Phys. Status Solidi B **250**, No. 6, 1128–1141 (2013) / **DOI** 10.1002/pssb.201248460

Part of Special Issue on
Metal–Substrate Interactions in Heterogeneous Catalysis

physica status solidi b
www.pss-b.com
basic solid state physics

Feature Article

MOF-FF – A flexible first-principles derived force field for metal-organic frameworks

Sareeya Bureekaew[1], Saeed Amirjalayer[1,2], Maxim Tafipolsky[1,3], Christian Spickermann[1,4], Tapta Kanchan Roy[1,5], and Rochus Schmid[*,1]

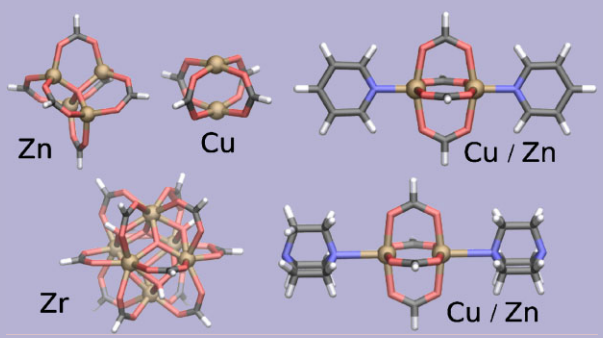
[1] Lehrstuhl für Anorganische Chemie 2, Computational Materials Chemistry Group, Ruhr-Universität Bochum, Universitätsstr. 150, 44780 Bochum, Germany
[2] van't Hoff Institute for Molecular Sciences, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands
[3] Institut für Physikalische und Theoretische Chemie, Universität Würzburg, Am Hubland, 97074 Würzburg, Germany
[4] R&D Electronics, Atotech Deutschland GmbH, Erasmusstraße 20, 10553 Berlin, Germany
[5] Institute of Chemistry, The Hebrew University of Jerusalem, 91904 Jerusalem, Israel

Received 28 September 2012, revised 21 January 2013, accepted 24 January 2013
Published online 1 March 2013

Keywords force-field, genetic algorithm, metal-organic frameworks, parametrization

[*] Corresponding author: e-mail rochus.schmid@rub.de, Phone: +49 234 3224166, Fax: +49 234 3214174

In this contribution the development, definition and selected applications of a new force field (FF) for metal-organic frameworks MOF-FF is presented. MOF-FF is fully flexible and is parameterized in a systematic and consistent fashion from first principles reference data. It can be used for a variety of different MOF-families and in particular – due to the reparametrization of a variety of organic linkers – also to explore isoreticular series of systems. The history of the development, leading to the final definition of MOF-FF is reviewed along with the application of the previous incarnations of the FF. In addition, the parametrization approach is explained in a tutorial fashion. The currently parametrized set of inorganic building blocks is constantly extended.



Formate models of currently covered inorganic building blocks.

1 Introduction The rather new class of porous materials, referred to as metal-organic frameworks (MOFs), has attracted a large interest in recent years, owing to their potential application in sensing, gas storage, chemical separations, and catalysis, to name a few key areas [1–8]. In particular, they are on the one side crystalline and ordered, but on the other hand flexible and able to respond to external stimuli. Thus, despite their lower thermal stability in comparison with zeolites – the long standing other class of porous and crystalline materials – MOFs will be of use in areas, where the variability in the choice of building blocks and the structural flexibility is of prime importance. However, in order to investigate their properties, which are sometimes not directly observable by experimental methods on an atomistic scale, theoretical support is needed [9–11]. For example, the mechanistic details of guest molecule adsorption in a flexible MOF system, leading to a hysteretic behavior ("gate opening") [12–14], is still not clear and hard to observe by spectroscopic or diffraction methods. Due to the partly large unit cells with many transition metal atoms, and the need to compute multiple configurations, accurate *ab initio* methods – like periodic density functional theory (DFT) calculations – are numerically often too demanding. A viable option is the approximate quantum mechanical

density functional tight binding (DFTB) approach, which was recently shown to give good results for MOFs without the need for reparametrization of existing parameter sets, and at a substantially lower numerical effort as compared to standard DFT [15]. However, a computationally even more efficient alternative is the use of a molecular mechanics (MM) type model, or force field (FF), which is aimed to replace the complicated electronic interactions by a set of analytic terms, usually considering two-, three-, and four-body interactions. In general, the atomic connectivity (covalent bonds) are used to separate the energy expression into bonded terms and pairwise non-bonded interactions. This formulation restricts the method to investigations of processes where no bonds are broken. For hybrid materials like MOFs, the metal-ligand bonds can either be described with bonded terms or as purely ionic interactions using non-bonded terms. Whereas a bonded FF allows to consider also orbital directing effects [16], the latter non-bonded approach has the advantage that the breaking and formation of coordination bonds can be simulated [17]. In both cases, many structural and dynamic properties of MOFs, like for example elastic constants, vibrational spectra or thermal expansion, can be investigated. As pointed out by us also relative stabilities of supramolecular isomers of MOFs can be treated in this context, allowing for a structure prediction [18, 19]. In addition, all host–guest interactions with only weak interactions (physisorption) can in principle be investigated this way.

Initially, a large number of theoretical FF investigations have used experimental geometries and kept them rigid. This is useful when computing adsorption isotherms by Grand Canonical Monte Carlo (GCMC) methods [10]. In this case, the term ''force field'' really refers only to the non-bonded interactions. However, if structure optimizations need to be performed, *e.g.*, in case of predicting structures of yet unknown systems, or when flexibility needs to be modeled, a fully flexible FF including a proper parametrization of bonded terms is needed. In a large number of MOF simulations, generic FFs such as the universal force field (UFF) [20] or the DREIDING FF [21] have been employed. The critical part of the parameter set is here the description of the inorganic nodes and in particular the binding of *e.g.* the carboxylate linker to this building block. The softness or stiffness of these parameters largely determines the quality of the overall FF. Unfortunately, well-parametrized FFs like OPLS – AA [22] contain only parameters for the purely organic part, which might be sufficient for the guest molecules, but not for the complete framework. A generic FF like UFF, on the other hand, has the advantage that it can be used essentially for any kind of MOF, since the parameters are derived ''on the fly'' by a rule based system from atomic parameters. This makes it a perfect tool for a qualitative structural modeling. However, it is clear that this approach can never deliver the same accuracy as a direct fit of the parameters to either experimental or high level theoretical reference data. Moreover, the method is a ''black box'' approach, and there is no clear route for assessing its accuracy in case of unknown and new systems or to improve its accuracy in a systematic way. Thus, different authors have proposed extensions of existing FFs for specific MOF systems. Greathouse and Allendorf combined CVFF [23] with a non-bonded parametrization of ZnO for the simulation of the Zn_4O-based MOF-5 [17, 24]. The same CVFF was combined with parameters from DREIDING for a non-bonded FF for the IRMOF series by Dubbeldam et al., who first discovered the strong negative thermal expansion (NTE) of these systems [25]. A number of authors have focused on the ''breathing'' MIL-53 [26] system and extended their FFs either by using first principles computed or experimental reference data [27–29]. Also for copper paddle-wheel based MOFs such an FF was proposed [30]. Despite the usefulness of these FFs for specific systems, they usually do not share a common energy expression. Also the parametrization strategy is very different, often mixing theoretical and experimental reference data. As a consequence, parameters of these FFs cannot be combined with each other, nor can the results be compared directly. Thus, a systematic strategy, using only theoretical higher level reference data in a consistent way for a wide variety of MOFs, is needed. This is why we set out to develop MOF-FF, which we present in this contribution. Note, that MOF-FF and its predecessors are bonded FFs, using a fixed connectivity also for the metal-ligand coordination bonds. We started in 2007 with a first generation FF for MOF-5 and related IRMOFs by extending the well-known MM3 [31, 32] with mostly manually fitted parameters for the inorganic part of the Zn_4O building block [33]. The second generation FF was then based on our new genetic algorithm (GA) parameter optimization approach, and was parametrized for Zn_4O and copper paddle-wheel based MOFs [34, 35]. The FF expression was still largely analogous to MM3 and, correspondingly, the parameters for the organic linker parts were taken from MM3. The current and third FF generation has now an energy expression that deviates from MM3 in important respects, and a complete GA-based reparametrization has been performed. Because of this it can no longer be considered as an extended MM3 FF, and we refer to it as an MOF-FF.

In the meantime, a few other groups have followed our general strategy and have developed FFs by a systematic approach from first principles. Grosch and Paesani developed an *ab initio* parametrized FF for the zinc paddle-wheel based $Zn_2(bdc)_2(dabco)$ (dabco = 1,4-diazabicyclo-[2.2.2]octan) layer-pillar MOFs, very much in the same way as MOF-FF [36]. The bonded FF terms describing the framework were determined by a GA-based fit to the corresponding *ab initio* data, calculated for non-periodic model systems of the inorganic building blocks. The remaining non-bonded terms were described by Lennard–Jones (LJ) potentials taken from GAFF [37]. The parametrization for the benzene guest molecule was taken from Refs. [38, 39], whereas the isopropyl alcohol FF was reparametrized. The simulation results could reproduce the structural deformations induced by benzene and isopropyl alcohol as a function of temperature and loading.

More recently, Vanduyfhuys et al. reported an accurate FF to describe a flexible MIL-53(Al) [40]. Two well-chosen cluster models, calculated on the DFT level including dispersion corrections, were used as references. The FF has three main contributions: an electrostatic term based on atomic charges derived with a modified Hirshfeld-I method, a van der Waals (vdW) term with parameters taken from the MM3 model and a valence FF, whose parameters were estimated with a new methodology that uses the gradients and Hessian matrix elements retrieved from the cluster calculations. The optimized FF was able to predict geometries and cell parameters for the breathing effect in good accord with the experimental values.

Note in passing, that the current paper focuses only on MOFs with inorganic building blocks. However, we have also used our approach to treat so called covalent organic frameworks (COFs). On the level of the second generation FF we have predicted the energetic preference of COF-102 [41] for the experimentally observed topology [19]. Later, the FF was extended to the full range of known 3D-connected boron containing COFs [42]. Recently, we have used the prescription of MOF-FF to derive an FF specifically for boroxin-based COFs and predicted the structures of yet unknown, hypothetical COFs [43].

The intention of this contribution is to give full reference to the current MOF-FF FF, but also to describe its development by a topical review of its predecessors. Thus, in the following section, both the first and second generation of FFs and their application will be described. In addition, their deficiencies, leading to the next generation, will be mentioned. In the third section, the energy expression of MOF-FF and the details about the technical implementation will be given. The last section is focused on the parametrization procedure, including the first principles reference calculations, the GA FF fitting and a survey of the current scope of MOF-FF, giving also some computed results for experimentally well-known systems.

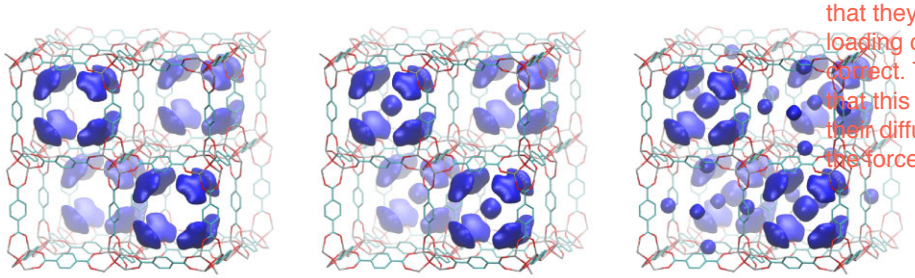
2 Predecessors of MOF-FF

2.1 The first generation force field: An extended MM3 for MOF-5 On the outset, we intended to simulate large guest molecules, like volatile metal-organic molecules, serving as metal precursors for a gas phase loading of MOFs with metal nanoparticles. These systems had just been presented by the group of Fischer and coworkers [44, 45]. It was evident, that such large molecules would lead to deformations of the MOF matrix and an appropriate FF would be needed here. Most empirical FFs used in common MM and molecular dynamics simulations are equipped with parameter sets for modeling biomolecules but, to our knowledge, none had been developed for MOFs. For the organic linker part this is not so much of a problem, because very reliable parameters exist, but for the inorganic building blocks, parameters are mostly not available. Due to our previous experience [46] with the MM3 FF [31], which is well known to accurately predict structures and conformational energies of organic molecules, we decided to extend it to cover the inorganic vertex of MOF-5 [47] ($[Zn_4O(BDC)_3]_n$, with BDC = 1,4-benzenedicarboxylate).

Since the three-dimensional periodic MOF-5 can be represented as being composed of a zinc-oxide cluster, $Zn_4O(O_2CR)_6$, the non-periodic benzoate model system with $R = C_6H_5$ was considered as the reference model. The model was fully optimized on the DFT level using the hybrid functional B3LYP [48]. Atomic point charges have been determined by the Merz–Kollman sampling Scheme [49] and were converted to bond dipoles to be used in MM3. The FF parameters missing in the MM3 set were mainly fitted manually. For details on the parametrization procedure we refer to Ref. [33]. Despite this simple procedure, lattice constants as well as infrared spectra of MOF-5 could be well reproduced by the FF.

This FF was extensively used in the following to investigate self-diffusion of benzene in MOF-5 [50, 51]. To our knowledge, we could for the first time, reproduce experimentally determined diffusion coefficients with our simulations [52]. By computing free energy maps as shown in Fig. 1, the adsorption and diffusion mechanism could be clarified on a molecular level. Interestingly, the topology of MOF-5 leads to a concentration of the guest molecules in the larger A-cells of MOF-5 and the diffusion occurs by a diagonal motion through one of the adjacent B-cells. Because of stabilizing guest–guest interactions, further sites of low free energy appear at a higher load, first, in the center of the A-cell and then also in the B-cell. In Fig. 2 the ΔG profile for such a motion is plotted along one of the crystallographic axes. This is the result of a total of about 50 ns of MD simulation. The free energy barriers fitted well to the self-diffusion at low loading, assuming a simple hopping mechanism.

Yeah, and then in 2011 they realized with their NMR study that they couldn't reproduce any diffusivities above a loading of 10 molecules per unit cell, so it wasn't really correct. They chose not to even cite that paper here. Not that this means the force field is wrong, it just means that their diffusion calculations calculated with it don't support the force field being correct.



a) 10 Bz/u.c. b) 32 Bz/u.c. c) 56 Bz/u.c.

Figure 1 (online color at: www.pss-b.com) 3D-Map of the free energy ΔG for different loadings of MOF-5 with benzene (Bz) guest molecules per unit cell (u.c.). The isosurface is drawn for $\Delta G = 5\,\mathrm{kJ\,mol^{-1}}$. (Reprinted from Ref. [51] with permission from Elsevier)

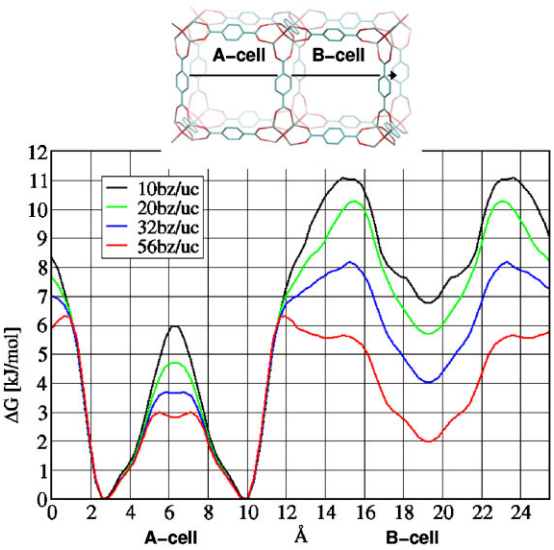


Figure 2 (online color at: www.pss-b.com) Free energy profile along one of the crystallographic axes. (Reprinted from Ref. [51] with permission from Elsevier.)

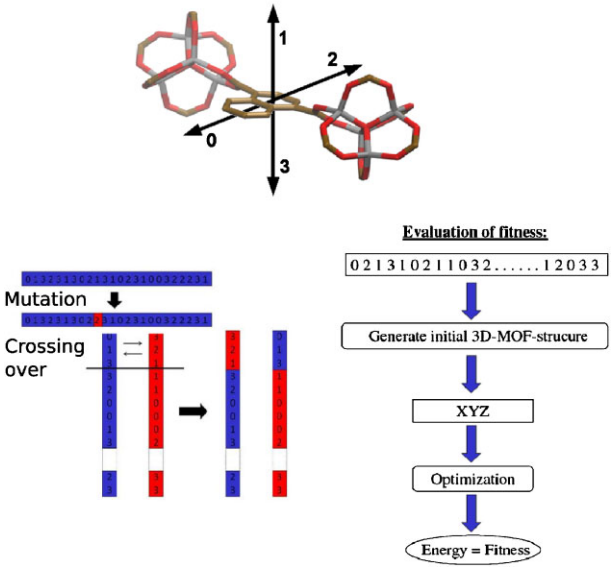


Figure 3 (online color at: www.pss-b.com) Schematic representation of the GA-based global minimum search of conformational isomers of IRMOF-7. The orientation of the naphthyl bridge with respect to the cell axes is encoded by an integer number. The genome of 24 linker orientations is subjected to mutation and cross-over operations.

From the results of benzene diffusion in MOF-5 we realized that in principle isomeric forms of MOF-5 should exist, where the Zn_4O tetrahedra are inserted in a different way. These isoreticular isomers share the same network topology, but have different physical properties. In particular, the self-diffusion for systems without the alternating pattern of A- and B-cells could be entirely different. Thus, already in 2008 we investigated the phenomenon of isoreticular isomerism in MOFs [18]. It was straightforward to extend the FF to treat also other representatives of the IRMOF series [53]. Especially for IRMOF-7 with its napthalene bridge in the linker we had to use a GA-based global minimum search strategy, as it is shown schematically in Fig. 3, to locate the lowest energy conformational isomer. From the results we could conclude, that isoreticular isomerism is energetically unfavorable for the parent MOF-5/IRMOF-1, but *e.g.*, for IRMOF-7 the energy difference between the isomers is about an order of magnitude lower. This results experimentally in a disorder of the position of the Zn_4O tetrahedra in the crystal structure of IRMOF-7.

Thus, the MM3 FF could successfully be extended for simulating MOF-5 and further IRMOFs. However, due to the wide variety of MOFs, it would be very time consuming and impractical to manually adjust parameters by hand for further systems. Thus, it was necessary to come up with an automated and systematic approach for force-field parametrization, which lead to the next generation of FFs.

2.2 A second generation force field: Systematic parameter fit using a genetic algorithm approach Due to the strong coupling of FF parameters and the fact that sometimes two different potential terms affect the same property, manual fitting is very difficult. In addition, many equally good FFs exist (local minima in the optimization space) and standard optimization methods mostly fail. Our target was to fit the bonded terms, with the non-bonded predetermined, on the basis of first principles reference data of non-periodic model systems. Importantly, not only the structural data but also the curvature information (matrix of second derivatives or Hessian) should be included into the fit. A new parameter set is generated for each MOF, sacrificing transferability for accuracy. There have been a number of approaches to parametrize an FF directly from the quantum chemical calculations (see references in Ref. [34]). GAs are a very efficient stochastic optimization method for problems such as conformational search or molecular docking (see also the previous section). GAs also have been widely used for FF parametrization. The major obstacle here is that in contrast to a conformational search, where the energy is the clear objective function, in case of FFs the target function is not so clearly defined. Often just structural deviations (sometimes in Cartesian space [54]) and partly relative energies have been used, ignoring curvature information. Structural comparison in internal coordinates is clearly superior to a comparison in Cartesian space, since in the latter case less relevant changes in *e.g.*, a dihedral angle can lead to large deviations. In principle, the Hessian can easily be projected into internal coordinate ("Z-matrix") space. This approach had been used in a single paper before [55]. However, the use of a non-redundant internal coordinate space is not unique and can break the symmetry *e.g.* for rings. Therefore, we developed an approach using a projection of structural and curvature information into a complete (redundant) space of internal coordinates, which is

naturally defined by the connectivity used in the FF. Note that the Hessian in redundant internal coordinate space contains non-zero elements mainly on and close to the diagonal, given that the internal coordinates are properly sorted topologically. We used a generalized inverse of the Wilson B-matrix [56] for the projection, which is unique for a given set of internal coordinates. For further details of the approach we refer here to Ref. [34]. The novel target function is the inverse of the mean-square-deviation (MSD) χ^2 between *ab initio* reference and FF for selected structural data (mainly bond distances and valence angles) and selected elements of the Hessian matrix (both diagonal and off-diagonal) in redundant internal coordinate representation. It is crucial to choose those internal coordinates (values and curvatures), which are actually influenced by the fitted parameters. The advantage of this strategy is, that parts of the parameter set can be fitted for a smaller model system and then kept fixed (and excluded from the fitness function) in a larger model, in order to add additional parameters.

The major difference in the FF expression for the ''second'' versus the ''first generation'' FF was to abandon the representation of Coulombic interactions by bond dipoles and to directly use atomic point charges. In addition, we used smaller formate model systems for the inorganic building blocks and combined it with parameters, derived from lithium or sodium carboxylates, in the ''building block'' type of approach. On this basis an FF for the Zn_4O-based MOFs like MOF-5 and the IRMOFs and a parametrization for copper paddle-wheel based MOFs like HKUST-1 was generated [34, 35, 57]. The resulting FF reproduced structures and lattice parameters of known MOFs like MOF-5 and HKUST-1 very well. More importantly, *e.g.*, for HKUST-1 the experimentally observed NTE [58] could quantitatively be predicted by the FF, demonstrating its accuracy also for complex dynamic properties of a porous framework. Very recently, the mechanical properties of HKUST-1 were investigated by nano-indentation of surface grown MOF thin films [59]. A very good match with the previous predictions of our FF was observed in this case. The copper paddle-wheel FF was later used to predict the steric preference of a given MOF for a network topology [60]. In this context also hypothetical topological isomers of known MOFs where structurally characterized. The calculations clarified why HKUST-1 with the smaller trimesic acid linker BTC prefers the tbo topology, whereas the extended linker in MOF-14 [61] leads to the – in this case ''interwoven'' – pto topology. It is very important to point out, that the potential of accurate and flexible FFs goes beyond the mere simulation of dynamic effects of known MOFs, but they allowed to explore and predict new and not yet synthesized systems, which could be of prime importance for a future truly rational design of MOFs.

At this point, it would have been possible to extend this second generation FF for further types of MOFs by fitting parameters for the corresponding inorganic building blocks. However, the combination of these parameters, fitted to first principles computed reference data, with those from the MM3 parameter set for the organic linker part is still somewhat inconsistent. On the other hand, the GA approach allowed us now to efficiently parametrize also the linker part. Thus, we decided to abandon the MM3 parameter set completely. Because of the necessary reparametrization this allowed us also refine the underlying energy expression and to remove all the obsolete or problematic legacy from the MM3 FF. In particular, different from most other FFs, the vdW interaction site of hydrogen is shifted along the X–H bond by a factor of 0.923 in MM3. This has been introduced already by Williams [62] in order to better describe the location of the X–H bond electron density. However, this introduces a torque on the X–H system from the vdW interactions of the hydrogen atom and makes it necessary to consider connectivity in the computation of the non-bonded interactions. For the first and second generation FFs, all the calculations have been performed with a modified version of the TINKER program package [63], which is one of the few MM codes that properly implements the MM3 FF. The code is non-parallel and limits to system sizes of about 3000 atoms, whereas more efficient parallel molecular dynamics packages like DL-Poly [64, 65] or LAMMPS [66] do not provide this shift of the vdW interaction site. Thus, we decided to discard this feature of MM3 for the third generation MOF-FF.

3 Definition of MOF-FF

3.1 Overview Essentially, the MOF-FF energy expression is still derived from the form originally proposed in MM3. Pairwise non-bonded vdW interactions (E^{vdW}) represent atomic repulsion and dispersive attraction. Electrostatic interactions (E^{coul}) between the atoms are modeled by spherical Gaussian charge distributions, instead of bond dipoles used in MM3. These had already been replaced by atomic point charges in the second generation FF. Bonded terms are described by bond (E^{stretch}), angle (E^{bend}) and torsion (E^{tors}) interactions plus out-of-plane (opb) bending (E^{opb}) for trigonal centers. In addition, a limited number of cross terms ($E^{\mathrm{str-str}}$; $E^{\mathrm{str-bnd}}$) are included mainly to improve the agreement with the reference vibrational frequencies. The overall energy expression can thus be written as

$$E_{\mathrm{MM}} = \sum_b E_b^{\mathrm{stretch}} + \sum_a \left(E_a^{\mathrm{bend}} + E_a^{\mathrm{str-str}} + E_a^{\mathrm{str-bnd}}\right) + \sum_o E_o^{\mathrm{opb}} + \sum_t E_t^{\mathrm{tors}} + \sum_n \left(E_n^{\mathrm{vdW}} + E_n^{\mathrm{coul}}\right), \quad (1)$$

where b runs over all bonds, a over all bond angles, o over all out-of-plane bends on trigonal centers, and t over all torsions. In the following each of these terms will be discussed.

3.2 Non-bonded van der Waals interactions Following our previous work, we use the vdW parametrization of Allinger et al., tabulated in Ref. [67], in MOF-FF as well. In contrast to other FFs, which often use a much more

repulsive LJ type potential, a Buckingham term with an exponential repulsive part is employed. Interestingly, the LJ term with its r^{-12} repulsive term has initially been used for numeric reasons since it is the square of the attractive dispersive r^{-6} term. However, current codes like DL-Poly use precomputed lookup-tables for the vdW interactions, and arbitrarily complex functions can be used without any problem for the numerical efficiency. In MM3 and our previous FFs the following term was used

$$E_{ij}^{\mathrm{vdW,MM3}} = \varepsilon_{ij}\left[1.84\times 10^{5}\exp\left(-12\frac{d_{ij}}{d_{ij}^{0}}\right) - 2.25\left(\frac{d_{ij}^{0}}{d_{ij}}\right)^{6}\right], \quad (2)$$

where d_{ij} is the interatomic distance and d_{ij}^{0} and ε_{ij} are the vdW minimum distance and well depth, respectively. This form introduces a fixed steepness of the repulsive part and reduces the three parameters of the Buckingham potential to the usual two, as in most FFs. Together with Lorentz–Berthelot mixing rules (arithmetic mean for minimum distance and geometric mean for the well depth of non-equal atom pairs) a good model for the non-bonded vdW interaction is found. Also in accord with MM3 we discard any vdW interactions between 1–2 and 1–3 bonded atoms, if not stated otherwise (an exception of this is used in the coordination environment, discussed below). By an extensive benchmark we could show that this holds, beyond the purely organic linker–linker interactions, also for the interaction of an aromatic guest molecule and an inorganic building block [68]. In Fig. 4 the potential curves for our FFs using the MM3 Buckingham potential (“Buck”) are compared with a quantum mechanic reference computed by a dispersion corrected double-hybrid functional. For both, the smaller formate model, and the larger system with three phenyl rings, the Buckingham potential reproduces the reference surprisingly well. The too repulsive LJ potential are also visible in this comparison.

At this point, two further issues need to be discussed. First, pairwise vdW potentials fail to account for many-body effects, which are known to be substantial. For example, in case of crystalline benzene, the dispersive interaction is reduced by about 10% by three-body terms [69]. Highly accurate intermolecular interaction energies are, for numerical reasons, often only available for smaller systems like dimers. As a consequence, for an accurate parametrization of vdW interactions of liquid or bulk systems (high density) usually experimental data needs to be used for the fit, making the pairwise terms an effective potential for the true many-body interaction. An example is the TraPPE FF [70], which is fitted with respect to vapor–liquid equilibria and is used in adsorption simulations of MOFs [10]. On the other hand, parametrization *e.g.* for molecular conformations (gas phase, low density) often overestimate the interaction in condensed phases. Second, the current approach – which is common for the most FFs – neglects the short-range electrostatic interactions due to charge penetration effects. Thus, the “repulsive” part of the current vdW potential is actually a combination of this attractive electrostatic potential (ESP) and the physically correct stronger repulsive part. An improved treatment of non-bonded interactions has recently been suggested by Tafipolsky and Engels [71].

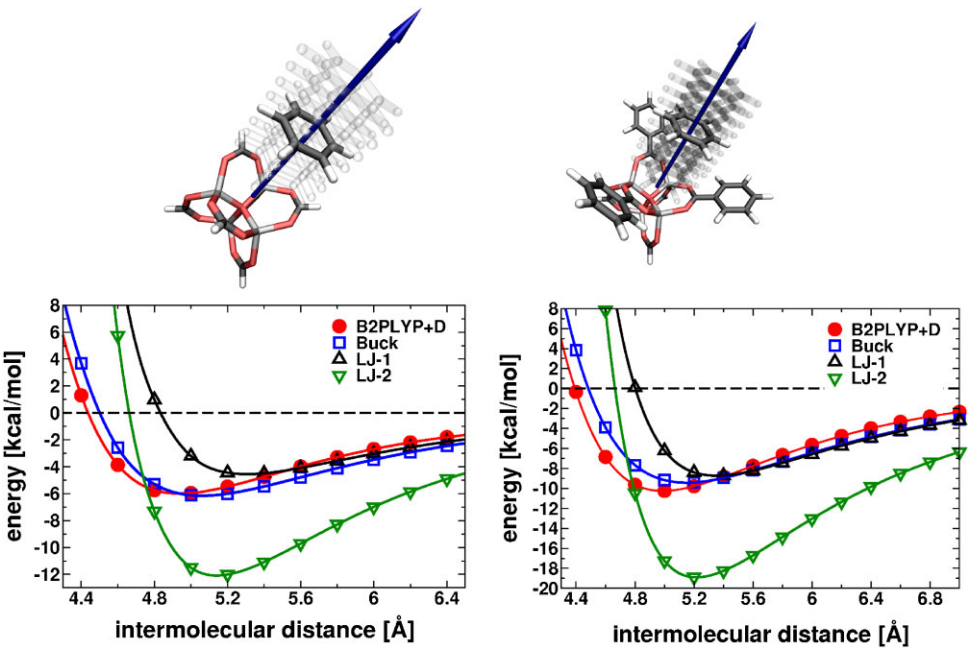


Figure 4 (online color at: www.pss-b.com) Interaction energy of benzene with two different models for the inorganic building block of MOF-5 for a quantum mechanic reference and three published FFs for benzene in MOF-5 (Reprinted with permission from Ref. [68]. Copyright 2012 American Chemical Society).

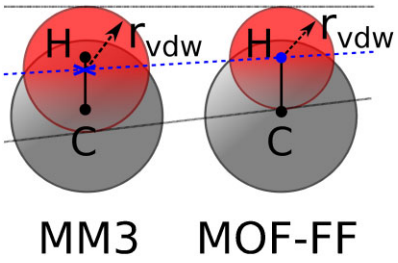


Figure 5 (online color at: www.pss-b.com) The vdW radius of hydrogen is reduced in MOF-FF to compensate the inward shift used in MM3.

As already mentioned in the previous section, a complication arises from the fact, that in MM3 the vdW interaction site of hydrogen is shifted toward the adjacent heavy atom. Since this shift is not implemented in DL-Poly and most other FF codes, we have discarded it and instead reduced the vdW radius of hydrogen from 1.62 Å in MM3 to 1.50 Å in MOF-FF, as schematically shown in Fig. 5. To validate this modification, we optimized structure and lattice parameters of bulk benzene with the original MM3 and MOF-FF (see Table 1). The results reveal that reducing the vdW radius of hydrogen affects the intermolecular interactions of benzene only slightly. However, by comparison with the experimental results, MOF-FF actually improves the results, justifying our approach.

In the MOF-FF expression for the vdW interaction we addressed a further problem, which arises from the Buckingham potential itself. Due to the fact that the r^{-6} term rises more quickly then the exponential repulsive part, the original Buckingham potential, as given in Eq. (2), reaches a maximum at very short distances and falls off to negative infinity. For regular MD simulations, this artifact is not a problem, because the barrier is usually not thermally accessible. However, for soft vdW potentials and high temperatures, numerical instabilities can arise. This gets even worse if the FF should be used in GCMC calculations, where close contacts can result from random insertions of guest molecules. Note that this is another reason for the wide spread use of the LJ potential, which does not suffer from this artifact. In order to cure this problem we have borrowed from the dispersion corrections to DFT introduced recently [73, 74]. There, in order to avoid double counting, the dispersive interaction is switched off smoothly at close distances. We used the same switching function, with a cutoff defined by the vdW radius. As a consequence, dispersive interactions vanish at short distances and the potential is completely repulsive. The final expression for the vdW interactions in MOF-FF is a dispersion damped Buckingham potential defined as follows:

$$E_{ij}^{\mathrm{vdW,MOF-FF}} = \varepsilon_{ij}\left\{1.85\times10^{5}\exp\left(-12\frac{d_{ij}}{d_{ij}^{0}}\right) - 2.25\left(\frac{d_{ij}^{0}}{d_{ij}}\right)^{6}\left[1+6\left(\frac{0.25d_{ij}^{0}}{d_{ij}}\right)^{14}\right]^{-1}\right\}. \quad (3)$$

To illustrate this, the vdW interaction energy between two sp^2 carbon atoms with identical parameters are plotted in Fig. 6 for the LJ potential, the original MM3 Buckingham term and the dispersion damped Buckingham in MOF-FF for comparison. Clearly, the artifact at short distances for the Buckingham potential is removed without noticeable changes in the region of the minimum (see inset). Furthermore, the less repulsive part as compared to the LJ potential is also obvious. We believe this to be the optimal solution until an efficient first principles parametrization is available to supersede the original MM3 parametrization. We are currently working on an approach, combining a fitted repulsive term with the dispersion term introduced by Grimme et al. in their most recent incarnation of dispersion corrections "D3" [73].

3.3 Electrostatic interactions Instead of point charges, used in the second generation FFs, in MOF-FF spherical Gaussian type charge distributions are employed. This effectively dampens the electrostatic interactions for high charges at close distances. Such a situation occurs for example in case of MOFs, where large and alternating charges are often found in the inorganic building blocks. As a consequence, we are now able to include all atom-atom interactions into the Coulomb sum, and the usual 1–2 and 1–3 exclusion can be skipped. Note, that many FFs use a scaled

Table 1 Cell parameters of solid benzene.

method	a (Å)	b (Å)	c (Å)	volume (Å^3)
MM3 (2000)	7.037	9.638	7.138	484.11
MOF-FF	7.471	9.194	6.790	466.40
expt. [72]	7.360	9.375	6.703	462.52

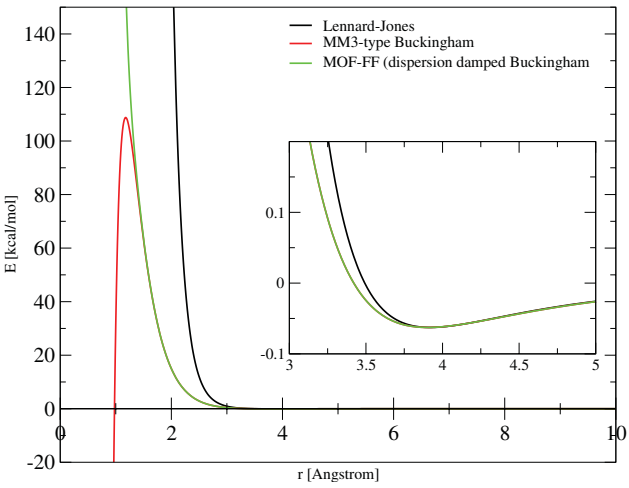


Figure 6 (online color at: www.pss-b.com) Comparison of vdW interaction potentials for two carbon atoms. Within the inset, the curves of MM3 and MOF-FF vdW potential coincide.

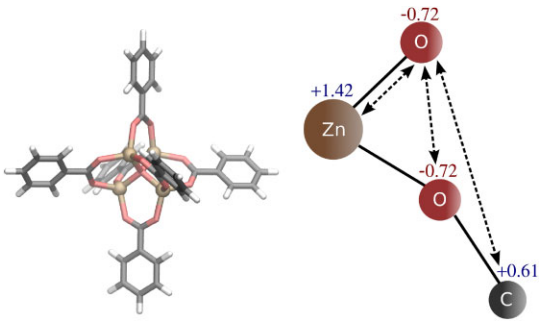


Figure 7 (online color at: www.pss-b.com) Coulombic interactions between the highly charged metal atom and the carboxylate atoms on the example of the Zn_4O SBU present in MOF-5. Note the alternating sign of the charges and the resulting compensations between 1–2, 1–3, and 1–4 interactions, respectively.

Coulomb interaction for 1–4 interactions. We found this to be crucial *e.g.*, for the inorganic building block of MOF-5, where a proper parametrization can only be achieved with a somewhat damped 1–4 interaction (see also Fig. 7). In case of Gaussian charge distributions and inclusion of all pairs, this was no longer needed. However, this means that in all bond and angle interactions a certain Coulombic contribution is included. Without the automated parameter fitting by the GA approach, this would not be feasible. A further benefit of not excluding any interactions is the possibility to extend the FF with a fluctuating charge approach, like, *e.g.*, the QTPIE method proposed by Chen and Martínez [75]. We are working on a polarizable extension of MOF-FF on this basis. However, in the current incarnation of MOF-FF fixed atomic charges are employed. They are derived from a fit of the ESP by using the sampling scheme proposed by Merz and Kollmann [49]. The following Coulomb interaction term is used in MOF-FF:

$$E_{i,j}^{\text{coul}} = \frac{1}{4\pi\varepsilon} q_i q_j \frac{\operatorname{erf}\left(\frac{d_{ij}}{\sigma_{ij}}\right)}{d_{ij}}. \tag{4}$$

Here d_{ij} is the distance of the two atoms of interaction n, q_i, and q_j are the atoms charges. The Gaussian charge distribution width σ_{ij} is computed by $\sqrt{\sigma_i^2 + \sigma_j^2}$. The values are taken from a fit to Slater functions by Chen and Martínez [75].

3.4 General bonded terms and cross terms For the regular bond stretches, the quartic term from MM3

$$E_b^{\text{stretch}} = \frac{1}{2} k_b (r_b - r_b^{\text{ref}})^2 \times \left[1 - 2.55(r_b - r_b^{\text{ref}}) + \frac{7}{12}\left(2.55(r_b - r_b^{\text{ref}})\right)^2\right] \tag{5}$$

is also used in MOF-FF. Here r_b is the bond distance, k_b the force constant, and r_b^{ref} is the corresponding reference distance. It contains anharmonic contributions in a fixed form, and only the two parameters (force constant and equilibrium value) need to be determined. In the same way, for valence angle bendings the six-order polynomial form of MM3 is used

$$E_a^{\text{bend}} = \frac{1}{2} k_a (\theta_a - \theta_a^{\text{ref}})^2 \times \left[1 - 0.14(\theta_a - \theta_a^{\text{ref}}) + 5.6 \times 10^{-5}(\theta_a - \theta_a^{\text{ref}})^2 - 7 \times 10^{-7}(\theta_a - \theta_a^{\text{ref}})^3 + 2.2 \times 10^{-8}(\theta_a - \theta_a^{\text{ref}})^4\right], \tag{6}$$

where θ_a is the bond angle, k_a the force constant, and θ_a^{ref} the reference angle. Note that in MM3 for trigonal centers the angle θ_a is taken to be the angle in the projection onto the trigonal plane. In MOF-FF we have dropped this additional complexity and directly use the interatomic angle. Correspondingly, for these sites an additional out-of-plane bending term, using the Wilson–Decius definition for the out-of-plane angle θ_o, is employed. In contrast to MM3, where again the six-order bending potential is used for the out-of-plane bending, in MOF-FF a simpler purely harmonic potential is used, together with a unique value of the force constant k_o for each trigonal center:

$$E_o^{\text{opb}} = \frac{1}{2} k_o \theta_o^2. \tag{7}$$

Torsions are computed in MOF-FF using the usual Fourier series term

$$E_t^{\text{tors}} = \sum_n \frac{V_t^n}{2}[1 + \cos(n\tau_t + \tau_t^n)], \tag{8}$$

where V_t^n and τ_t^n are the energy barrier and the phase shift, respectively, for the n-fold term with τ_t being the torsion angle of torsion t. In MOF-FF multiplicities up to $n = 4$ are supported.

Further changes with respect to the original MM3 concern the cross terms. First of all, the bend-bend and stretch-torsion cross terms, employed in MM3, have been completely discarded. Note, that the stretch–torsion term in MM3 is actually ill defined [76]. However, especially to improve the GA-based FF fit, stretch–stretch and stretch–bend cross terms were found to be beneficial [57]. In difference to MM3, MOF-FF allows to use two different stretch–bend force constants k_{sb1} and k_{sb2} for asymmetric situations. These cross terms are only included for adjacent bonds [34]. For three connected atoms A–B–C with a bond angle θ_a the distance r_{a1} refers to the distance A–B, and r_{a2} to the distance B–C. Note, that the reference distances r_{a1}^{ref} and r_{a2}^{ref} are the same as r_b^{ref} for the same bond type.

$$E_a^{\text{str-str}} = k_{ss}(r_{a1} - r_{a1}^{\text{ref}})(r_{a2} - r_{a2}^{\text{ref}}), \tag{9}$$

$$E_a^{\text{str-bend}} = (\theta_a - \theta_a^{\text{ref}}) \times \left[k_{sb1}(r_{a1} - r_{a1}^{\text{ref}}) + k_{sb2}(r_{a2} - r_{a2}^{\text{ref}})\right]. \tag{10}$$

3.5 Special force field terms for coordination environments In case of coordination bonds (like M–O and M–N) in the inorganic building blocks, somewhat different terms are employed in MOF-FF. We found that for the softer coordination bonds, the predefined anharmonicity of the stretch term (Eq. 5) is not sufficient [11]. Here we use a superior Morse-bond potential

$$E_{b,\mathrm{coord}}^{\mathrm{stretch}} = \frac{1}{2\alpha^2} k_b \left[1 - e^{-\alpha(r_b - r_b^{\mathrm{ref}})}\right]^2. \quad (11)$$

In this case, the additional parameter α is calculated using the relationship $E_{\mathrm{diss}} = k_b/2\alpha^2$, where E_{diss} is the approximate dissociation energy for the corresponding bond. Note, that the GA method is not able to fit this parameter, and it needs to be specified in advance.

In coordination complexes with coordination numbers beyond four or in case of a square planar coordination, polynomial angle bending terms with a single minimum cannot be used [77]. In the second generation FF we have, therefore, used a special Fourier-type angle bending term proposed by Rappe et al. [78], which avoids an artificial minimum at zero degrees. However, in case of MOF-FF we include not only 1–3 electrostatic interactions, but – only in a coordination environment – also 1–3 vdW interactions, which prevents these low angles due to atomic repulsion. As a consequence, the simpler Fourier term

$$E_{a,\mathrm{coord}}^{\mathrm{bend}} = \frac{V_a}{2}\left[1 + \cos(n\theta_a + \theta_a^0)\right] \quad (12)$$

is used, where V_a is the energy barrier, θ_a^0 the phase shift, and n is the multiplicity. Due to the presence of vdW and electrostatic 1–3 interactions, this term can actually be neglected (or a very small V_a can be used) if no orbital directing effects are present. This is the case, *e.g.*, in the zinc paddle-wheel systems with the Zn(II) in a d^{10}-electron configuration. In contrast to that, the corresponding Cu(II) paddle-wheel with their d^9-configuration strongly prefer a square planar coordination [16].

3.6 Technical implementation: The pydlpoly code We decided to use the DL-Poly 2 (V2.20) package [79], which is implemented in a work-parallel fashion, but retains the data like atomic positions or velocities redundant on all nodes. It is suitable for simulations of up to 30 000 atoms. With its simpler parallelization scheme, as compared to the domain decomposition used in DL-Poly 3 [65] or in LAMMPS [66], it was easier to adapt the code to our needs. For the implementation of MOF-FF, the MM3-type stretch and bend terms (Eqs. 5 and 6), the MOF-FF vdW term (Eq. 3), and the Fourier bending term for coordination environments (Eq. 12) had to be added, whereas all other terms are already supported by the *DL-Poly 2* code. In addition, the use of Gaussian charge distributions (Eq. 4) was implemented for Ewald-, SPME-, and cutoff-based schemes to compute the Coulombic interactions. Note that the potential energy terms can be implemented in a similar way in any other FF code.

In addition to the changes in the potential terms, the way of using the code was completely restructured. *DL-Poly 2* is written in the *Fortran90* programming language as a monolithic binary, which reads a certain input, performs a simulation task and ends. By wrapping the code, using the *f2py* code generator [80], it is controlled now by an object oriented scripting language *Python* [81, 82]. All data structures like atom positions or forces are accessible from *Python* and complex tasks can be realized. The derived code *pydlpoly* is fully parallelized and as efficient as the original code. Most importantly, *pydlpoly* automatically parses molecular structure and FF definition files and generates the *DL-Poly* FIELD-file (which defines all individual potential energy terms) on the fly. In addition, the *Python* wrapper was used to add an efficient low memory L-BFGS geometry optimizer, since *DL-Poly* itself does not provide efficient routines for geometry optimization.

4 Parametrization of MOF-FF

4.1 Reference calculations of the model systems The parametrization scheme for MOF-FF is based on the building block approach, and non-periodic model systems are used as a reference. However, because of the closer similarity with respect to the true periodic systems and the inclusion of dispersion correction in our reference calculations, we went back to the larger benzoate models as opposed to the formate models, which we used for the second generation FFs, if numerically possible. The model systems were completely optimized and the Hessian matrix was computed using the TURBOMOLE (V6.3) [83] package. The hybrid functional B3LYP was used, throughout. A finer numerical integration grid (''m5'') was employed to improve the computed low frequency normal modes. For main group elements (C, H, N, O) cc-pVDZ basis sets were used [84], and in case of transition metals (Cu, Zn, Zr) the corresponding valence basis sets cc-pVDZ-PP together with the appropriate Stuttgart-Dresden effective core potentials (ECPs) [85, 86] were employed. If permissible, the augmented versions of the basis sets were used to improve the description of the carboxylates [57]. In contrast to the previous FF generations, the most recent D3 dispersion correction according to Grimme et al. was added [73]. The Hessian matrix was computed analytically in all cases (*aoforce*). The effective atomic charges were obtained from a fit of the ESP by the Merz–Kollman sampling scheme [49]. Here, vdW exclusion radii of Cu, Zn, and Zr of 1.960, 1.946, and 2.80 Å were used, respectively. In general, we slightly modified the charges, resulting from the ESP fit, in order to generate ''neutral groups''. For example, in case of the Zn_4O building block, atomic charges were altered by ± 0.02 or less, in order maintain the Zn_4O core together with the carboxylate groups and the α-C atoms of the phenylene groups in total charge neutral.

In Fig. 8 the geometries of the inorganic model systems are shown. Note that due to the use of benzoate models, the

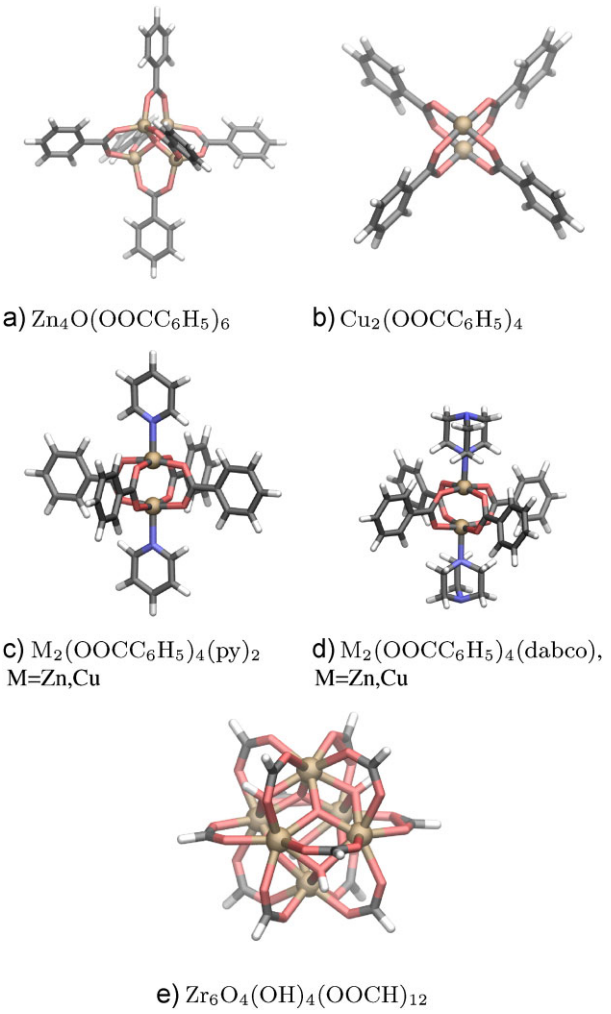


Figure 8 (online color at: www.pss-b.com) Non-periodic model systems of the inorganic building blocks currently covered by MOF-FF.

lithium or sodium carboxylate models were not needed for MOF-FF. However, in order to assemble a wide variety of different MOFs, also some purely organic model systems were included. In Fig. 9 a selection of possible linkers is shown for illustration. Note that the above described "additive" parametrization scheme is also used here (Section 2.2). For example, the parameter for the phenylene groups were derived from isolated benzene and not refitted for each system. This parameter set for the organic linkers can already cover a wide variety of systems, but it is also constantly extended. Because of the consistent parametrization scheme, all linker FF parameters can be combined with any of the inorganic building blocks. The parameter set does not grow quadratically with the number of combinations of inorganic building blocks and linkers, included. Furthermore, it is possible to simulate for example MOFs with different inorganic building blocks [87] or core–shell systems with varying linkers [88]. In the following we have limited ourselves – in order to demonstrate the application of

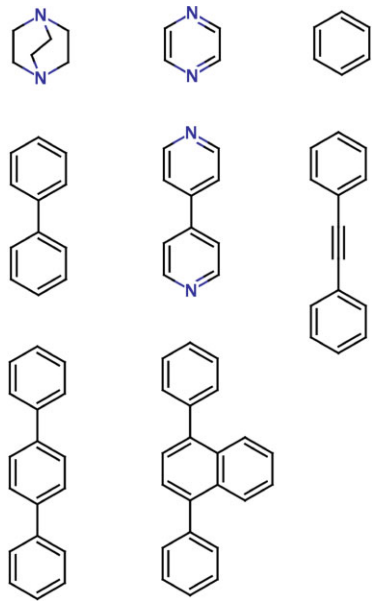


Figure 9 (online color at: www.pss-b.com) Selection of organic linker backbones that can be treated by the current parameter set of MOF-FF.

MOF-FF – on experimentally known and well-studied MOFs.

4.2 GA parameter optimization After defining the atomic charges and selecting the potential terms (*e.g.*, cross terms) the GA optimizer is used to optimize the parameter set. As already mentioned in Section 2.2, it is crucial to include only those internal coordinates into the fitness function, which correspond to actually fitted parameters. A further requirement is that the chosen potential energy function is able to reproduce the local shape of the potential energy surface of the DFT reference. Otherwise a stall of the optimization is observed. As a consequence, a clear and noticeable rise of the target function in the evolution of the GA is a clear indication that the fit will give a good result. In some cases, the allowed ranges for the parameters in the GA need to be adjusted. We also found that the optimization can be improved by running a number of independent GA runs and combine the best species for a final run. In all cases, the final parameter set is tested to reproduce both geometry and vibrational normal modes of the DFT reference. In general, deviations in bond length and valence angles are below 0.5% and low normal modes deviate up to around 10 cm^{-1}. In rare cases, larger deviations are due to missing cross-terms (not available in MOF-FF) and not to an insufficient fit.

In order to illustrate a successful fit, in Fig. 10 the development of the target function during a typical GA run is shown for the case of the Zn_4O building block. A total of 4 stretch, 7 bend, 3 torsion, 2 opb, and 5 cross-terms had to be fitted. Correspondingly, 11 internal coordinate values (bond distances and angles) and 16 elements of the redundant internal coordinate Hessian matrix were included in the

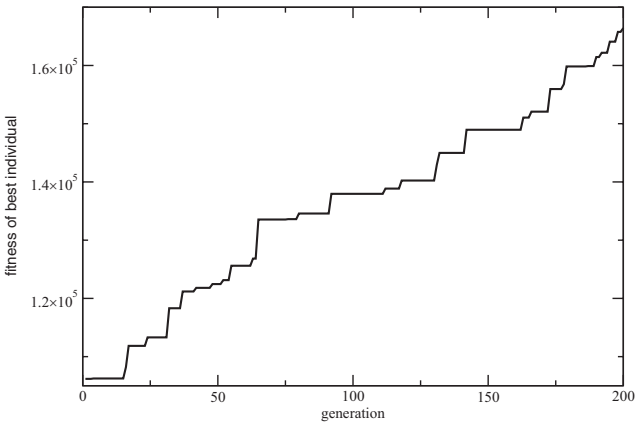


Figure 10 Development of the objective function during a typical fit for the $Zn_4O(O_2CC_6H_5)_6$ model system.

fitness function. A detailed listing of the final quality of the fit for each element of the target function is given in the Supporting Information (online at www.pss-b.com).

4.3 MOF-FF parameter set For brevity, the detailed parameter tables are listed in the Supporting Information. Since this set of parameters will constantly extended and improved, we recommend to contact the authors for a most recent version. In the following, further details with respect to the individual "families" of MOFs (basically defined by their inorganic building block) will be given.

Zn_4O-based MOFs For the IRMOFs and UMCMs we used, as in the "first generation", the benzoate $Zn_4O(OOCC_6H_5)_6$ as a basis for the FF parametrization. The model was optimized in T_d-symmetry in order to reduce the numerical effort. The previously fitted benzene parameters were transferred to the phenyl part. In order to maintain neutral groups, the charge of the α-carbon connected to C_{coo} had to be increased from 0.12 to 0.18. The coordination bond between Zn–O was described with a Morse potential with a dissociation energy E_{diss} of 50.0 kcal mol^{-1}. Due to a strong coupling of Zn and O in the inorganic part, stretch-bend cross terms on Zn–O_{cen}–Zn and O_{cen}–Zn–O_{carb} were also considered in the fitting procedure.

Copper paddle-wheel based MOFs with open metal sites A paddle-wheel Cu_2 $(OOCC_6H_5)_4$ molecular unit in D_{4h}-symmetry was used for the parametrization. The structure provides two open metal sites, which are very useful for catalytic applications. According to our previous experience, the paddle-wheel structure of d^9-Cu could be preserved due to an orbital directing effect. Note that with Zn (d^{10}) it would be difficult to obtain Zn paddle wheel without axial ligands [16].

Similar to the Zn_4O case, the benzene parameters were transferred and the charge of the α-carbon was increased to 0.15. A Morse potential is used for the Cu–O_{carb} bond stretch with 50 kcal mol^{-1} dissociation energy. Between the two Cu ions a bond term was introduced, which has, however, a small force constant compared with other covalent bonds. Since the angle of O_{carb}–Cu–O_{carb} has multiple minima (90°, 180°, 270°), the Fourier-type angle bending potential function is used. Moreover, due to the very close distance of the oxygen atoms, for the O_{carb}–Cu–O_{carb} angle bending the vdW 1–3 interaction was explicitly included.

Paddle-wheel based MOFs (M = Cu,Zn) with pillar linkers In this case, the two axial positions of the paddle-wheels are coordinated with a pillar linker. MOF-FF is parametrized to treat both sp^3 donor nitrogen pillar linkers like in dabco or sp^2 nitrogen donor linkers as bipyridine. Pillar linker parameters were separately determined and then transferred to the reference models $M_2(OOCC_6H_5)_4(dabco)_2$ and $M_2(OOCC_6H_5)_4(py)_2$. The coordination environment is treated analogous to the open metal site paddle-wheel with a Fourier-angle bending term, as well as explicit inclusion of the vdW 1–3 interactions between the oxygen atoms. The coordination bonds M–O_{carb} and M–N_{dab} were described with a Morse potential with a bond dissociation energy of 50.0 and 25.0 kcal mol^{-1}, respectively.

$Zr_6O_4(OH)_4$-based MOFs Because of the size of the benzoate model, $Zr_6O_4(OH)_4$ $(OOCH)_{12}$ in T_d-symmetry was considered as a reference model for the UIO-type MOFs. All Zr–O bonds are described by a Morse potential with a dissociation energy of 50.0 kcal mol^{-1}. In this case, the remaining parameters to treat aromatic carboxylates were taken from Zn_4O system.

4.4 Application of MOF-FF The current parameter set has been used to optimize the structure of a number of well-known MOF systems shown in Fig. 11. All calculations have been performed with the *pydlpoly* code using periodic boundary conditions. We optimized the systems without symmetry constraints (P_1) and the result corresponds to the zero Kelvin structure. All structure and lattice optimizations were carried out with the short range forces cut off of 12.0 Å. Lattice parameters of MOF-5, HKUST-1, $[Zn_2(bdc)_2(dabco)]$, $[Cu_2(bdc)_2(dabco)]$, UIO-66 obtained from MOF-FF together with the experimental values are shown in Table 2. All calculated numbers show reasonably good agreement with experimental observations. Similar deviations are found if periodic DFT calculations are used with similar functionals as used in our parametrization. Note that the crystal data of $[Cu_2(bdc)_2(dabco)]$ is not available due to the difficulty in growing single crystals of this material. However, the cell parameter should be comparable with the analog $[Zn_2(bdc)_2(dabco)]$.

5 Conclusions We have presented a new, fully flexible FF for MOFs, which is consistently parametrized on the basis of first principles reference data. Because of the GA fitting procedure it can be seen as the best possible

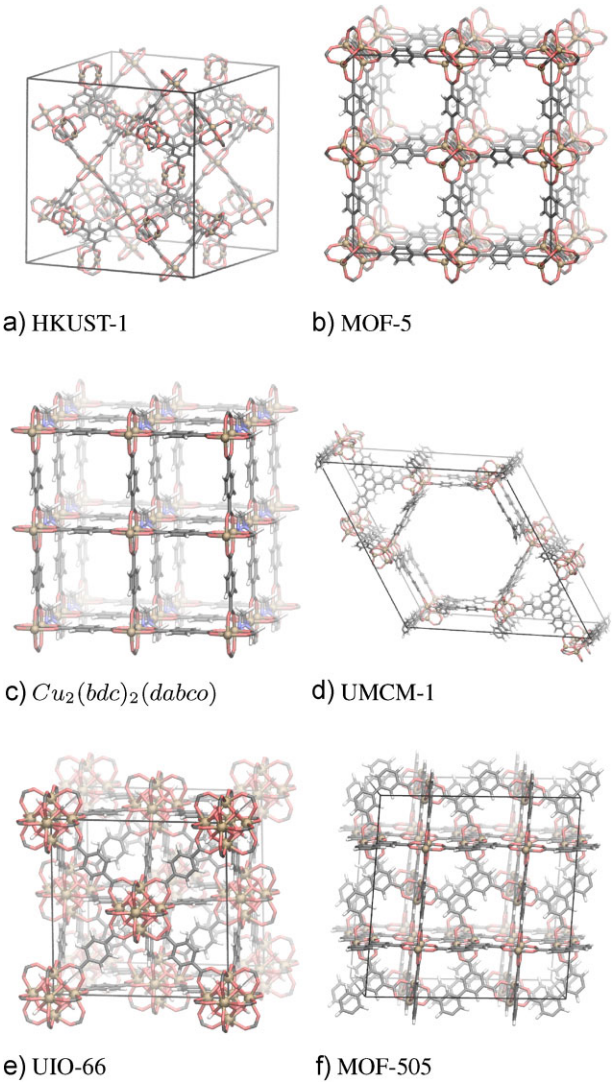


Figure 11 (online color at: www.pss-b.com) Structures of the optimized MOFs.

Table 2 Cell parameters of MOFs calculated from MOF-FF compared with available experimental data.

	MOF-FF a (c) (Å)	exptl. (XRD) a (c) (Å)	Ref.
MOF-5	26.08	25.885	[47]
HKUST-1	26.43	26.3046	[89]
$[Zn_2(bdc)_2(dabco)]$	11.01 (9.75)	10.9288 (9.6084)	[90]
$[Cu_2(bdc)_2(dabco)]$	10.92 (9.60)	–	–
UiO-66	20.90	20.7004	[91]

representation of the first principles potential energy surface in the vicinity of the reference structure for the given FF energy expression. Further improvements in accuracy can only be achieved by extending this energy expression *e.g.*, with further cross terms. Deviations in bond length and angles are generally below 0.5% and low normal modes differ in most cases less than 10 cm^{-1} compared to the DFT reference data. In contrast to other such FFs, the parameters can be exchanged and combined, because they have been generated by the same recipe. As a consequence, properties computed for different MOFs can safely be compared with each other. MOF-FF is currently parametrized for Zn_4O and $Zr_6O_4(OH)_4$-based MOFs and copper and zinc paddle-wheel systems with and without pillar linkers. Parameters for most common linkers as well as for aliphatic and aromatic guest molecules or CO_2 are available, but the parameter set will constantly be extended to further systems in the future. Already with the current parameter set, not only a large number of known MOFs can be investigated, but also yet unsyntesized systems can be simulated. Together with the reverse topological approach, introduced recently by us [60], this opens the road for a computational prescreening. The FF allows to compute structures, also in the presence of strain (large guest molecules or large linker substituents). Dynamic properties like elastic constants or thermal expansion can be predicted, and effects like "breathing" or the "gate opening" effects can be studied.

A further important field of application of MOF-FF is the use in a combined QM/MM treatment to study catalytic processes within MOFs. When the QM part is computed at a similar theoretical level as the reference for the MM parametrization, the FF acts as an ideal mechanical embedding.

Because MOF-FF uses bond terms for the metal–ligand interactions, it is not able to simulate processes, where the coordination of the metal atoms changes (including coordination of guest molecules to open metal sites). Future work is dedicated to an extension of MOF-FF to allow the breaking and forming of coordination bonds, however, with the inclusion of orbital directing effects.

Acknowledgements This work has been funded by the Deutsche Forschungsgemeinschaft (DFG) within the Sonderforschungsbereich SFB-558 at the Ruhr-University Bochum and SPP-1362. RS would like to thank Prof. Roland Fischer, Prof. Christoph Wöll, and Prof. Martin Muhler for inspiring discussions and continuous support. SA would like to thank the graduate school within SFB-558 for a generous scholarship.

References

[1] G. Ferey, Chem. Soc. Rev. **37**, 191–214 (2008).
[2] G. Ferey and C. Serre, Chem. Soc. Rev. **38**, 1380–1399 (2009).
[3] S. L. James, Chem. Soc. Rev. **32**, 276 (2003).
[4] C. Janiak, Dalton Trans. 2781–2804 (2003).
[5] S. Kitagawa, R. Kitaura, and S. Noro, Angew. Chem., Int. Ed. **43**, 2334–2375 (2004).
[6] L. J. Murray, M. Dinca, and J. R. Long, Chem. Soc. Rev. **38**, 1294–1314 (2009).
[7] J. L. C. Rowsell and O. M. Yaghi, Microporous Mesoporous Mater. **73**, 3–14 (2004).

[8] H. C. Zhou, J. R. Long, and O. M. Yaghi, Chem. Rev. **112**(2), 673–674 (2012).
[9] S. Keskin, J. Liu, R. B. Rankin, J. K. Johnson, and D. S. Sholl, Ind. Eng. Chem. Res. **48**, 2355–2371 (2009).
[10] T. Duren, Y. S. Bae, and R. Q. Snurr, Chem. Soc. Rev. **38**, 1237–1247 (2009).
[11] M. Tafipolsky, S. Amirjalayer, and R. Schmid, Microporous Mesoporous Mater. **129**, 304–318 (2010).
[12] D. Li and K. Kaneko, Chem. Phys. Lett. **335**(1–2), 50–56 (2001).
[13] R. Kitaura, K. Seki, G. Akiyama, and S. Kitagawa, Angew. Chem., Int. Ed. **42**(4), 428–431 (2003).
[14] S. Horike, D. Tanaka, K. Nakagawa, and S. Kitagawa, Chem. Commun. 3395–3397 (2007).
[15] B. Lukose, B. Supronowicz, P. St. Petkov, J. Frenzel, A. B. Kuc, G. Seifert, G. N. Vayssilov, and T. Heine, Phys. Status Solidi B **249**(2), 335–342 (2012).
[16] S. Bureekaew, S. Amirjalayer, and R. Schmid, J. Mater. Chem. **22**, 10249–10254 (2012).
[17] J. A. Greathouse and M. D. Allendorf, J. Am. Chem. Soc. **128**, 10678–10679 (2006).
[18] S. Amirjalayer and R. Schmid, J. Phys. Chem. C **112**, 14980 (2008).
[19] R. Schmid and M. Tafipolsky, J. Am. Chem. Soc. **130**(38), 12600–12601 (2008).
[20] A. K. Rappe, C. J. Casewit, K. S. Colwell, W. A. Goddard, and W. M. Skiff, J. Am. Chem. Soc. **114**, 10024–10035 (1992).
[21] S. L. Mayo, B. D. Olafson, and W. A. Goddard, J. Phys. Chem. **94**(26), 8897–8909 (1990).
[22] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, J. Am. Chem. Soc. **118**(45), 11225–11236 (1996).
[23] P. Dauber-Osguthorpe, V. A. Roberts, D. J. Osguthorpe, J. Wolff, M. Genest, and A. T. Hagler, Proteins: Struct. Funct. Genet. **4**, 31 (1988).
[24] J. A. Greathouse and M. D. Allendorf, J. Phys. Chem. C **112**, 5795 (2008).
[25] D. Dubbeldam, K. S. Walton, D. E. Ellis, and R. Q. Snurr, Angew. Chem., Int. Ed. **46**, 4496–4499 (2007).
[26] T. Loiseau, C. Serre, C. Huguenard, G. Fink, F. Taulelle, M. Henry, T. Bataille, and G. Férey, Chem. Eur. J. **10**(6), 1373–1382 (2004).
[27] F. Salles, A. Ghoufi, G. Maurin, R. G. Bell, C. Mellot-Draznieks, and G. Ferey, Angew. Chem., Int. Ed. **47**(44), 8487–8491 (2008).
[28] J. Cirera, J. C. Sung, P. B. Howland, and F. Paesani, J. Chem. Phys. **137**(5), 054704 (2012).
[29] D. S. Coombes, F. Cora, C. Mellot-Draznieks, and R. G. Bell, J. Phys. Chem. C **113**(2), 544–552 (2009).
[30] L. Zhao, Q. Yang, Q. Ma, C. Zhong, J. Mi, and D. Liu, J. Mol. Modell. **17**(2), 227–234 (2011).
[31] N. L. Allinger, Y. H. Yuh, and J. H. Lii, J. Am. Chem. Soc. **111**, 8551 (1989).
[32] N. L. Allinger, F. Li, L. Q. Yan, and J. C. Tai, J. Comput. Chem. **11**, 868–8895 (1990).
[33] M. Tafipolsky, S. Amirjalayer, and R. Schmid, J. Comput. Chem. **28**, 1169–1176 (2007).
[34] M. Tafipolsky and R. Schmid, J. Phys. Chem. B **113**, 1341–1352 (2009).
[35] M. Tafipolsky, S. Amirjalayer, and R. Schmid, J. Phys. Chem. C **114**(34), 14402–14409 (2010).
[36] J. S. Grosch and F. Paesani, J. Am. Chem. Soc. **134**(9), 4207–4215 (2012).
[37] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case, J. Comput. Chem. **25**(9), 1157–1174 (2004).
[38] C. D. Sherrill, B. G. Sumpter, M. O. Sinnokrot, M. S. Marshall, E. G. Hohenstein, R. C. Walker, and I. R. Gould, J. Comput. Chem. **30**(14), 2187–2193 (2009).
[39] W. L. Jorgensen and D. L. Severance, J. Am. Chem. Soc. **112**(12), 4768–4774 (1990).
[40] L. Vanduyfhuys, T. Verstraelen, M. Vandichel, M. Waroquier, and V. Van Speybroeck, J. Chem. Theory Comput. **8**(9), 3217–3231 (2012).
[41] H. M. El-Kaderi, J. R. Hunt, J. L. Mendoza-Cortés, A. P. Côté, R. E. Taylor, M. O'Keeffe, and O. M. Yaghi, Science **316**(5822), 268–272 (2007).
[42] S. Amirjalayer, R. Q. Snurr, and R. Schmid, J. Phys. Chem. C **116**(7), 4921–4929 (2012).
[43] S. Bureekaew and R. Schmid, CrystEngComm **15**, 1551–1562 (2013).
[44] S. Hermes, M. K. Schröter, R. Schmid, L. Khodeir, M. Muhler, A. Tissler, R. W. Fischer, and R. A. Fischer, Angew. Chem., Int. Ed. **44**(38), 6237–6241 (2005).
[45] D. Esken, H. Noei, Y. Wang, C. Wiktor, S. Turner, G. Van Tendeloo, and R. A. Fischer, J. Mater. Chem. **21**, 5907–5915 (2011).
[46] M. Tafipolsky and R. Schmid, Surf. Coat. Technol. **201**(22–23), 8818–8824 (2007), Euro CVD 16: 16th European Conference on Chemical Vapor Deposition.
[47] H. Li, M. Eddaoudi, M. O'Keeffe, and O. M. Yaghi, Nature **402**, 276 (1999).
[48] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).
[49] B. H. Besler, K. M. Merz, and P. A. Kollman, J. Comput. Chem. **11**, 431 (1990).
[50] S. Amirjalayer, M. Tafipolsky, and R. Schmid, Angew. Chem., Int. Ed. **46**(3), 463–466 (2007).
[51] S. Amirjalayer and R. Schmid, Microporous Mesoporous Mater. **125**(1–2), 90–96 (2009).
[52] F. Stallmach, S. Gröger, V. Künzel, J. Kärger, O. M. Yaghi, M. Hesse, and U. Müller, Angew. Chem., Int. Ed. **45**(13), 2123–2126 (2006).
[53] M. Eddaoudi, J. Kim, N. Rosi, D. Vodak, J. Wachter, M. O'Keeffe, and O. M. Yaghi, Science **295**, 469–472 (2002).
[54] T. Strassner, M. Busold, and H. Radrich, J. Mol. Model. **7**, 374–377 (2001).
[55] S. Mostaghim, M. Hoffmann, P. Konig, T. Frauenheim, and J. Teich, in: Proc. of the Congress on Evolutionary Computation (CEC'04), Portland, USA, 20–23 June 2004, pp. 212–219.
[56] P. Pulay and G. Fogarasi, J. Chem. Phys. **96**, 2856–2860 (1992).
[57] M. Tafipolsky and R. Schmid, J. Chem. Theory Comput. **5**, 2822–2834 (2009).
[58] V. K. Peterson, G. J. Kearley, Y. Wu, A. J. Ramirez-Cuesta, E. Kemner, and C. J. Kepert, Angew. Chem., Int. Ed. **49**(3), 585–588 (2010).
[59] S. Bundschuh, O. Kraft, H. K. Arslan, H. Gliemann, P. G. Weidler, and C. Woll, Appl. Phys. Lett. **101**(10), 101910 (2012).
[60] S. Amirjalayer, M. Tafipolsky, and R. Schmid, J. Phys. Chem. C **115**(31), 15133–15139 (2011).
[61] B. Chen, M. Eddaoudi, S. T. Hyde, M. O'Keeffe, and O. M. Yaghi, Science **291**, 1021 (2001).
[62] D. E. Williams, J. Chem. Phys. **43**, 4424 (1965).

[63] J. W. Ponder, P. Ren, R. V. Pappu, R. K. Hart, M. E. Hodgson, D. P. Cistola, C. E. Kundrot, and F. M. Richards, Tinker software tools for molecular design, version 4.2, June 2004, Washington University School of Medicine, 2004. Available at http://dasher.wustl.edu/tinker/.
[64] W. Smith and T. Forester, J. Mol. Graph. **14**(3), 136–1141 (1996).
[65] I. T. Todorov and W. Smith, Philos. Trans. A, Math. Phys. Eng. Sci. **362**(1822), 1835–1852 (2004).
[66] S. Plimpton, J. Comput. Phys. **117**(1), 1–19 (1995).
[67] N. L. Allinger, X. F. Zhou, and J. Bergsma, Theochem **118**, 69 (1994).
[68] S. Amirjalayer and R. Schmid, J. Phys. Chem. C **116**(29), 15369–15377 (2012).
[69] O. A. von Lilienfeld, and A. Tkatchenko, J. Chem. Phys. **132**(23), 234109 (2010).
[70] J. J. Potoff and J. I. Siepmann, AIChE J. **47**(7), 1676 (2001).
[71] M. Tafipolsky and B. Engels, J. Chem. Theory Comput. **7**(6), 1791–1803 (2011).
[72] G. A. Jeffrey, J. R. Ruble, R. K. McMullan, and J. A. Pople, Proc. R. Soc. Lond., Ser. A **414**(1846), 47–57 (1987).
[73] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, J. Chem. Phys. **132**(15), 154104 (2010).
[74] S. Grimme, J. Comput. Chem. **27**(15), 1787–1799 (2006).
[75] J. Chen and T. J. Martínez, Chem. Phys. Lett. **438**(4–6), 315–320 (2007).
[76] C. H. Lee and S. S. Zimmerman, J. Biomol. Struct. Dyn. **13**(2), 201–218 (1995).
[77] V. S. Allured, C. M. Kelly, and C. R. Landis, J. Am. Chem. Soc. **113**, 1 (1991).
[78] A. K. Rappe, L. M. Bormann-Rochotte, D. C. Wiser, J. R. Hart, M. A. Pietsch, C. J. Casewit, and W. M. Skiff, Mol. Phys. **105**, 301 (2007).
[79] W. Smith, C. Yong, and P. Rodger, Mol. Simul. **28**(5), 385–471 (2002).
[80] P. Peterson, Int. J. Comput. Sci. Eng. **4**(4), 296–305 (2009).
[81] D. Ascher, P. F. Dubois, K. Hinsen, J. Hugunin, and T. Oliphant, Numerical Python, Lawrence Livermore National Laboratory, Livermore, California, USA, 2001. Available at http://www.pfdubois.com/numpy/.
[82] G. van Rossum, and F. L. Drake (eds.), Python Reference Manual (PythonLabs, Virginia, USA, 2001). Available at http://www.python.org.
[83] TURBOMOLE V6.3 2011, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989–2007, TURBOMOLE GmbH, since 2007 available from http://www.turbomole.com.
[84] A. K. Wilson, D. E. Woon, K. A. Peterson, and T. H. Dunning, J. Chem. Phys. **110**, 7667–7676 (1999).
[85] P. Fuentealba, H. Preuss, H. Stoll, and L. V. Szentpály, Chem. Phys. Lett. **89**(5), 418–422 (1982).
[86] G. Igel-Mann, H. Stoll, and H. Preuss, Mol. Phys. **65**(6), 1321–1328 (1988).
[87] O. Shekhah, K. Hirai, H. Wang, H. Uehara, M. Kondo, S. Diring, D. Zacher, R. A. Fischer, O. Sakata, S. Kitagawa, S. Furukawa, and C. Woll, Dalton Trans. **40**, 4954–4958 (2011).
[88] S. Furukawa, K. Hirai, K. Nakagawa, Y. Takashima, R. Matsuda, T. Tsuruoka, M. Kondo, R. Haruki, D. Tanaka, H. Sakamoto, S. Shimomura, O. Sakata, and S. Kitagawa, Angew. Chem., Int. Ed. **48**(10), 1766–1770 (2009).
[89] V. K. Peterson, Y. Liu, C. M. Brown, and C. J. Kepert, J. Am. Chem. Soc. **128**, 15578 (2006).
[90] D. N. Dybtsev, H. Chun, and K. Kim, Angew. Chem., Int. Ed. Engl. **43**(38), 5033–55036 (2004).
[91] J. H. Cavka, S. Jakobsen, U. Olsbye, N. Guillou, C. Lamberti, S. Bordiga, and K. P. Lillerud, J. Am. Chem. Soc. **130**(42), 13850–13851 (2008).

Text