:1 and 64:64:1 on the VS_3 dataset; structural and mechanical properties; the potential energy surface of CO_2 binding to the unfunctionalized Mg site; vibrational properties of bare Mg_2(dobpdc); details of the method employed to compute thermal properties; thermal expansion of unfunctionalized Mg_2(dobpdc) using force constants obtained with displacements at 1 K; thermal properties of Mg_2(dobpdc) for force constants generated at different temperatures; thermal properties of Mg_2(dobdc); Grüneisen parameters of m2m-Mg_2(dobpdc) with and without CO_2.
[44] P. E. Blöchl, Projector augmented-wave method, Phys. Rev. B **50**, 17953 (1994).
[45] J. P. Perdew and W. Yue, Accurate and simple density functional for the electronic exchange energy: Generalized gradient approximation, Phys. Rev. B **33**, 8800 (1986).
[46] O. A. Vydrov and T. V. Voorhis, Nonlocal van der Waals density functional: The simpler the better, J. Chem. Phys. **133**, 244103 (2010).
[47] J.-H. Lee, P. Hyldgaard, and J. B. Neaton, An assessment of density functionals for predicting CO_2 adsorption in diamine-functionalized metal-organic frameworks, J. Chem. Phys. **156**, 154113 (2022).
[48] H. J. Monkhorst and J. D. Pack, Special points for Brillouin-zone integrations, Phys. Rev. B **13**, 5188 (1976).
[49] N. Marzari, D. Vanderbilt, A. De Vita, and M. C. Payne, Thermal Contraction and Disordering of the Al(110) Surface, Phys. Rev. Lett. **82**, 3296 (1999).
[50] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, ArXiv:1412.6980 (2014).
[51] Y. Shaidu, E. Küçükbenli, R. Lot, F. Pellegrini, E. Kaxiras, and S. de Gironcoli, A systematic approach to generating accurate neural network potentials: The case of carbon, npj Comput. Mater. **7**, 52 (2021).
[52] R. Hill, The elastic behaviour of a crystalline aggregate, Proc. Phys. Soc. Section A **65**, 349 (1952).
[53] K. Lee, J. D. Howe, L.-C. Lin, B. Smit, and J. B. Neaton, Small-molecule adsorption in open-site metal-organic frameworks: A systematic density functional theory study for rational design, Chem. Mater. **27**, 668 (2015).
[54] W. Zhang, H. Liu, Y. Sun, J. Cakstins, C. Sun, and C. E. Snape, Parametric study on the regeneration heat requirement of an amine-based solid adsorbent process for post-combustion carbon capture, Appl. Energy **168**, 394 (2016).
[55] S. M. Moosavi, B. Á. Novotny, D. Ongari, E. Moubarak, M. Asgari, Ö. Kadioglu, C. Charalambous, A. Ortega-Guerrero, A. H. Farmahini, and L. Sarkisov, *et al.*, A data-science approach to predict the heat capacity of nanoporous materials, Nat. Mater. **21**, 1419 (2022).
[56] A. Togo and I. Tanaka, First principles phonon calculations in materials science, Scr. Mater. **108**, 1 (2015).
[57] A. Carreras, phonoLAMMPS: A python interface for LAMMPS phonon calculations using phonopy (2021), https://zenodo.org/record/5668319.
[58] S. G. Subraveti, S. Roussanaly, R. Anantharaman, L. Riboldi, and A. Rajendran, How much can novel solid sorbents reduce the cost of post-combustion CO_2 capture? A techno-economic investigation on the cost limits of pressure-vacuum swing adsorption, Appl. Energy **306**, 117955 (2022).
[59] P. Canepa, K. Tan, Y. Du, H. Lu, Y. J. Chabal, and T. Thonhauser, Structural, elastic, thermal, and electronic responses of small-molecule-loaded metal-organic framework materials, J. Mater. Chem. A **3**, 986 (2015).
[60] W. L. Queen, C. M. Brown, D. K. Britt, P. Zajdel, M. R. Hudson, and O. M. Yaghi, Site-specific CO_2 adsorption and zero thermal expansion in an anisotropic pore network, J. Phys. Chem. C **115**, 24915 (2011).
[61] T. Kamencek, B. Schrode, R. Resel, R. Ricco, and E. Zojer, Understanding the origin of the particularly small and anisotropic thermal expansion of MOF-74, Adv. Theory Simul. **5**, 2200031 (2022).
[62] E. Grüneisen, Theorie des festen zustandes einatomiger elemente, Ann. Phys. **344**, 257 (1912).

[63] J. Fabian and P. B. Allen, Thermal Expansion and GrÜneisen Parameters of Amorphous Silicon: A Realistic Model Calculation, Phys. Rev. Lett. **79**, 1885 (1997).
[64] A. Togo, L. Chaput, and I. Tanaka, Distributions of phonon lifetimes in Brillouin zones, Phys. Rev. B **91**, 094306 (2015).
[65] E. J. Candes and M. B. Wakin, An introduction to compressive sampling, IEEE Signal Process. Mag. **25**, 21 (2008).
[66] F. Zhou, W. Nielson, Y. Xia, and V. Ozoliš, Lattice Anharmonicity and Thermal Conductivity from Compressive Sensing of First-Principles Calculations, Phys. Rev. Lett. **113**, 185501 (2014).
[67] F. Eriksson, E. Fransson, and P. Erhart, The hiphive package for the extraction of high-order force constants by machine learning, Adv. Theory Simul. **2**, 1800184 (2019).
[68] J. Brorsson, A. Hashemi, Z. Fan, E. Fransson, F. Eriksson, T. Ala-Nissila, A. V. Krasheninnikov, H.-P. Komsa, and P. Erhart, Efficient calculation of the lattice thermal conductivity by atomistic simulations with ab initio accuracy, Adv. Theory Simul. **5**, 2100217 (2022).
[69] D. West and S. K. Estreicher, First-Principles Calculations of Vibrational Lifetimes and Decay Channels: Hydrogen-Related Modes in Si, Phys. Rev. Lett. **96**, 115504 (2006).
[70] E. Fransson, F. Eriksson, and P. Erhart, Efficient construction of linear models in materials modeling and applications to force constant expansions, npj Comput. Mater. **6**, 135 (2020).
[71] K. Lee, E. D. Murray, L. Kong, B. I. Lundqvist, and D. C. Langreth, Higher-accuracy van der Waals density functional, Phys. Rev. B **82**, 081101 (2010).