

Dalton Transactions

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

Coarse Graining of Force Fields for Metal-Organic Frameworks

Johannes P. Dürholt¹, Raimondas Galvelis² and Rochus Schmid^{1*}

¹ Computational Materials Chemistry group,
Lehrstuhl Anorganische Chemie 2, Ruhr-Universität
Bochum, Bochum, Germany

² RIKEN Theoretical Molecular Science Laboratory,
2-1 Hirosawa, Wako, Saitama, Japan

* rochus.schmid@rub.de

Abstract

We have adapted our Genetic Algorithm based optimization approach, originally developed to generate force field parameters from quantum mechanic reference data, to derive a first coarse grained force field for a MOF, taking the atomistic MOF-FF as a reference. On the example of the copper paddle-wheel based HKUST-1, a maximally coarse grained model, using a single bead for each three and four coordinated vertex, was developed as a proof of concept. By adding non-bonded interactions with a modified Buckingham potential, the resulting MOF-FF-CGNB is able to predict local deformation energies of the building blocks as well as bulk properties like the **tbo** vs. **pto** energy difference or elastic constants in a semi-quantitative way. As expected, the negative thermal expansion of HKUST-1 is not reproduced by the maximally coarse grained model. At the expense of atomic resolution, substantially larger systems (up to tens of nanometers in size) can be simulated with respect to structural and mechanical properties, bridging the gap to the mesoscale. As an example the deformation of the [111] surface of HKUST-1 by a “tip” could be computed without artifacts from periodic images.

1 Introduction

The swiftly growing class of porous metal-organic frameworks (MOFs) is of high interest for various applications, ranging from gas storage and separation to sensing and catalysis.¹⁻⁶ One of the reasons for this is the wide variety of organic and inorganic building units that can be combined, leading to a range of network topologies and pore geometries.⁷ A further, very important advantage over the well established zeolite-type porous materials is

their flexibility. Especially the sometimes large volume change upon guest adsorption, which is called “breathing”, has triggered a lot of research on this phenomenon.⁸ For a further – ideally rational – development of MOFs towards technical application, an atomistic understanding of the underlying mechanisms of interaction and conformational flexibility is pivotal, making theoretical calculations a valuable and important tool.⁹⁻¹¹ However, many MOF systems push quantum mechanic (QM) methods like periodic density functional theory (DFT) type calculations¹² to their limit because of the size and number of atoms in the unit cell. As a consequence computationally more efficient parameterized force fields (FFs) have been employed where possible.^{11,13} As long as covalent bond breaking is not involved, these more approximate molecular mechanics (MM) methods can be used to investigate larger systems for longer timescales. The difficulty here is to define a proper energy expression to describe the relevant part of the potential energy surface (PES) in the correct way and to determine the corresponding parameters. For computing the conformational flexibility of the porous MOF matrix (as well as for the host-guest interactions, dominated by physisorption) conventional non-reactive force fields, that employ a separation in bonded and non-bonded terms, is a sufficiently good approximation. However, the determination of parameters for the inorganic building blocks remains a challenging problem. A number of system specific force fields for particular MOFs have been developed meanwhile.¹⁴⁻¹⁷ More general force fields, which are able to describe various MOF families in the same way are still scarce. A frequently employed solution is to use so-called generic force fields, like UFF,¹⁸ where the parameters are generated by a rule based system from a much smaller set of atomic parameters. This allows the consistent treatment of a wide range of systems, including a number of MOFs, however, with a limited and uncontrolled accuracy. Very recently, the UFF atomic parameter set was augmented by the Heine group with specific environments for a series of MOFs, referred to as UFF4MOF, which significantly improves the accuracy in this area.¹⁹ A complementary approach has been derived by us over the years.²⁰⁻²² By a Genetic Algorithm (GA) based global optimization scheme using a novel objective function, specific parameters for each indi-

vidual MOF system are generated in a systematic way, using reference data computed on DFT level for non-periodic model systems. By trading transferability for accuracy, force field parameters with controlled accuracy can be generated, however, with a larger effort as compared to the rule based generic force fields. The resulting force field MOF-FF is available for a range of MOF systems and will be constantly enlarged.²² Very recently a similar, but technically different approach has been introduced by Vanduyfhuys *et al.*²³ In the QuickFF method no global optimization of a target function is used but the force field parameters are directly derived from the QM potential energy surface.

Despite these advancements in the computationally accessible range of length and time scales, achieved by atomistic force fields, system sizes are still limited to a few unit cells, especially for larger MOFs. For a full understanding of the “breathing” effect of MOF crystallites, which seems e.g. to be affected by morphology,²⁴ it might be necessary to go beyond the periodic boundary approximation.²⁵ Recently, indications for meso-porosity in defect-engineered MOFs were found.²⁶ Thus, in order to achieve a meso-scale modeling of MOFs computational approaches for even larger time and length scales are desirable. This is in analogy to the field of biomolecular and polymer systems simulation. There, a further “coarse graining” (CG) is employed and multiple atoms are mapped on single interaction sites, which are often referred to as beads.^{27–30} In the simplest case the hydrogen atoms of e.g. a $-CH_3$ group are mapped on the carbon atom, leading to an effective methyl group potential in a so-called united atom force field.^{31,32} Such FFs have already been employed for MOFs, in particular for describing guest molecules. However, a true coarse graining of the MOF matrix, combining the inorganic building blocks and linkers into one or more effective interaction sites, has to our knowledge not been investigated. Note, however, that the approach proposed by Sarkisov *et al.*³³, where the MOF is represented as a space frame truss, mapping the atomistic system on a mechanical system, can indeed be considered as a coarse graining. The limitation of this method is that it can distinguish only in a qualitative way between mechanically flexible (e.g. “breathing”) systems and rigid ones. In contrast to that, a properly parameterized coarse grained force field should at least semi-quantitatively reproduce the essential features of structure and mechanical flexibility of the MOF, like for example elastic constants.

In this contribution we propose a first and maximally coarse grained force field for copper paddle-wheel (PW) based MOFs, focusing on the archetypical Cu_3btc_2 ($btc = \text{benzene-1,3,5-tricarboxylate}$) or HKUST-1.³⁴ The square planar geometry of the

paddle-wheel is particularly challenging for a coarse grained representation as compared to e.g. an octahedral Zn_4O fragment in the IRMOF series.³⁵ For the fitting of the bonded parameters of the coarse grained force field we propose a strategy similar to the one used previously for the optimization of the atomistic MOF-FF, however, taking MOF-FF itself as a reference instead of QM computed data. Different levels of coarse graining are discussed and the influence of the non-bonded parameters on the results is investigated. The aim is to proof the concept and to assess which properties can be reproduced at this extreme level of coarse graining. This is only a first step towards a general coarse grained force field for MOFs, since neither the optimal level of coarse graining and potential energy expression, nor the best parameterization procedure are clear from the outset.

2 Computational Details

All force field calculations for both the atomistic and the coarse grained systems have been performed with our in house developed *pydlpoly* molecular mechanics code.²² *Pydlpoly* is based on the DL-POLY Classic package³⁶ written in the Fortran 90 programming language, but is completely restructured and interfaced by the object oriented scripting language Python^{37,38} using the f2py code generator.³⁹

The fit of the CG force field parameters was performed in analogy to the fit of the atomistic MOF-FF as described in ref. 20 by a Genetic Algorithm global optimization scheme using the *Pikaia*⁴⁰ code. The single objective function is based on accordingly weighted mean square deviations of structure and curvature information, represented in redundant internal coordinates (RIC)^{41–43} defined by the connectivity of the force field. We have used a complete rewrite of the force field generator code *ff-gen*, which is now written in Python in a strictly object oriented fashion. It combines the *pydlpoly* MM back-end with a RIC-module to compute the objective function and the *Python* wrapped *Pikaia* GA-module. Typically, a population of 200 species is propagated over 200 generations using a steady state reproduction plan with a selection probability based on fitness based ranks, a two point crossover operation with a probability of 85% and a uniform one-point (jump) and creep mutation with equal probabilities of 5%. In addition the elitism option is used in the GA optimization.

We employed the atomistic MOF-FF²² force field as a reference. Note that it is in principle possible, but numerically more demanding, to fit the CG force field directly to QM reference data. Further details on the generation of the reference data and the parameter fit are described in the next section.

All parameters for the CG force field are listed in the Electronic Supporting Information (ESI).

Periodic calculations of full MOFs with the atomistic MOF-FF have been performed using a cut-off of 12 Å for all non-bonded interactions and the Gezelter et al.⁴⁴ shifted force approach for truncating the Coulomb interactions. For systems with unit cell sizes comparable to the cut-off, a $2 \times 2 \times 2$ supercell was used.

3 Results and Discussion

3.1 Levels of Coarse Graining

In our proof-of-concept investigation we have used the well studied HKUST-1³⁴ copper paddle-wheel based MOF to illustrate the potential of coarse grained FFs. HKUST-1 exhibits the edge transitive 3,4-connected **tbo** network topology, with the paddle-wheel forming a four coordinated square planar vertex and the organic linker being three coordinated (see Fig. 1). In principle, also a **pto** topology would be possible, but it is energetically less favorable.⁴⁵

From a conceptional point of view, there are numerous ways to coarse grain a MOF like HKUST-1. Depending on the number of beads, used to represent the system, different levels of coarse graining can be distinguished, which is schematically shown in Fig. 1. The minimal model is defined by the **tbo** topology itself, i.e. each vertex is a bead. This also corresponds to the “blueprint” used in the reversed topological approach (RTA)^{46,47} as a starting point to generate the atomistic model. For HKUST-1 there are 56 beads in the unit cell, which means a reduction of the number of interaction sites (beads vs. atoms) by more than one order of magnitude. In Fig. 1 two further models are shown, consisting of 152 and 248 beads. The larger we named “building block” model, because it is built from building blocks, having itself the symmetry of the corresponding atomistic fragment. Consequently, the linker and the paddle-wheel secondary building unit (SBU) comprise of five and four beads, respectively. In the smaller intermediate representation, the linker consists only of one bead at the vertex site, whereas the SBU is mapped on a square planar “building block” fragment of five beads. Note that in this case the question arises, whether the connecting bead is positioned at the carboxylate-carbon atom or combines with the C_α of the phenyl ring. CG force fields could in principle be developed for all three models by the same parameterization bottom-up scheme, which is discussed in the following. It is obvious, that the less coarse-grained models with more beads will comprise of more force field parameters to be fitted, and it will be easier to capture the properties of the full atomistic model.

Thus, in order to test the CG concept for MOFs in this first proof of concept step, we decided to use the most coarse grained, minimal model as a starting point.

To reproduce the flexibility of the periodic system in the correct manner it is essential to describe the structural deformations of both vertices as accurate as possible. Because of the weaker coordination bonds in the paddle-wheel it can be expected that it has the softer deformation modes and is thus most important. From a topological point of view, the paddle-wheel unit is a square planar vertex with a D_{4h} symmetry. With its five sites it has nine internal degrees of freedom (DOFs). A normal mode analysis of this reduced system reveals – ignoring the four stretching modes – that in principle four different deformation modes shown in Fig. 2 are possible. One of the in-plane bending modes (Fig. 2a) is responsible for the “wine-rack” type breathing transformation of paddle-wheel based layer-pillar MOFs⁴⁸ and we refer to it in the following as the wine-rack mode. Note that the other in-plane bending mode (Fig. 2b) is two-fold degenerate. The two out-of-plane bending modes lead to a pyramidal (Fig. 2c) and a tetrahedral (Fig. 2d) distortion. These four deformation modes of the underlying square planar fragment correspond to low lying normal modes of the actual atomistic system. In particular, the lowest normal modes with 17.1 cm^{-1} and 18 cm^{-1} , respectively, of the non-periodic benzoate model correspond very well to the wine-rack and in-plane bending mode. It is clear that any CG model needs to be able to reproduce these structural deformations of the building blocks sufficiently accurate.

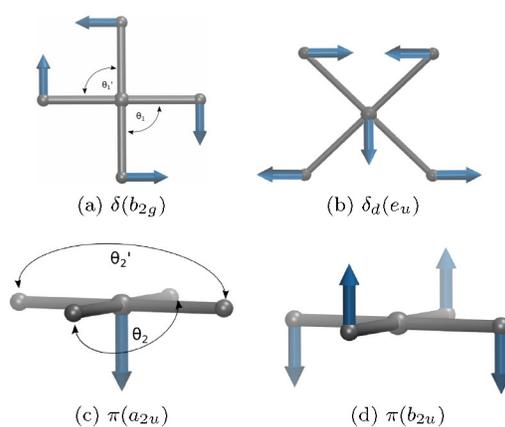


Fig. 2: Deformation modes of a square planar fragment representing the copper paddle-wheel unit. (a) “wine-rack” mode; (b) in-plane bending (two-fold deg.); (c) pyramidal distortion; (d) tetrahedral distortion

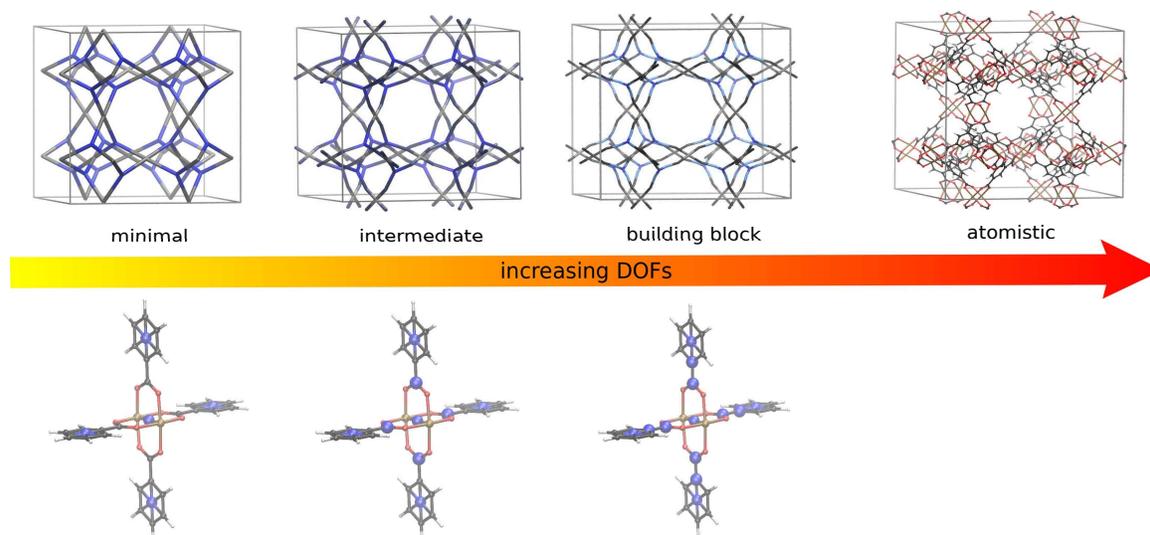


Fig. 1: Schematic representation of possible different levels of coarse graining of HKUST-1 (**tbo** topology). Blue spheres in the lower row indicate beads for the corresponding CG models.

3.2 Fitting of CG Parameters

In analogy to the atomistic MOF-FF, its CG variant MOF-FF-CG is parameterized by a bottom-up approach, which means the parameters are not fitted to reproduce global properties of the bulk system, as it is done for example in the BTW-FF,⁴⁹ but from local properties of the unperturbed building blocks, computed on a higher level of theory. This approach has the advantage that different supramolecular isomers, like for example different topologies, or even different combinations of building blocks can be investigated with a single FF. The non-periodic reference structures used for the parameterization of the minimal MOF-FF-CG for HKUST-1 are shown in Fig. 3. The benzoate model (Fig. 3a) serves as a reference system for the parameterization of the square planar SBU. Note that the same system was utilized also in Ref. 22 for the atomistic MOF-FF. In this case, also the necessary parameters for the organic linker could be derived from this model, whereas for MOF-FF-CG a further reference system to represent the three coordinated vertex is necessary, which is shown in Fig. 3b. The blue spheres in Fig. 3 represent the bead positions in the CG-FF, and they are added as virtual sites, defined by the center of mass (COM) of its defining atoms, in the atomistic model. The virtual sites are defined either by the six phenyl carbon atoms or by the two copper atoms, which is indicated in Fig. 3 by additional bonds between the defining atoms and the virtual sites.

In the very first step, a MOF-FF-CG was parameterized without any explicit inclusion of non-bonded van der Waals (vdW) and Coulomb interactions, in other words they were only implicitly

included in the bonded interactions. As for MOF-FF, the bonded parameters were obtained by a GA optimization of an objective function, constructed from weighted mean square deviations of geometric data and Hessian matrix elements, represented in redundant internal coordinates defined by the connectivity of the CG system. After a free optimization of the atomistic model the geometric data like bond length and angles can be extracted from the COM positions of the virtual sites. The computation of the CG Hessian (H_{CG}) from the atomistic model, however, is not straight forward. To compute it, Lagrange constraints for the virtual site positions have been implemented. The CG Hessian with the dimension $[3N_b, 3N_b]$ (where N_b is the number of beads) was calculated by a double sided finite difference approach. Starting from the optimized geometry, all virtual sites are in turn displaced by a small distance d in one of the Cartesian directions (in both positive and negative direction). In order to maintain the COM constraint also the corresponding defining atoms are moved by the same displacement d . Then a geometry optimization with the L-BFGS algorithm⁵⁰ is performed, with all virtual sites constrained to their positions. For this purpose, the force acting on each virtual site was projected out by subtracting it from the force on its defining atoms. The convergence criterion for the L-BFGS optimizer has to be chosen very tight in this case, in order to get converged remaining forces on the virtual sites. A threshold of $5 \cdot 10^{-7}$ kcal/mol/Å was used together with a displacement of $d = 0.001$ Å. In total, $6N_b$ such constrained optimizations have to be performed and the Matrix H_{CG} is constructed in the usual way from the change in the remaining gradient on the

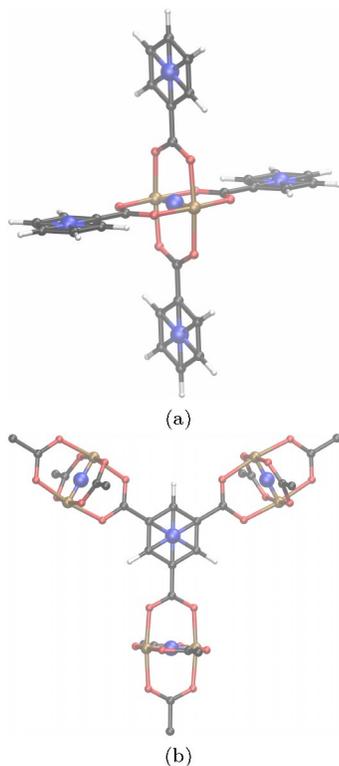


Fig. 3: Models used for the parameterization of the CG FFs (color scheme: black: carbon; white: hydrogen; red: oxygen; brown: copper; blue: virtual site, COM of its defining atoms).

virtual sites. In the following GA optimization of the MOF-FF-CG parameters, only the diagonal elements of H_{CG} (represented in internal coordinates) were considered in the objective function.

Originally we intended to use the same potential energy expressions, used in the atomistic MOF-FF, also for MOF-FF-CG. However, initial tests revealed an important difficulty for the square planar D_{4h} symmetric SBU site, which is of particular relevance for the minimal CG model. For the in-plane bending a Fourier term

$$E_{a,coord}^{bend} = \frac{V_a}{2} [1 + \cos(4\theta_a + \theta_a^{ref})], \quad (1)$$

is used, with θ_a the valence angle, θ_a^{ref} the phase shift, V_a the energy barrier. Especially because of the absence of non-bonding interactions, the difference between in-plane and out-of-plane bending can not be captured properly, since out-of-plane bending terms in force fields are only considered for trigonal centers. Note that, in case of the atomistic MOF-FF, the explicit use of 1,3-vdW interactions for such square planar coordination environments helps to reduce this problem. To resolve the problem within MOF-FF-CG we introduced a special out-of-plane bending term for four coordinated

square planar systems, which is explained in detail in the ESI. This term can also be used to improve the description of a square planar coordination geometry in an atomistic force field, and demonstrates that the potential energy expression in a CG force fields needs to be chosen with care. For the pseudo octahedral $Zn_4O(O_2C)_6$ moiety, a Fourier angle bending term (Eq. 1) would be sufficient, whereas a more complex bending term with multiple minima would be necessary to coarse grain the pseudo-cuboctahedral $ZrO_4(OH)_4(O_2C)_{12}$ fragment. For bond stretching and angle bending the usual MM3 potentials with fixed anharmonicity contributions were used.⁵¹ The parameters for the torsional potential, describing the dihedral angle between two SBUs, were adopted from the original MOF-FF, and ill defined torsions were neglected. No cross-terms were included at this proof-of-concept stage.

The GA optimized parameters for the MOF-FF-CG are summarized in the ESI. It is worth to note that the optimized stretch force constant is with 1.19 mdyn/Å much lower than a typical covalent bond force constant in the atomistic force field. Even the Cu-O stretch force constant in MOF-FF is with 1.46 mdyn/Å more stiff. This is due to the fact that shortening the SBU to linker-bead distance can be achieved via angle bending, e.g. within the carboxylate group.

3.3 Validation

As mentioned before, the important low normal modes of the paddle-wheel unit correspond to the basic deformation modes of a square planar unit (see Fig. 2) and have to be reproduced by the CG force field. Thus, the quality of MOF-FF-CG for the minimal model was validated first by potential energy scans along the wine-rack and pyramidal deformations for the SBU model (Fig. 3a). In case of the atomistic system, the scans were performed by constraining the virtual sites to the angles θ_1 or θ_2 and a similar constraint was employed for the beads in the CG-model. The results are shown in Fig. 4. The potential curves for the other in-plane bending and the tetrahedral deformation look nearly identical and are not shown. Note that these deformations have not been part of the training set of the CG FF. Because of the Hessian fit the curves calculated with the CG model match the atomistic ones especially well in the minimum region. However, also for rather large deformations, where various stretch and bend DOFs of the atomistic model are involved, the approximate CG model captures the potential energy very well. Larger deviations are seen for the wine-rack mode, which is due to the wrong behavior of the Fourier bending term (Eq. 1) for valence angles approaching zero in the absence of 1,3-vdW interactions. In contrast, the pyrami-

dal distortion shows perfect agreement up to large deformations.

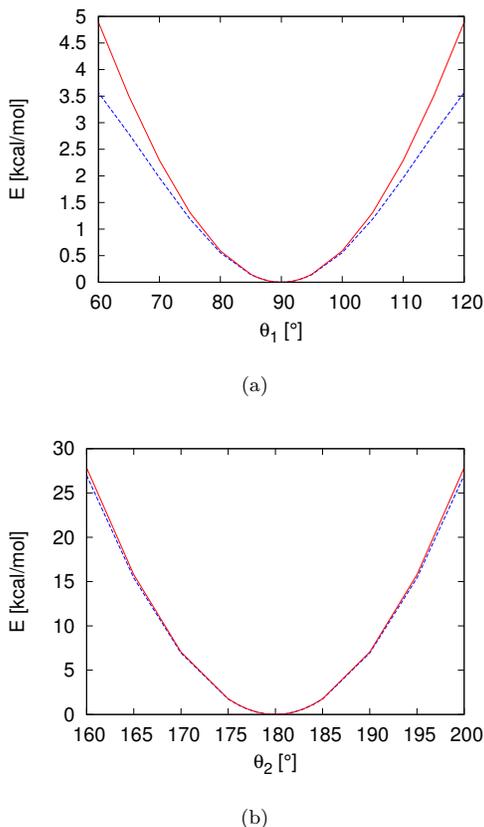


Fig. 4: Potential energy scans along the wine-rack (a) and pyramidal distortion (b) obtained with MOF-FF-CG (dashed blue line) and MOF-FF (solid red line).

As a further test, full 3D periodic network structures were considered. As already mentioned, two different edge-transitive topologies are possible for 3,4-connected systems with a square planar vertex like HKUST-1: the **tbo** and the **pto** topology. HKUST-1 crystallizes in the **tbo** topology, whereas MOF-14⁵², which differs from HKUST-1 by an extended linker, prefers the **pto** topology.⁴⁵ In order to assess the quality of MOF-FF-CG, the energy difference between these supra-molecular isomers was calculated with MOF-FF and MOF-FF-CG. Furthermore, lattice constants and elastic properties for both topologies have been evaluated with the atomistic and the CG model. The results are shown in Table 1 and Table 2. The Young modulus of the cubic systems E_x was calculated via a one dimensional scaling operation with the scaling factor s_x . The size of the simulation box in x direction was changed from $L_{0,x}$ to $s_x L_{0,x}$, and the atomic positions were relaxed. The Young modulus is then given by:

$$E_x = \frac{1}{L_0^3} \frac{\partial^2 E(r')}{\partial s_x^2}. \quad (2)$$

The scale factor s_x was varied from 0.98 to 1.02 in steps of 0.005. The second derivative in Eq. 2 was determined from a harmonic fit of the curve. The Bulk modulus is related to the curvature of $E(V)$,

$$B = V_0 \frac{d^2 E}{dV^2} \quad (3)$$

where V is the volume of the unit cell and E is the energy/unit cell at volume V . The second derivative was derived from a Birch-Murnaghan equation of state fit to the energy vs. volume curve.

In Table 1 the energy differences between the **tbo** and the **pto** topology are summarized, together with the deviations between the atomistic and CG force field. All energies are given with respect to the formula unit S_3T_4 (where S represents the square planar and T the trigonal building block). The **tbo** unit cell consists of eight such units, whereas the **pto** net is formed by two. The CG model correctly predicts the **tbo** system to be more stable, but it overestimates the energy difference by a factor of two. In Table 1 also the individual energy contributions to the energy are given. They are only of limited significance, since in the CG model certain terms are missing and are implicitly included in others. For example, the Coulomb interaction is entirely missing, but contributes about 4 kcal mol^{-1} to the total energy difference in the atomistic calculations. As expected, the main reason for the preference of **tbo** is the torsional contribution in both MOF-FF and MOF-FF-CG.⁴⁵ However, a close inspection reveals that the main source of error is the lack of vdW interactions in the CG model. These non-bonded terms stabilize the slightly denser **pto** system in the atomistic calculation, reducing the overall energy difference.

A comparison of the lattice constants (see Table 2) shows a very good agreement between the two models, which is expected due to the structural fit of the building blocks. The elastic properties, on the other hand, reveal again certain shortcomings of the CG approximation as summarized in Table 2. Both Bulk and Young moduli are underestimated for MOF-FF-CG. For the less dense **tbo** the moduli are consistently lower for both models, but differ more strongly between the atomistic and the CG model, as compared to **pto**.

3.4 Inclusion of non-bonded interactions

Overall, the CG model is able to reproduce the PES of deformations (see Fig. 4) as well as properties of the actual MOF fairly well. The largest deficiency, however, is the lack of vdW interactions between

the beads. Consequently, we developed a revised model MOF-FF-CGNB, which includes these interactions in an approximate way. Following current strategies in CG models for bio-polymers⁵³ we avoided the complexity arising from non-spherical bead models and tried to mimic non-bonded interactions (both Coulomb and vdW) from the atomistic model into a spherical Buckingham type potential, similar to the one used in MOF-FF. First, non-bonded energy curves for the approach of rigid building blocks (benzene and a simple paddle-wheel system truncated at the carboxylate carbon atoms) were determined using MOF-FF. Only a single, most relevant relative orientation was considered at this point. Further details and the resulting curves are given in the ESI. Each pairwise interaction was fitted individually by a least squares method, avoiding the use of combination rules. Interestingly, we found it difficult to represent these curves by a regular Buckingham potential, since the dispersive attraction is due to the summed effect of all atoms mapped to the bead, whereas the repulsive part is mainly determined by the outermost atom. A much improved fit could be achieved by using a somewhat modified Buckingham potential of the form

$$E_{ij}^{vdW+C} = Ae^{-B(d_{ij}-d_{ij}^0)} - \frac{C}{(d_{ij}-d_{ij}^0)^6}, \quad (4)$$

using an offset parameter d_{ij}^0 . The parameter A and B determine the strength and steepness of the repulsive part, whereas C adjusts the strength of the dispersive interaction. The inclusion of these vdW interactions in MOF-FF-CGNB did not require any refit of the bonded parameters, because 1-2 and 1-3 non-bonded interactions were excluded, similar to MOF-FF. The non-bonded interactions, present within the building blocks, are already implicitly contained in the bonded terms in MOF-FF-CG. In Table 1 and Table 2 the results for relative energies of **tbo** vs. **pto** and the elastic properties of both phases are shown for comparison also with MOF-FF-CGNB. In particular, the overall energy difference between **tbo** and **pto** is substantially improved by adding approximate non-bonded interactions in this way.

The lattice constants show only small changes, but especially the Young modulus of the **pto** phase is now in an almost perfect agreement, whereas for **tbo** only a small improvement was achieved. As mentioned before, the **tbo** phase is less dense i. e., the distances between the beads are larger than in the **pto** case. In addition, the configurations used for the potential energy scans to fit the CG vdW parameters resemble more closely to the configurations present in the **pto** phase. Overall, the addition of vdW interactions improved the quality of the CG force field markedly even for the minimal model, where only a few spherical beads are used.

Note that the largest deviations for the Young and Bulk moduli for **tbo** are in the order of 10%, which is well in the range of discrepancies observed for different theoretical methods (ranging from periodic DFT to different force fields) as well as for measured values. Thus, even the minimal model is able to achieve our aim of semi-quantitative accuracy.

3.5 Limitations

In order to assess the applicability of MOF-FF-CGNB for dynamic properties we have computed the thermal expansion coefficient of desolvated HKUST-1. With an atomistic FF (predecessor of MOF-FF) the negative thermal expansion (NTE) behavior could be predicted quantitatively.²¹ In Fig. 5 the temperature dependence of the lattice constant for MOF-FF-CGNB is shown. It was obtained from a series of 4 ns sampling runs in the NPT ensemble between 50 and 500 K at a pressure of 1 atm, using a $3 \times 3 \times 3$ supercell. In contrast to the atomistic FF, no NTE is observed for our maximally coarse grained model. Wu *et al.*⁵⁴ have proposed three different vibrational motions to be responsible for the NTE in HKUST-1: the translation and libration motion of the benzyl ring of the linker and the local twisting of the PW unit. Clearly, in the minimal model these motions are not present because both the PW unit and the linker consist of only one bead. Thus, in order to reproduce the effect of NTE with a CG FF a model has to be used, which is able to resolve these motions, like for example the above proposed less coarse grained building block model (Fig. 1).

This example illustrates a general limitation of coarse grained FFs. The ability to simulate systems on larger time and length scales comes with the price of a loss of resolution. The minimal model of HKUST-1 is able to reproduce the mechanical properties but not the NTE. Note that in case of MOF-5 the mechanical stability comes largely from the orthogonal orientation of the adjacent planar linkers, leading to a large and small pore.⁵⁵ Thus, it can be expected that a maximally coarse grained MOF-5 force field might not even be able to reproduce the mechanical properties.

From this follows the general rule that the coarse grained model has to be chosen carefully with respect to the required property and the system of interest.

3.6 Application: When Size Matters

Certain properties of crystalline materials cannot be computed accurately under the approximation of periodic boundary conditions using small unitcells. An example are mesopores in a MOF resulting from the introduction of defect linkers.²⁶ In order to investigate the overall mechanical stability depend-

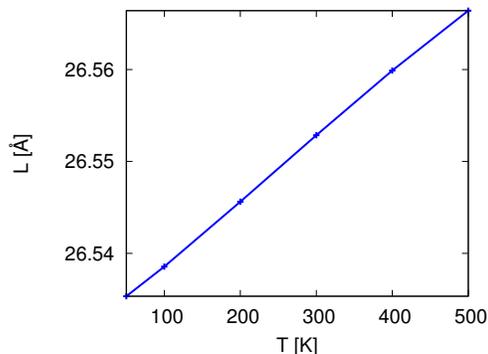


Fig. 5: Thermal expansion of HKUST-1 computed with MOF-FF-CGNB (cell parameter vs. temperature).

ing on the type and size of mesopores, it will be necessary to embed them into a sufficiently large non-defective material. Similar large system sizes would be needed to simulate correlated disorder as observed recently in UiO-66 type materials.⁵⁶ Such applications will be the target in our further investigations using CG force fields for MOFs.

Because of the success in describing the elastic properties of the bulk material, we illustrate the application of MOF-FF-CGNB here by computing the effect of a nanosized “tip” deforming the surface of an HKUST-1 crystal. This scenario is inspired by nano-indentation experiments,⁵⁷ where the applied load is recorded with respect to the depth of penetration of an indenter tip into the material in order to extract elastic properties. Such experiments have been performed for a thin film of HKUST-1 experimentally.⁵⁸ Note, however, that in the experiments “dead material” is formed by plastic deformation upon lowering the indenter, resulting in a hysteretic behavior. This can of course not be reproduced by our non-reactive force field and the scenario is chosen just to demonstrate the potential use of a CG FF where size effects led to a slow convergence with system size.

A slab model exposing the most stable [111] surface⁵⁹ with a thickness of six unitcells along the z -axis was constructed, using supercell sizes of 1×1 up to 5×5 in the lateral (x and y) direction. The surface is assumed to be terminated by acetate groups, which are represented in MOF-FF-CGNB by spherical beads for the CH_3 -groups.⁵⁹ In the smallest case ($1 \times 1 \times 6$) the corresponding system has dimensions of about $3.8 \times 3.3 \times 25.3 \text{ nm}^3$, whereas the largest slab ($5 \times 5 \times 6$) is about $18.8 \times 16.3 \times 25.3 \text{ nm}^3$ in size and consists of 23,500 beads (see Fig. 6). For simplicity the indenter was modeled by an atomistic model using entirely repulsive tungsten atoms cut from a primitive cubic packing (cell constant of 3.16 Å. The first atom of the tip was positioned

5.82 Å above the surface of the slab. It was moved 20 Å into the material performing 40 steps with a stepsize of 0.5 Å. In every step the positions of the framework beads were completely relaxed with the lowest layer and the tip atoms frozen. In Fig. 7 the total energy with respect to the penetration depth is shown for all systems.

In the case of the smallest $1 \times 1 \times 6$ supercell, the indenter tip is nearly as large as the exposed surface patch and thus it is not penetrating into the material, but compressing the whole slab nearly uniaxial along the z -coordinate (see Fig. S3). A harmonic fit of the energy vs. penetration depth curve, shown in Fig. 7, results in a force constant of 9.70 kcal/mol/Å² (see also Table S4). In contrast to that, as shown in Fig. 6 on the right, a shear type deformation in the area around the tip is observed, due to the point-like indenter tip. Because of the size of the system, this local deformation is not coupling to its periodic images in the 2D slab model. This is reflected by the energy vs. penetration depth curves in Fig. 7 converging with system size. Interestingly, the corresponding force constants converge with 18.5 kcal/mol/Å² to a value that is about twice as high as compared to the nearly uniaxial compression in case of the smallest $1 \times 1 \times 6$ system (Table S4).

The largest $5 \times 5 \times 6$ system with its 23,500 beads represents about 258,000 atoms of the corresponding atomistic model. The computation of the complete indentation curve took about 10 min with a single energy and force calculation in about 0.2 s on a single core of a current desktop GNU/LINUX workstation. Note that in a molecular dynamics simulation additional savings in the cpu-time can be achieved due to the fact that a larger time step can be employed, since the high frequency motions like C-H bonds have been removed. Interestingly, Banlusan *et al.*⁶⁰ have recently studied the plastic deformation of MOF-5 by molecular dynamics simulations with the reactive force field ReaxFF⁶¹ and an atomistic slab model of 183,000 atoms. In this study a massively parallel computing cluster and extensive computation time were used, which are not available for routine calculations. The example shows, that a CG FF can be used to gain insight at a substantially larger length and time scale at moderate or even small computational costs, under the condition that the CG model is appropriate for the given problem.

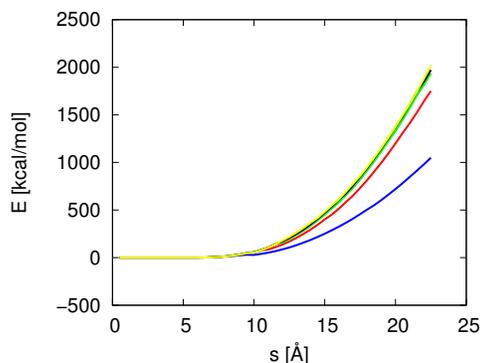


Fig. 7: Energy wrt the penetration depth s of the “tip” for the $1 \times 1 \times 6$ (blue line), $2 \times 2 \times 6$ (red line), $3 \times 3 \times 6$ (green line), $4 \times 4 \times 6$ (black line) and $5 \times 5 \times 6$ (yellow line) supercell slab models.

4 Conclusion and Outlook

We have presented a first parameterization of a coarse grained force field for a MOF on the example of the copper paddle-wheel based HKUST-1. As a proof of concept, the minimal representation (using one bead per vertex) was developed, adapting the “bottom-up” systematic parameterization approach, previously employed to derive parameters for the atomistic force field from QM reference data. It was validated by testing the accuracy of the CG force field to reproduce the fundamental low energy deformation modes of the isolated square planar SBU. In addition, relative stabilities, lattice parameters and elastic properties of both topologies **tbo** and **pto** were computed. Especially after adding approximate steric interactions to the CG model by fitting a modified Buckingham potential for bead-bead interactions to the non-bonded interaction energies of the atomistic building blocks, a semi-quantitative prediction of the structural and mechanical properties could be achieved. All tendencies computed by the atomistic MOF-FF were reproduced even for the maximally coarse grained force field MOF-FF-CGNB. However, as expected, the NTE effect in HKUST-1, which can be quantitatively be predicted by an atomistic FF, is absent in the maximally coarse grained model, since the relevant deformation modes can not be resolved. Because of the substantial reduction in the number of interaction sites, large savings in the cpu time can be achieved, making it possible to compute systems with tens of nanometers in size. As an example, the local deformation by a point like “tip”, lowered into a [111] surface of HKUST-1, could be investigated by a large $5 \times 5 \times 6$ supercell slab model.

It is very encouraging to see that these promising results could be achieved even for a maximum coarse graining. Depending on the system and

property of interest, theoretical simulations of sufficient accuracy were possible at a substantially larger length and timescale. The more approximate CG approach is naturally limited, where effects are due to atomistic details, which are lost in the coarse graining. Besides the NTE of HKUST-1, it can be expected that the current maximally coarse grained MOF-FF-CGNB with its few beads is not able to capture host-guest interactions (and thus adsorption isotherms) properly. A further lesson to be learned is the fact that the potential energy terms established for atomistic FFs can not always be employed as such. Especially the coordination environments in the inorganic fragments can be challenging here. On the other hand, we could show that the GA optimization method, originally developed to parameterize atomistic force fields, can successfully be employed also for the CG models.

It is evident that our proposed MOF-FF-CGNB is only a first step towards a general CG FF for MOFs. Our next step will be to use a slightly less coarse grained model with more beads, at the expense of a somewhat higher computational cost. We are aiming for a coarse graining at the building block level for the inorganic part and will borrow concepts developed in CG FFs for (bio-)polymers, like for example the “Martini” force fields, where a 4:1 ratio of atoms to beads is found to be optimal.⁵³ This will not only improve the mechanical properties, but also allow a more accurate calibration of CG non-bonded interactions, since less atoms are mapped on a single spherical bead. In addition, a more elaborate strategy to fit these non-bonded interactions, including multiple relative orientations of the fragments will be developed. This should allow at least a semi-quantitative representation of host-guest interactions, enabling for example to study the coupling between mechanical deformations and guest adsorption for nanosized crystallites without using periodic boundary conditions.

Overall we could show that sufficiently accurate coarse grained force fields can be achieved, allowing to extend both length and time scale in the simulation of MOFs. They can be of use either as such or in context of concurrent multiscale simulations in a MM-CG or even QM-MM-CG fashion and can facilitate to bridge from atomistic to mesoscale simulations.

Acknowledgment

This project has financially been supported by the Deutsche Forschungsgemeinschaft (DFG, Priority Program 1362). Further financial support from the Cluster of Excellence RESOLV (EXC 1069) funded by the DFG is acknowledged.

References

- [1] R. Kitaura, K. Seki, G. Akiyama and S. Kitagawa, *Angew. Chem. Int. Ed.*, 2003, **42**, 428.
- [2] S. L. James, *Chem. Soc. Rev.*, 2003, **32**, 276.
- [3] J. L. C. Rowsell and O. M. Yaghi, *Micropor. Mesopor. Mater.*, 2004, **73**, 3.
- [4] S. Kitagawa, R. Kitaura and S. Noro, *Angew. Chem., Int. Ed.*, 2004, **43**, 2334.
- [5] G. Ferey, *Chem. Soc. Rev.*, 2008, **37**, 191.
- [6] H.-C. Zhou, J. R. Long and O. M. Yaghi, *Chem. Rev.*, 2012, **112**, 673.
- [7] M. Li, D. Li, M. O. Keffe and O. M. Yaghi, *Chem. Rev.*, 2014, **114**, 1343.
- [8] G. Ferey and C. Serre, *Chem. Soc. Rev.*, 2009, **38**, 1380.
- [9] S. Keskin, J. Liu, R. B. Rankin, J. K. Johnson and D. S. Sholl, *Ind. Eng. Chem. Res.*, 2009, **48**, 2355.
- [10] T. Duren, Y.-S. Bae and R. Q. Snurr, *Chem. Soc. Rev.*, 2009, **38**, 1237.
- [11] M. Tafipolsky, S. Amirjalayer and R. Schmid, *Micropor. Mesopor. Mater.*, 2010, **129**, 304.
- [12] W. Koch and M. C. Holthausen, *A Chemist's Guide to Density Functional Theory*, Wiley-VCH, 2nd edn, 2001.
- [13] D. Dubbeldam, A. Torres-Knoop and K. S. Walton, *Mol. Simul.*, 2013, **39**, 1253.
- [14] M. Tafipolsky, S. Amirjalayer and R. Schmid, *J. Comput. Chem.*, 2007, **28**, 1169.
- [15] L. Vanduyfhuys, T. Verstraelen, M. Vandichel, M. Waroquier and V. Van Speybroeck, *J. Chem. Theory Comput.*, 2012, **8**, 3217.
- [16] J. S. Grosch and F. Paesani, *J. Am. Chem. Soc.*, 2012, **134**, 4207.
- [17] F. Salles, A. Ghoufi, G. Maurin, R. G. Bell, C. Mellot-Draznieks and G. Ferey, *Angew. Chem. Int. Ed. Engl.*, 2008, **47**, 8487.
- [18] A. K. Rappe, C. J. Casewit, K. S. Colwell, W. A. Goddard and W. M. Skiff, *J. Am. Chem. Soc.*, 1992, **114**, 10024.
- [19] M. A. Addicoat, N. Vankova, I. F. Akter and T. Heine, *J. Chem. Theory Comput.*, 2014, **10**, 880.
- [20] M. Tafipolsky and R. Schmid, *J. Phys. Chem. B*, 2009, **113**, 1341.
- [21] M. Tafipolsky, S. Amirjalayer and R. Schmid, *J. Phys. Chem. C*, 2010, **114**, 14402.
- [22] S. Bureekaew, S. Amirjalayer, M. Tafipolsky, C. Spickermann, T. K. Roy and R. Schmid, *Phys. Stat. Sol. B*, 2013, **250**, 1128.
- [23] L. Vanduyfhuys, S. Vandenbrande, T. Verstraelen, R. Schmid, M. Waroquier and V. Van Speybroeck, *J. Comput. Chem.*, 2015, **36**, 1015.
- [24] C. Zhang, J. A. Gee, D. S. Sholl and R. P. Lively, *J. Phys. Chem. C*, 2014, **118**, 20727.
- [25] F.-X. Coudert, *Chem. Mater.*, 2015, **27**, 1905.
- [26] Z. Fang, J. P. Dürholt, M. Kauer, W. Zhang, C. Lochenie, B. Jee, B. Albada, N. Metzler-Nolte, A. Pöpl, B. Weber, M. Muhler, Y. Wang, R. Schmid and R. A. Fischer, *J. Am. Chem. Soc.*, 2014, **136**, 9627.
- [27] B. A. Merchant and J. D. Madura, *Annual Reports in Computational Chemistry*, Elsevier, 2011, vol. 7, p. 67.
- [28] G. Voth, *Coarse-Graining of Condensed Phase and Biomolecular Systems*, CRC Press/Taylor and Francis Group, Boca Raton, FL, 2009.
- [29] S. Izvekov, A. Violi and G. A. Voth, *J. Phys. Chem. B*, 2005, **109**, 17019.
- [30] H. I. Ingólfsson, C. A. Lopez, J. J. Uusitalo, D. H. de Jong, S. M. Gopal, X. Periole and S. J. Marrink, *WIREs Comput. Mol. Sci.*, 2014, **4**, 225.
- [31] W. L. Jorgensen and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1988, **110**, 1657.
- [32] M. G. Martin, and J. I. Siepmann, *J. Phys. Chem. B*, 1998, **102**, 2569.
- [33] L. Sarkisov, R. L. Martin, M. Haranczyk and B. Smit, *J. Am. Chem. Soc.*, 2014, **136**, 2228.
- [34] S. S. Chui, M. Samuel and P. Jonathan, *Science*, 1999, **283**, 1148.
- [35] M. Eddaoudi, J. Kim, N. Rosi, D. Vodak, J. Wachter, M. O'Keeffe and O. M. Yaghi, *Science*, 2002, **295**, 469.
- [36] W. Smith and T. Forester, *J. Mol. Graph.*, 1996, **14**, 136.
- [37] G. van Rossum and F.L. Drake (eds), *Python Reference Manual*, PythonLabs, Virginia, USA, 2001. Available at <http://www.python.org>.
- [38] D. Ascher, P.F. Dubois, K. Hinsin, J. Hugunin and T. Oliphant, *Numerical Python*, Lawrence Livermore National Laboratory, Livermore, California, USA, 2001. Available at <http://www.pfdubois.com/numpy/>.
- [39] P. Peterson, *Int. J. Comput. Sci. Eng.*, 2009, **4**, 296.
- [40] P. Charbonneau, *Astrophys. J. Suppl. S.*, 1995, **101**, 309.
- [41] P. Pulay and G. Fogarasi, *J. Chem. Phys.*, 1992, **96**, 2856.
- [42] C. Peng, P. Y. Ayala, H. B. Schlegel and M. J. Frisch, *J. Comput. Chem.*, 1996, **17**, 49.

- [43] V. Bakken and T. Helgaker, *J. Chem. Phys.*, 2002, **117**, 9160.
- [44] C. J. Fennell and J. D. Gezelter, *J. Chem. Phys.*, 2006, **124**, 234104.
- [45] S. Amirjalayer, M. Tafipolsky and R. Schmid, *J. Phys. Chem. C*, 2011, **115**, 15133.
- [46] S. Bureekaew and R. Schmid, *CrystEngComm*, 2013, **15**, 1551.
- [47] S. Bureekaew, V. Balwani, S. Amirjalayer and R. Schmid, *CrystEngComm*, 2015, **17**, 344.
- [48] A. U. Ortiz, A. Boutin, A. H. Fuchs and F.-X. Coudert, *J. Chem. Phys.*, 2013, **138**, 174703.
- [49] J. K. Bristow, D. Tiana and A. Walsh, *J. Chem. Theory Comput.*, 2014, **10**, 4644.
- [50] D. C. Liu and J. Nocedal, *Math. Program.*, 1989, **45**, 503.
- [51] N. L. Allinger, Y. H. Yuh and J. H. Lii, *J. Am. Chem. Soc.*, 1989, **111**, 8551.
- [52] B. Chen, M. Eddaoudi, S. T. Hyde, M. O’Keeffe and O. M. Yaghi, *Science*, 2001, **291**, 1021.
- [53] S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman and A. H. de Vries, *J. Phys. Chem. B*, 2007, **111**, 7812.
- [54] Y. Wu, A. Kobayashi, G. J. Halder, V. K. Peterson, K. W. Chapman, N. Lock, P. D. Southon and C. J. Kepert, *Angew. Chem. Int. Ed.*, 2008, **47**, 8929.
- [55] S. Amirjalayer and R. Schmid, *J. Phys. Chem. C*, 2008, **112**, 14980.
- [56] M. J. Cliffe, W. Wan, X. Zou, P. A. Chater, A. K. Kleppe, M. G. Tucker, H. Wilhelm, N. P. Funnell, F.-X. Coudert and A. L. Goodwin, *Nat. Commun.*, 2014, **5**, year.
- [57] A. Fischer-Cripps, *Surf. Coat. Tech.*, 2006, **200**, 4153.
- [58] S. Bundschuh, O. Kraft, H. K. Arslan, H. Gliemann, P. G. Weidler and C. Wöll, *Appl. Phys. Lett.*, 2012, **101**, 101910.
- [59] S. Amirjalayer, M. Tafipolsky and R. Schmid, *J. Phys. Chem. Lett.*, 2014, **5**, 3206.
- [60] K. Banlusan, E. Antillon and A. Strachan, *J. Phys. Chem. C*, 2015, **119**, 25845.
- [61] A. C. T. van Duin, S. Dasgupta, F. Lorant and W. A. Goddard, *J. Phys. Chem. A*, 2001, **105**, 9396.

Table 1: Energy differences $\Delta E = E(\mathbf{tbo}) - E(\mathbf{pto})$ per formula unit S_3T_4 for HKUST-1, calculated with MOF-FF, MOF-FF-CG and MOF-FF-CGNB. $\Delta\Delta E$ gives the difference between the atomistic and the CG methods. All energies are stated in kcal mol⁻¹.

	MOF-FF	MOF-FF-CG		MOF-FF-CGNB	
	ΔE_{AA}	ΔE_{CG}	$\Delta\Delta E$	ΔE_{CG}	$\Delta\Delta E$
total	-12.50	-23.68	11.17	-16.17	3.67
vdW	15.41	0.00	15.41	7.52	7.90
Coulomb	-4.24	0.00	-4.24	0	-4.24
stretch	0.98	0.00	0.98	-0.02	0.99
bend	-0.62	3.03	-3.64	3.17	-3.78
out-of-plane bend	0.90	0.52	0.38	0.46	0.44
torsion	-24.93	-27.22	2.90	-27.30	2.36

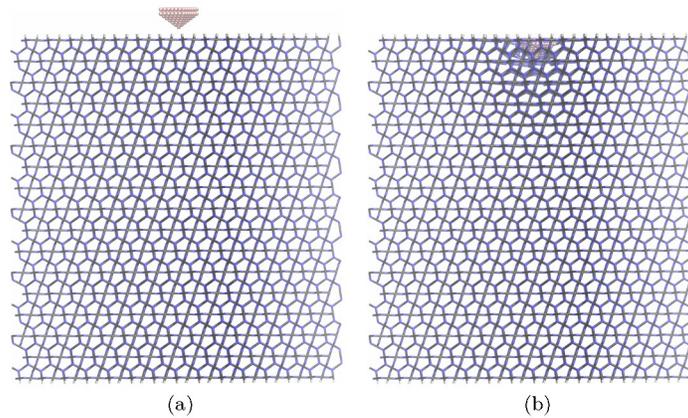


Fig. 6: Penetration of an atomistic “tip” into the [111] surface of a coarse grained slab model of HKUST-1 ($5 \times 5 \times 6$ supercell), where (a) shows the initial and (b) the final state (color scheme: black: SBU bead; blue: linker bead; white: CH_3 -bead; brown: Tungsten).

Table 2: Lattice constants and Young moduli for both the atomistic model and the minimal CG representation of HKUST-1 in the \mathbf{tbo} and \mathbf{pto} topology.

	\mathbf{tbo}			\mathbf{pto}		
	MOF-FF	MOF-FF-CG	MOF-FF-CGNB	MOF-FF	MOF-FF-CG	MOF-FF-CGNB
$L_x[\text{\AA}]$	26.43	26.56	26.56	30.61	30.71	30.68
$E_x[\text{GPa}]$	23.83	20.29	20.12	28.49	27.58	28.06
$B[\text{GPa}]$	19.59	17.64	17.75	26.86	25.91	26.72

