Systematic First Principles Parameterization of Force Fields for Metal−Organic Frameworks using a Genetic Algorithm Approach

Maxim Tafipolsky and Rochus Schmid*

Lehrstuhl für Anorganische Chemie 2, Organometallics and Materials Chemistry, Ruhr-Universität Bochum, Universitätsstrasse 150, D-44780 Bochum, Germany

Received: August 21, 2008; Revised Manuscript Received: October 30, 2008

A systematic strategy is proposed to derive the necessary force field parameters directly from first principles calculations of nonperiodic model systems to reproduce both the structure and curvature of the reference potential energy surface. The parameters are determined using a genetic algorithm combined with a novel fitness criterion based on a representation of structure and curvature in a set of redundant internal coordinates. Due to the efficiency of this approach it is possible to abandon the need for transferability of the parameters. The method is targeted for the application on metal−organic frameworks (MOFs), where parameters for molecular mechanics force fields are often not available, because of the wide range of possible inorganic fragments involved. The scheme is illustrated for Zn_4O−based IRMOF materials on the example of MOF-5. In a "building block" approach parameters are derived for the two model systems basic zinc formate ($Zn_4O(O_2CH)_6$), and dilithium terephthalate with reference data obtained from density functional theory. The resulting potential gives excellent agreement with the structure, vibrational frequencies, thermal behavior and elastic constants of the periodic MOF-5.

1. Introduction

The recently developed nanoporous metal-organic frameworks (MOFs) have attracted considerable interest owing to their potential applications in sensing, gas storage, chemical separations and catalysis.[1−4] The many useful properties of this new class of coordination polymers are due to their hybrid nature: they are built from metal (or metal−oxide) clusters interconnected by various organic linkers (mainly polycarboxylates). By self-assembling of the starting materials (metal salts and organic molecules), a combination of the inorganic secondary building units (SBU) with organic linkers leads to a variety of 1-, 2- and 3-dimensional periodic structures.

For porous materials to be used and optimized for a particular application, a detailed knowledge of atomic structure and lattice dynamics is essential. To understand diffusion and adsorption within the porous material, a careful study of the nature of host−guest interactions is needed. Molecular simulations are indispensable in the rational design of the properties of these materials. But due to their size and complexity, one has to resort to force field (FF) methods. Previous experience accumulated from the zeolite modeling using empirical FF has proven to be of paramount importance to achieve this ambitious goal. For MOFs it can be expected that theoretical calculations will be of similar relevance, especially because of the fact that a systematic variation of the organic linkers opens the possibility for a true rational design of various properties. In the case of zeolites accurate force field parametrizations are nowadays available.[5] Due to the large variability of inorganic and organic building blocks in MOFs,[6−8] however, there is a need for a much larger FF parameter set. For the organic linker part this is not so much of a problem, because very reliable parametrizations exist, but for the inorganic building blocks, parameters are mostly not available. Moreover, because experimental data on the structures and vibrational frequencies are very scarce for this class of compounds, ab initio calculations should be used for the FF parametrization. To be predictive, theoretical methods should even be able to describe structures not yet synthesized.

A number of previous theoretical studies of MOF systems have used generic FFs like UFF[9] or DREIDING,[10] where the force field are derived by a number of established rules from atomic parameters. Unfortunately, these methods are usually not very accurate and large deviations from experimental structures are sometimes observed. Thus, in the simulations, the framework is usually kept frozen at the experimentally determined structure. For the rather rigid zeolite materials this kind of approximation is often justified, but a number of cases where lattice flexibility is relevant are known.[5] MOFs are in contrast much more flexible and a strong influence of the lattice dynamics on the cell parameters[11] has been observed. Many physical properties like elastic constants or thermal conductivity depend intrinsically on the lattice dynamics.[12] Furthermore, certain frameworks can actively respond to the inclusion or removing of guest molecules thus demonstrating guest-dependent dynamic behavior where the host framework can reversibly shrink upon inclusion of organic guest molecules and expand upon guest removal, the so-called "breathing" effect,[8,13−15] which could be of relevance for applications like sensing devices. In addition, to investigate interfaces, isomeric structures of networks or various defects therein, which cannot (or only with great difficulty) be observed by experimental methods, a force field describing accurately deformed nonequilibrium structures is needed. It must be noted that for all the above-mentioned properties and effects no bonds are broken or made. This allows the use of traditional "nonreactive" FFs as opposed to approaches that allow for bond breaking.[16,17] The latter are either much more complex in the definition of the FF or more simplistic but less accurate in the representation of the potential energy surface.[18,19] Thus, we focus on "bonded" or "covalent" FFs. However, our parametrization strategy itself can be applied for other kinds of FFs as well.

* Corresponding author. E-mail: rochus.schmid@rub.de.

10.1021/jp807487f CCC: $40.75
Published on Web 01/09/2009

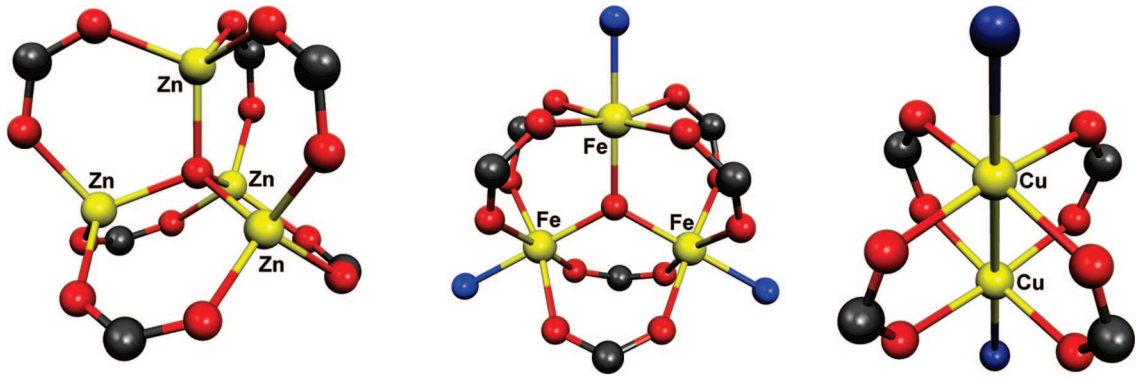

Figure 1. Some representative examples of inorganic secondary building units used in MOF materials: tetranuclear (left) and trinuclear (middle) oxo-centered metal clusters and paddle-wheel-like dinuclear metal clusters (right). Color code: metal, carbon, oxygen and donor atoms are shown in yellow, gray, red and blue, respectively.

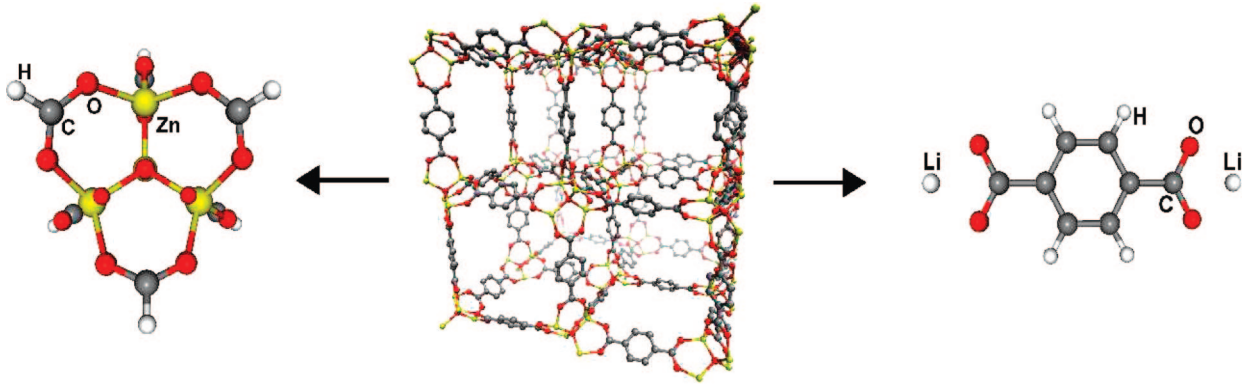

Figure 2. System reduction for the FF parametrization. Middle: structure of MOF-5 with eight cubic pores (hydrogen omitted). Left: structure of basic zinc formate. Right: structure of dilithium terephthalate.

Deriving accurate FF parameters for coordination complexes is still a challenge.[20] In our recent work on the FF parametrization for a particular member of MOFs (MOF-5) encouraging results concerning structures and infrared spectra were obtained.[21] Good agreement with experiment was also found on the benzene diffusion in MOF-5.[22] Most of the previous approaches describe the metal−ligand interactions in MOFs by taking only van der Waals and electrostatic interactions into account (only two-body terms), ignoring directionality of bonding. The van der Waals parameters and effective point charges were usually adjusted to match experimental data (such as adsorption isotherms),[11] or using first principles calculations,[23] or even transferred from other systems.[18] Recently, Huang et al.[12] have adjusted the FF parameters based on the reproduction of relative energies of distorted structures of MOF-5 calculated with DFT. There, in addition to two-body (nonbonded) interactions some three- and four-body potentials were also included in their force field, whereas the phenylene rings were treated using a united atom approach. We also mention in passing that some previous attempts exist to parametrize the Zn−ligand interactions as covalent bonds.[24,25] Also, using the analytical bond-order formalism, an interatomic potential for zinc oxide is derived and validated by reproducing bulk properties of various solid structures of zinc oxide including cohesive energies, lattice parameters, and elastic constants.[26] More recently, a force field for the MIL-53(Cr) framework, treating the Cr−O interactions as bonded, is derived and validated by molecular dynamics simulations, which allows reproduction of the structural transformations ("breathing") of the framework upon thermal activation and CO_2 adsorption.[27]

The objective of the present work is to develop a strategy to derive FF parameters for simulations of various hybrid materials, especially for their inorganic part, in a consistent way automatically (with minimal user intervention) on the basis of high level ab initio calculations of well chosen model systems. Our overall target is to develop a consistent force field for a variety of MOF materials, based on well-defined motifs for inorganic SBUs such as those shown in Figure 1, with a specialized parametrization for the respective chemical environment. Thus, we fully rely on the first principles reference data and thus seek FF parameters, which reproduce as close as possible the DFT results (structure and vibrational modes). As a consequence we do intentionally not correct for any deficiencies of the first principles method as, for example, a systematic overestimation of bond lengths of transition metal−ligand bonds in hybrid DFT.[28] As a proof of concept, and as an extension to our prior work,[21] we demonstrate the procedure for the Zn_4O based IRMOF series, using a slightly modified form of the well-established MM3 force field.[29]

2. Methodology

2.1. Ab Initio Calculations of the Model Systems. Our model systems are chosen as follows: the 3-dimensional periodic structure of MOF-5 can be represented as being composed of a zinc−oxide cluster as the inorganic SBU, $Zn_4O(O_2CR)_6$, and a terephthalate anion as the organic linker, as depicted schematically in Figure 2. The zinc−oxide cluster capped with hydrogen atoms (formally known as basic zinc formate) serves as our model system for the inorganic SBU, whereas the linker is modeled by the terephthalate anion with two lithium atoms as counterions. Thus, as an extension to our previous work,[21] where we used the phenyl capped zinc−oxide cluster (basic zinc benzoate) as a single reference system, we show here that the "building block" approach can also be used for the force field

parametrization. By using the $Zn_4O(O_2CH)_6$ as the model system for the parametrization of the zinc−oxide cluster within MOF-5, we rely on the assumption that the parameters will be transferable to a larger and periodic system. Note that electronic structure calculations of the periodic systems showed that the electronic communication between the inorganic SBUs is basically negligible, and MOF-5 can faithfully be seen as a 3D polymer of the electronically unchanged monomeric building blocks.[30] Lithium ions have been used previously to model the environment for the terephthalate in MOF-5.[31,32] The lithium carboxylate group $-CO_2Li$ is a good choice to represent a carboxylate bonded to a cation like Zn^{2+} (Li^+ has a very similar ionic radius), superior to the free carboxylate $-CO_2^-$ or the acid form $-CO_2H$, where the bonding situation within the carboxylate group is different from that observed in the zinc−oxide cluster. Note that the parameters for the carboxylate group itself in the final force field are taken from the zinc−oxide cluster model.

The model systems were fully optimized by DFT calculations using the hybrid functional B3LYP,[33−35] which was found to be reliable for zinc coordination compounds.[36] In contrast to our previous study, we used here the correlation consistent basis sets for all atoms employing energy-consistent pseudopotential for the Zn atom denoted as cc-pVXZ-PP (X = D, T), because this type of basis set has the advantage of systematically improving the results of ab initio calculations (i.e., errors introduced by an incomplete basis set can be controlled by using a more extended one in this series).[37−39] Additional calculations were performed with an all-electron triple-ζ correlation consistent basis set, denoted as cc-pVTZ-DK.[40] All calculations were carried out with the Gaussian program package.[41] Due to the presence of low vibrational modes (on the order of 10 cm^{-1}), we used, in all DFT calculations, a tighter convergence criterion (Opt = Tight) and a finer grid for the integration (Int = Ultrafine). The optimized structures for the $Zn_4O(O_2CR)_6$ (R = H or Ph) model systems of T_d symmetry were confirmed to be true minima by vibrational frequency calculations.The effective atomic charges were obtained from the fitting of the electrostatic potential based on the Merz−Kollman sampling scheme[42] using ca. 2000 grid points per atom. The van der Waals radius of 2.29 Å for the Zn atom, as recommended in ref 43, was adopted here as in our previous study.

2.2. Force Field. The MM3(2000) force field of Allinger et al.[29] was used as implemented in the TINKER program package.[44] Within our covalent model, the metal−oxygen interactions were treated with the usual bonded terms. Four new atomic types have been introduced: zinc, central oxygen (O_{cent}), carboxylate oxygen (O_{carb}) and carbon (C_{carb}) atoms. For the phenylene moiety, the original MM3 atom types 2 and 5 were assigned to the carbon (C_{ph}) and hydrogen atoms, respectively. As in our previous work, the conjugation in the phenylene moiety is taken into account by adjusting the values for the $C_{ph}-C_{ph}$ bond stretching force constant and the corresponding reference bond length together with the 2-fold torsional parameters for the rotation around $C_{ph}-C_{ph}$ bonds by performing a self-consistent field molecular orbital calculation for the π-system using a modified Pariser−Parr−Pople method as implemented in TINKER.[45] In principle, these values can be also optimized but we decided to keep them unaltered for the present study to have the number of FF parameters, subjected to a fitting procedure, as small as possible. A reliable description of the local hypersurface around a potential energy minimum, and through this a good description of the molecular vibrations, require a force field with a certain number of off-diagonal (or cross) terms of the Hessian matrix.[29,46] To this end, we have modified the original MM3 stretch−bend cross-term (we introduced two different force constants, k_{sb}^a and k_{sb}^b, for two bonds that comprise the same bond angle) and implemented a new stretch−stretch cross-term as suggested by Mapple et al.[46] Their functional forms are given below, where r_{ref} and α_{ref} are the reference values for the bond length and bond angle, respectively.

for stretch−stretch cross-terms:

$$E_{ss} = k_{ss}(r^a - r_{ref}^a)(r^b - r_{ref}^b) \quad (1)$$

and for stretch−bend cross-terms:

$$E_{sb} = k_{sb}^a(r^a - r_{ref}^a)(\alpha - \alpha_{ref}) + k_{sb}^b(r^b - r_{ref}^b)(\alpha - \alpha_{ref}) \quad (2)$$

The importance of the stretch−stretch cross-terms, especially for conjugated systems, has been recognized and implemented in the MM4 force field.[47]

As in our previous study, the van der Waals parameters for the Zn atom were taken from ref 43. For the newly introduced carbon and oxygen atom types the same values as in the original MM3 force field were used. The effective atomic charges were fitted to reproduce the electrostatic potential. In contrast to our previous work, the long-range Coulomb interactions were described directly via atomic charge−charge interactions. This is done to compare our results with those known from the literature. We found that describing electrostatics via the bond centered dipoles, calculated from effective atomic charges as in our previous work, produced similar results. Because the MM3 force field utilizes dipole−dipole electrostatic interactions via bond dipoles, we have scaled down the charge−charge interactions between 1−4 connected atoms (atoms separated by three covalent bonds) for our covalent model. The multiplicative scale factor of 0.48 was chosen so that the contribution from electrostatic interactions to the strain energy, described either by dipole−dipole or by charge−charge interactions, is of approximately the same value. We note that different molecular force fields treat the 1−4 Coulomb interactions differently and, for example, a scale factor of 0.75 is used in the MMFF94 force field,[48] whereas a value of 0.5 is adopted in the OPLS-AA force field.[49]

2.3. Genetic Algorithm. There have been a number of approaches to parametrize a force field directly from the quantum chemical calculations.[50] Because emphasis is placed on obtaining both accurate geometries and accurate vibrational frequencies within the molecular mechanics (MM) method, the Hessian matrix (second derivatives of the energy with respect to atomic coordinates) is needed. A simple and straightforward strategy was developed by Leonard and Ashman,[51] who converted the diagonal elements of the Hessian matrix, calculated at the equilibrium geometry and expressed in *Z*-matrix (internal) coordinates for small model systems, to the corresponding bond stretching and angle bending force constants to be used with the MM2 force field. This procedure, however, can result in unbalanced FF parameters if some kind of redundancy occurs between internal coordinates. Dasgupta and Goddard[52] combined the Hessian from ab initio calculations with the structural and spectroscopic data from experiment to generate a new (biased) Hessian that is then used to derive an accurate FF. This method is not applicable in cases where experimental data are not available or scarce. The Hessian matrix in Cartesian

coordinates was utilized in another method, developed by Seminario,[53] to extract the force constants based on the analysis of pair atomic interactions in a molecular system of any complexity. Here, only local information stored in the Hessian matrix is used. Moreover, this procedure becomes cumbersome if some off-diagonal terms need to be extracted. Palmo et al.[54] have proposed a direct mathematical transformation of the ab initio structure and Hessian matrix to produce a completely general valence FF. However, a drawback of this latter direct procedure is that a large number of force constants is produced initially, which has to be reduced afterward in some way to be of practical use.

Besides the above-mentioned more or less direct conversions of the Hessian matrix, a number of approaches have been developed to fit the (relative) energies and/or the first and second derivatives of the energy (for a given MM analytical function) to the corresponding values obtained from ab initio calculations on a small number of distorted configurations.[46] Such fitting can in principle be accomplished by a systematic search or by a number of stochastic approaches, based on a carefully chosen fitness criterion or objective function.[55] Among the latter, in particular the so-called evolutionary algorithms have found widespread use in various optimization problems. Genetic algorithms, being the most popular class of evolutionary algorithms, are heuristic search techniques that incorporate in a computational setting the biological notion of evolution by means of natural selection using three basic operations of selection, crossover, and mutation, operating on "genotypes" encoded as strings.[56] This type of global optimizers is particularly well suited for nonlinear multiple-optima (minima or maxima) optimization problems of high dimensionality of the search landscape with possibly strongly correlated adjustable parameters. GAs have been used previously for force field parametrization.[55,57−63] In the present work we adopted a particular implementation of GAs known as PIKAIA and widely used in astronomy and astrophysics.[64] This particular GA implementation has already been used in FF development for molecular dynamics simulations of bulk silica materials.[65] Some particular features of this implementation include (1) two-point crossover, (2) creep mutation, and (3) the mutation rate, which can be dynamically controlled by monitoring the difference either in fitness between the current best and median in the population or in metric distance in parameter space. It uses decimal (rather than binary) encoding and establishes selection probability in terms of fitness-based ranks, making use of the Roulette Wheel Algorithm. Three reproduction plans are available: full generational replacement, steady-state-delete-random, and steady-state-delete-worst. Elitism is incorporated as a default option. Further details can be found in ref 66.

2.4. Objective Function. In most of the work on GAs for the parametrization of FFs from first principles calculated data, objective functions based on structural deviations and partly on relative energies (often of a set of reference systems) have been used, ignoring any curvature information of the reference potential energy surface.[57−59] To include curvature information, one could in principle utilize deviations of the raw Hessian matrix in Cartesian coordinates in the objective function. However, this means considering a large number of matrix elements with the relevant information buried under many small off-diagonal values. It is, thus, a common practice to include vibrational frequencies in the objective function.[55,63,67] For large systems, however, some internal degrees of freedom can be strongly coupled and the normal mode assignment becomes an important issue. To resolve this problem, we use instead a complete (redundant) internal coordinate system and a direct transformation of the Cartesian Hessian matrix into this set of internal coordinates. Besides substantially reducing the size of the problem, this avoids any ambiguity in the assignment of the modes, because we use the projected Hessian matrix elements directly in the fitness function. As will be shown below, a very good correspondence in terms of both frequencies and normal modes can be achieved, provided a reasonable number of Hessian matrix elements, both diagonal and off-diagonal, reproduce their reference counterparts. Our approach is in contrast to that used recently by Mostaghim et al.,[60] where an advanced multi-objective evolutionary algorithm was applied for the FF parametrization, based on a fitness function composed of the diagonal elements of the Hessian matrix, projected into a nonredundant set of internal coordinates. In our approach, we found it to be sufficient to stick to the single-objective paradigm. However, instead of the minimal nonredundant internal coordinate system we use a much more natural redundant coordinate system. The projection of both structural and curvature (Hessian matrix) data into the redundant internal coordinate system, defined by the systems connectivity, is motivated by the underlying principles behind the molecular mechanics concept. From a practical point of view this is advantageous, because the model systems can have quite complex topologies (vide infra) and hence the selection of a particular nonredundant set is sometimes not trivial, whereas the redundant set can be defined fully automatically using the bond topology information. More importantly, the point group symmetry of the structure and corresponding force field is preserved; i.e., all symmetrically equivalent bond stretchings or angle bendings in a molecule have the same force constants. In contrast to some previous studies,[46,60,65] where quantum-mechanically calculated forces (first derivatives of the energy) at nonstationary points were included in the objective function, we utilize the internal coordinates (bond distances, bond angles and, optionally, out-of-plane bends and torsions) at the energy minimum.

We project the Hessian matrix in Cartesian coordinates (F_X) into the complete (redundant) set of primitive internal coordinates (F_R) using a generalized inverse technique.[68] Using the following relationship (in matrix notation):

$$R = BX \tag{3}$$

where B is a well-known Wilson's rectangular matrix ($3N \times M$; usually $M > 3N$),[69] which transforms the rectilinear Cartesian coordinates (X) of an N-atomic system to a (redundant) set of M curvilinear internal coordinates (R), this transformation reads as (at stationary points):

$$F_R = (B^T)^{-1} F_X B^{-1} \tag{4}$$

where B^{-1} is a generalized inverse of a B matrix (superscript T denotes transposition). This transformation is unique for a given set of internal coordinates, the only requirement being that the B matrix should have full rank (=min ($3N$,M)).The translational and rotational degrees of freedom are completely projected out.

The novel objective function is the inverse of the mean-square-deviation (MSD) χ^2 of the metric parameters (bond distances r, valence angles α, torsional angles ϕ and out-of-plane bendings θ) and the elements F of the projected Hessian

matrix (both diagonal and off-diagonal) between ab initio and MM optimized structures for the model system:

$$\chi^2 = \frac{w_{\text{bonds}}}{N_{\text{ref}}^{\text{bonds}}} \sum_{i=1}^{N_{\text{ref}}^{\text{bonds}}} (r_i^{\text{MM}} - r_i^{\text{ref}})^2 + \frac{w_{\text{bends}}}{N_{\text{ref}}^{\text{bends}}} \sum_{j=1}^{N_{\text{ref}}^{\text{bends}}} (\alpha_j^{\text{MM}} - \alpha_j^{\text{ref}})^2 + \frac{w_{\text{wags}}}{N_{\text{ref}}^{\text{wags}}} \sum_{k=1}^{N_{\text{ref}}^{\text{wags}}} (\theta_k^{\text{MM}} - \theta_k^{\text{ref}})^2 + \frac{w_{\text{tors}}}{N_{\text{ref}}^{\text{tors}}} \sum_{l=1}^{N_{\text{ref}}^{\text{tors}}} (\phi_l^{\text{MM}} - \phi_l^{\text{ref}})^2 + \frac{w_{\text{diag}}}{N_{\text{ref}}^{\text{diag}}} \sum_{m=1}^{N_{\text{ref}}^{\text{diag}}} (F_m^{\text{MM}} - F_m^{\text{ref}})^2 + \frac{w_{\text{cross}}}{N_{\text{ref}}^{\text{cross}}} \sum_{n=1}^{N_{\text{ref}}^{\text{cross}}} (F_n^{\text{MM}} - F_n^{\text{ref}})^2 \quad (5)$$

where N_{ref} gives the number of the corresponding reference data. Bond distances are expressed in Å, all angle terms are in radians and the force constants in mdyn/Å (bonds) or (mdyn Å)/rad^2 (angles), so that all adjustable FF parameters, both metric and force constants, are of the same order of magnitude. In principle, the weighting factors w could be adjusted for each term. However, for this choice of units we found it sufficient to use unity for all weighting factors.

2.5. Implementation. We have interfaced a serial version of PIKAIA with the TINKER package.The whole procedure for the FF parametrization is shown schematically in the flowchart in Figure 3.

The evaluation of the fitness of an individual parameter set comprises several basic steps:

(1) MM geometry optimization of the reference structure based on the selected functional form (here MM3) and a trial set of parameters;

(2) calculation of the Hessian matrix in Cartesian coordinates;

(3) transformation of both the Cartesian coordinates and the Hessian matrix into the principal axes system;

(4) transformation of the Hessian matrix from Cartesian to a predefined set of redundant internal coordinates;

(5) calculation of the sum of mean square deviations between a trial solution and the reference data.

Any redundant set of internal coordinates is possible, which must not necessarily be based on the connectivity actually used in the MM calculation. This means that both "bonded" and "nonbonded" parametrizations can be obtained using the same reference data. The MM strain energy is minimized using the optimally conditioned variable metric nonlinear optimization routine without line searches, as implemented in TINKER, with a convergence criterion of root-mean-square of 0.01 kcal/(mol Å) per atom in the gradient.The Hessian matrix is calculated numerically from analytical gradients with the stepsize of 10^{-5} Å.

The stopping criteria for the GA can be set up by separate convergence thresholds for the metric parameters (bonds, angles) and Hessian terms or by a fixed number of generations to be evolved. After extensive testing, we found the following GA control parameters to be well suited for rapid convergence:

(a) A steady-state reproduction plan (i.e., insertion individuals as they are being bred), with replacing the least fit member of the parent population by the offspring, was used.

(b) A two-point crossover operation (i.e., cutting between two randomly chosen points and exchanging a segment in the genome) was performed with a probability of 85%.

(c) A uniform one-point (jump) and creep mutation with equal probabilities was used with the fixed rate (probability) of 5%.

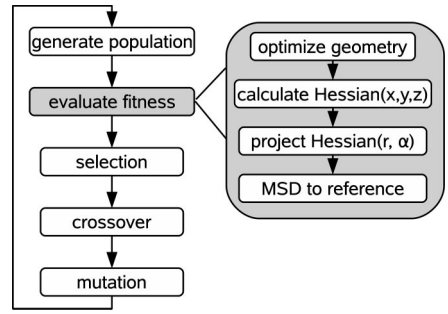

Figure 3. Flowchart for FF parametrization.

TABLE 1: Selected Bond Distances (in Å) for Model Systems $Zn_4O(O_2CR)_6$

	Zn−O_{cent}	Zn−O_{carb}	C_{carb}−O_{carb}
	B3LYP/cc-pVDZ-PP		
R = H	1.974	1.973	1.262
R = Ph	1.967	1.970	1.272
	B3LYP/cc-pVTZ-PP		
R = H	1.980	1.971	1.256
R = Ph	1.970	1.967	1.265
	B3LYP/cc-pVTZ-DK		
R = H	1.986	1.971	1.255
R = Ph	1.977	1.965	1.264
	MP2/cc-pVDZ-PP		
R = H	1.956	1.965	1.266
	RI-MP2/TZVPP[73]		
R = H	1.973	1.965	1.257
	X-ray[74]		
R = Ph	1.936−1.950	1.930−1.950	1.234−1.270

(d) elitism is used, i.e., the fittest member of the breeding population is recopied to the new population as the final step of the reproduction plan.

The ranges for various adjustable parameters are defined as follows: ±0.1 Å for bond lengths, ±10° for bond angles, ±1 mdyn/Å for bond stretchings, [0.0, 2.0] (mdyn Å)/rad^2 for bond bendings, [−10.0, 10.0] kcal/mol for torsional potential terms, [0.0, 3.0] (mdyn Å)/rad^2 for out-of-plane bends and [−3.0, 3.0] (in respective units) for all cross-terms. During GA runs, the calculated Hessian matrix is checked for negative eigenvalues to ensure the optimized structure to be always a true minimum.We select the best members (with the highest fitness) of the population of size 200 from 10 GA runs and then started the final GA run letting them evolve with the rest of the population chosen randomly. Our final results are those represented by the best member of the population from this last GA run.

3. Results and Discussion

3.1. Reference Systems. The most important structural parameters of the inorganic reference system, calculated with different basis sets, are compared in Table 1. Increasing the quality of the basis set from double- to triple-ζ causes slight variations in the Zn−O bond distances with a clear trend of elongating the Zn−O_{cent} bond and shortening the Zn−O_{carb} bond. It should also be noted that replacing inner electrons of the Zn atoms with the effective pseudopotential (cc-pVTZ-PP) causes only minor changes in the Zn−O bond distances as compared to the all-electron results (cc-pVTZ-DK). Overall, however, the variations are smaller than the systematic overestimation of all Zn−O bond lengths using the B3LYP approximation with respect to the experimental

TABLE 2: Selected Stretching Frequencies (in cm^{-1}) for the Core Vibrations of $Zn_4O(O_2CH)_6$

method	$\nu_s(Zn-O_{cent})$	$\nu_{as}(Zn-O_{cent})$	$\nu_s(C_{carb}-O_{carb})$	$\nu_{as}(C_{carb}-O_{carb})$
B3LYP/cc-pVDZ-PP	226	522	1395, 1397, 1404	1622, 1698
B3LYP/cc-pVTZ-PP	219	507	1397, 1398, 1405	1602, 1676
B3LYP/cc-pVTZ-DK	220	507	1401, 1402, 1410	1607, 1679
FF (without cross terms)	199	596	1342, 1342, 1343	1925, 1926
FF (with cross terms)	220	533	1407, 1407, 1409	1674, 1676

TABLE 3: Atomic Charges Used in the FF Parametrization

atom	charge (e)
O_{cent}	−1.41
Zn	1.30
O_{carb}	−0.68
C_{carb}	0.66
$C_{ph}(-C_{carb})$	0.02
C_{ph}(av)	−0.12
H(av)	0.12

data.We note that geometry optimization at the MP2/cc-pVDZ-PP level gives the Zn−O bond lengths much closer to the experimental values, however, the necessary frequency calculation is currently numerically too demanding (unless some approximations developed recently, like RI-MP2, are used; see, for example refs 70−72). Thus, the cc-pVDZ-PP basis set results have been used as a reference for the force field parametrization for this zinc−oxide cluster. Note in this context that we intentionally fit the force field purely to the theoretical reference data and do not correct for these systematic deviations from experimental results.

From the data shown in Table 1 it is also obvious that the substitution of R = Ph by R = H leads, for example, to a systematic lengthening of the $Zn-O_{cent}$ bond, but this effect is of about the same magnitude as the deviation between experimental and theoretical bond length. Therefore, the smaller system with R = H can as well be used for the FF parametrization.

The influence of the quality of the basis set on some characteristic vibrational frequencies for the core of basic zinc formate is only marginal as can be judged from the data presented in Table 2.

The reference data for dilithium terephthalate were generated from the optimized structure (D_{2h} symmetry) and Hessian at the same level of theory (B3LYP/cc-pVDZ). To estimate the barrier to internal rotation of the phenylene moiety, a "perpendicular" conformer (D_{2h} symmetry), constrained to have both carboxylate groups rotated by 90° with respect to the phenylene ring, was optimized. The torsional barrier was calculated as the energy difference between the global minimum (all atoms are in one plane) and the perpendicular conformer (16.1 kcal/mol).

Effective atomic charges are listed in Table 3. They are only slightly different from those used in our previous work due to the change in the basis set.

3.2. Parameterization. The reference data were generated from the optimized structures and Hessians of our model systems (basic zinc formate and dilithium terephthalate) based on the B3LYP/cc-pVDZ-PP calculations. With the reference data in hand, we adjusted the parameters for two- (bonds), three- (angles) and four-body (torsions and out-of-plane wags) terms assuming that the Zn atom is covalently bonded to the O_{cent} and O_{carb} atoms in the zinc−oxide cluster. During the refinement, we keep the number of FF parameters as small as possible, and a total of 25 parameters were adjusted simultaneously with a population size of 200 members in the GA runs. Our preliminary refinements indicated that the torsional parameters around all bonds containing the Zn atom can safely be set to zero. However, for example the torsions around $Zn-O_{carb}-C_{carb}-O_{carb}$ are explicitly included into the parametrization.These terms mimic the amount of directionality of the coordinative $Zn-O_{carb}$ bonds, keeping the Zn atom in the carboxylic plain. Recently, the necessity for such terms was highlighted in the parametrization of a force field for the "breathing" MIL-53 system.[8,27,75] We note in passing that the same reference data for basic zinc formate can be used also to parametrize a corresponding "nonbonded" force field model,[11,76] where only two-body electrostatic and van der Waals interactions are used to describe all Zn−O interactions and such directionality is not considered.

As stated above, during GA runs all reference data (metric parameters and Hessian terms) are treated with the same weighting factors (1.0) and the adjustable FF parameters (reference distances, angles and corresponding force constants) show rapid convergence from a completely random initialization (within the ranges mentioned above), as can be judged from a semilogarithmic plot in Figure 4. One GA run (population 200, 500 generations) took about 7−8 h on a standard LINUX PC. Because the multiple GA runs can be performed in parallel, the total time for optimizing the force field is comparable to the overall time needed to calculate the DFT reference data (optimization and frequency calculation) on a comparable parallel computer. Realistic MD simulations of the periodic network can run up to several days, and the time needed for parametrizing a new system is thus faster and can in principle be completely automated.

As it has been found before, to reproduce some asymmetric stretchings accurately, certain cross-terms must be included in the parametrization.We found that adjusting stretch−bend and stretch−stretch terms describing the $Zn-O_{cent}/Zn-O_{cent}-Zn$ and $Zn-O_{cent}/Zn-O_{cent}$ interactions and especially the $C_{carb}-O_{carb}/O_{carb}-C_{carb}-O_{carb}$ and $C_{carb}-O_{carb}/C_{carb}-O_{carb}$ interactions can significantly improve the agreement between the MM and DFT (B3LYP/cc-pVDZ-PP) stretching frequencies as shown in Table 2 for the basic zinc formate. This effect is most pronounced for the asymmetric bands.

Even though the vibrational frequencies were not included in the fitness function, the optimized FF parameters are able to reproduce both the normal modes (eigenvectors) and their respective frequencies as can be seen from the data presented in Table 4. The frequencies were ranked in accordance with the maximum overlap of the MM eigenvectors with the reference modes (DFT), taken as the cross product of the eigenvectors for two sets of modes. The calculated cross products were all very close to unity, indicating a good match. This is remarkable taking into account that only a very limited number of off-diagonal elements of the Hessian matrix was used in the fit.

With the final FF parameters, the geometrical parameters for basic zinc formate are fit to DFT results to within the

TABLE 4: Low Vibrational Modes (in cm^{-1}) for Basic Zinc Formate (<600 cm^{-1})

B3LYP/ cc-pVDZ-PP	force field	deviation[a]	mode degeneracy
36.34	45.02	−8.68 (23.9)	3
69.77	86.63	−16.86 (24.2)	1
73.48	82.90	−9.42 (12.8)	2
83.97	92.67	−8.69 (10.3)	3
122.06	134.22	−12.16 (10.0)	3
130.60	128.63	1.97 (1.5)	3
131.49	134.22	−2.74 (2.1)	3
146.68	136.75	9.92 (6.8)	3
157.18	142.74	14.44(9.2)	2
225.93	220.32	5.61 (2.5)	1
270.26	273.25	−2.99 (1.1)	3
293.38	288.35	5.03 (1.7)	3
377.19	363.88	13.30 (3.5)	3
386.05	365.34	20.71 (5.4)	1
392.39	433.71	−41.32 (10.5)	3
393.54	396.95	−3.41 (0.9)	2
401.06	422.02	−20.96 (5.2)	3
522.49	532.90	−10.41 (2.0)	3

[a] Relative deviation in percent is given in parentheses.

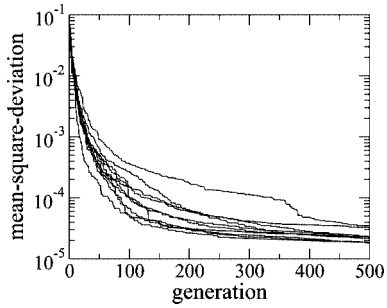

Figure 4. Sample convergence curves for 10 GA runs for the basic zinc formate model system. The variance (χ^2) of the best trial in the population as a function of generation is shown. Number of adjustable FF parameters: 25. Population number: 200.

following ranges of mean absolute deviation (maximum absolute deviation in parentheses): 0.002 (0.006) Å for bond lengths, 0.08 (0.2) deg for bond angles, 0.002 (0.005) for diagonal Hessian and 0.0006 (0.001) for off-diagonal Hessian terms in respective units. In Table 4 the deviations of the lowest normal modes (below 600 cm^{-1}) with respect to the DFT reference are listed in detail. The root-mean-square (rms) deviation for these modes is only 15 cm^{-1}. The remaining larger errors are mainly due to the still rather small number of cross terms included in the parametrization.

To parametrize the linker, the reference data for dilithium terephthalate were used. The main objective here is to supplement the MM3 force field with the parameters for the benzene ring attached to the carboxylate group. Two lithium atoms served to mimic the electronic environment of the carboxylate group. Therefore, the parameters of the following terms were adjusted by the GA: $C_{carb}-O_{carb}$ and $C_{carb}-C_{ph}$ bonds, $O_{carb}-C_{carb}-O_{carb}$, $O_{carb}-C_{carb}-C_{ph}$ and $C_{carb}-C_{ph}-C_{ph}$ angles, and the respective out-of-plane bends for a trigonal C_{carb} center. In addition, stretch−stretch and stretch−bend cross terms around the trigonal C_{carb} center were included in the optimization. The stretch−stretch and stretch−bend terms for carbon atoms of the benzene ring represented with the atom type 2 (with a special treatment for conjugation, see above) were adjusted as well with the final values similar to those suggested by Nevins et al.[47] To describe the torsional potential of the phenylene group, we used a single cosine term of the form $V(\varphi) = 1/2V_0(1 - \cos 2\varphi)$, where φ is the $C_{ph}-C_{ph}-C_{carb}-O_{carb}$ dihedral angle and the parameter V_0 was optimized to reproduce the respective torsional force constant (dilithium terephthalate). Overall, 24 parameters were refined with the GA against the respective reference data.With the value for the torsional parameter, V_0, we can reproduce the barrier for the internal rotation of the phenylene moiety as obtained by our DFT calculations (DFT, 16.1 kcal/mol; MM, 15.8 kcal/mol), which indicates that our single cosine term is adequate for describing this torsional motion.The torsional barrier is slightly overestimated by our calculations as compared with a recently determined experimental value of 11 ± 2 kcal/mol.[77]

The final parameters from our two model systems (basic zinc formate and dilithium terephthalate) were combined without readjustment. The parameters for the carboxylate group ($C_{carb}-O_{carb}$ bond stretching, $O_{carb}-C_{carb}-O_{carb}$ angle bending, $C_{carb}-O_{carb}/C_{carb}-O_{carb}$ stretch−stretch and $C_{carb}-O_{carb}/O_{carb}-C_{carb}-O_{carb}$ stretch−bend), adjusted independently for both model systems, were found to be very similar. Thus, the final values were taken from the parametrization of basic zinc formate, because this cluster resembles more closely the chemical environment of the carboxylate group in MOF-5. The final parameters are listed in Table 9 and were used in all molecular dynamics simulations.

3.3. Validation of the Force Field. To judge the validity of our "building block" strategy, and thus the transferability of our FF parameters, the geometry and vibrational frequencies for the basic zinc benzoate (see Figure 5) were compared with the results from DFT calculations carried out at the same level of theory.

As can be seen from the data listed in Table 5, the most important structural parameters are in excellent agreement.

The vibrational frequencies (see Table 6) calculated using the force field match their respective DFT counterparts quite well with the rms deviation for the low frequency modes (<600 cm^{-1}) being only 13 cm^{-1}.

As compared to our previous force field for basic zinc benzoate,[21] a significant improvement is achieved especially for the Hessian terms and vibrational frequencies, as judged from the matching to our DFT results. Using our previous force field, the rms deviation for the low frequency modes (<600 cm^{-1}) is much larger (50 cm^{-1}). This discrepancy is mainly due to the fact that most of the torsional parameters in the zinc−oxide cluster were not optimized.

To validate our force field, we carried out a series of static minimizations and molecular dynamics simulations of the periodic MOF-5 structure, using a cutoff value of 12 Å for the van der Waals interactions, and the electrostatic interactions were calculated by a smooth particle mesh Ewald summation with a cutoff value of 12 Å. Our preliminary tests with a somewhat larger cutoff radius gave similar results.

The full crystal energy minimization of MOF-5 was performed by optimizing over atomic coordinates of the cubic unit cell (424 atoms, space group *P*1) and the lattice constant with the rms gradient at convergence of 0.01 kcal/(mol Å). The optimized lattice size of 26.04 Å (which corresponds to $T = 0$ K) is in good agreement with the value of 25.894 Å[78] for evacuated MOF-5 measured at low temperatures ($T =$ 30 K) and 25.909 Å ($T = 3.5$ K).[79] The higher simulated value for the lattice parameter is mainly due to the overestimation (∼0.03 Å) of the Zn−O distances in our DFT reference (∼1.97 Å) as compared to the experimental value (∼1.94 Å).[78] However, the value is very close to the calculated value of 26.09 Å obtained recently for the fully

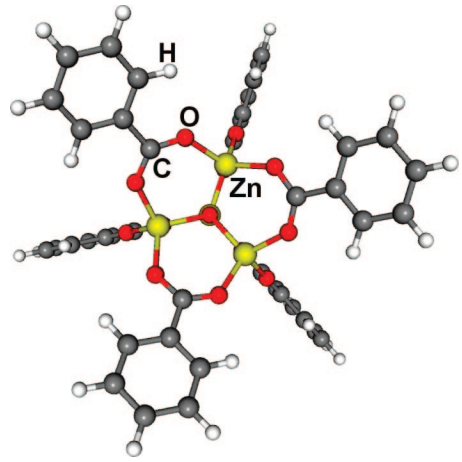

Figure 5. Structure of basic zinc benzoate.

TABLE 5: Selected Distances and Angles for Basic Zinc Benzoate

	force field	B3LYP/cc-pVDZ-PP
	Bond Distances (Å)	
$Zn-O_{cent}$	1.961	1.967
$Zn-O_{carb}$	1.975	1.970
$C_{carb}-O_{carb}$	1.264	1.272
$C_{carb}-C_{ph}$	1.510	1.496
	Bond Angles (deg)	
$O_{cent}-Zn-O_{carb}$	111.4	111.3
$Zn-O_{carb}-C_{carb}$	130.6	131.2
$O_{carb}-C_{carb}-O_{carb}$	126.5	125.5

optimized periodic structure of MOF-5 with the same DFT method (B3LYP) employing all-electron basis sets[80] and to those calculated within the framework of a different generalized gradient approximation to DFT by using the projector augmented-wave method and a plane-wave basis set (26.06 Å[81] and 26.04 Å[82]).

The normal coordinate analysis for MOF-5 based on our force field showed very good correspondence with the DFT results of Civalleri et al.,[80] as can be judged from Table 7 where some characteristic normal modes in both low- and high-frequency regions are listed. We note that our simple effective point charge model gives the lowest IR active mode at 55 cm^{-1} assigned to an in-phase translation of the phenylene moiety toward the center of the face of the cubic cell of MOF-5 ("breathing" mode), in agreement with the periodic DFT calculations (at 52 cm^{-1}).[80] Moreover, our present force field is able to reproduce the overall pattern of IR active modes calculated recently.[80]

In Figure 6 the simulated infrared spectrum is shown in comparison to an experimental spectrum of MOF-5. The IR intensities have been obtained from the dipole moment time derivative autocorrelation function. The dipole moment derivatives were calculated from the atomic velocities saved every 1 fs, weighted with the fixed point charges. The system was subjected to 100 ps of temperature equilibration at $T = 300$ K (*NVT*) followed by a 1-ns collection run (*NVE*) with a step size of 1.0 fs. The IR spectrum obtained by the autocorrelation function of the cell dipole moment gave similar results (data not shown). Despite the use of a simple electrostatic model based on the effective point charges, the simulated infrared spectrum is able to reproduce the salient features of the experimental data especially for the lower vibrational modes. In Table 7 the corresponding peak positions are given. Due to anharmonic contributions at finite temperature the modes are slightly lower

TABLE 6: Low Vibrational Modes (in cm^{-1}) for Basic Zinc Benzoate (<600 cm^{-1})

B3LYP/cc-pVDZ-PP[a] force field	deviation[b]	mode degeneracy
11.1	11.7	−0.6 (5.4)	3
13.0	13.2	−0.2 (1.5)	3
31.7	32.0	−0.3 (0.9)	1
36.1	34.0	2.1 (5.8)	3
36.8	34.7	2.1 (5.7)	2
45.2	43.3	1.9 (4.2)	3
63.4	62.4	1.0 (1.6)	3
69.9	68.7	1.2 (1.7)	2
81.8	85.4	−3.6(4.4)	3
92.5	108.2	−15.7 (17.0)	1
100.5	100.6	−0.1 (0.1)	3
125.3	111.7	13.6 (10.8)	3
127.7	140.2	−12.5 (9.8)	3
149.9	147.6	2.3 (1.5)	2
152.9	161.4	−8.5 (5.6)	3
153.2	179.9	−26.7 (17.4)	3
157.6	143.7	13.9 (8.8)	1
190.5	196.1	−5.6(2.9)	3
199.3	197.1	2.2 (1.1)	3
231.8	231.8	0.0 (0)	3
239.8	247.0	−7.2 (3.0)	2
268.3	274.5	−6.2 (2.3)	1
319.8	321.7	−1.9 (0.6)	3
321.0	323.5	−2.5 (0.8)	3
418.7	386.1	32.6 (7.8)	1
418.9	386.1	32.8(7.8)	3
419.0	386.2	32.8 (7.8)	2
455.4	441.9	13.5 (3.0)	3
456.1	442.5	13.6 (3.0)	3
491.6	483.8	7.8 (1.6)	3
496.2	492.8	3.4 (0.7)	2
498.4	482.0	16.4 (3.3)	1
527.3	541.3	−14.0 (2.7)	3

[a] Average values are given for the degenerated modes due to a small numerical inaccuracy. [b] Relative deviation in percent is given in parentheses.

than the vibrational frequencies from the normal coordinate analysis in the harmonic approximation. Note that the two distinct peaks at 1460 and 1650 cm^{-1}, denoted as symmetric and asymmetric $C_{carb}-O_{carb}$ stretches in Table 7, are actually strongly coupled to deformation modes of the phenylene units, which have not been reparameterized. This is the reason for these modes to be slightly shifted to higher wavenumbers. However, also the periodic DFT calculation overestimates these bands.[84,85] We have estimated (B3LYP/cc-pVDZ)[41] the anharmonic contribution to the stretching modes (1300−1700 cm^{-1}) of dilithium terephthalate to be in the range of 30−40 cm^{-1}, which brings the calculated values closer to the experimental data.

To further validate the force field, molecular dynamics simulations are carried out to study thermal properties of MOF-5. Thermal expansion data were obtained from a series of *NPT* (isothermal−isobaric sampling; 0.1 ps thermostat relaxation time and 2.0 ps barostat relaxation time) simulations between 10 and 600 K and 1 atm. We used a Berendsen bath coupling method[86] for thermostating/barostating the system as implemented in TINKER. At each temperature, the system was subjected to 100 ps of temperature equilibration at a constant volume (*NVT*) followed by a 1 ns production run (*NPT*) with a step size of 1.0 fs. Longer runs (10 ns) gave statistically indistinguishable results.

Using the present covalent force field, our simulations confirm the trend of negative thermal expansion as is evident from the plot in Figure 7 where experimental data for MOF-5

TABLE 7: Comparison of Calculated Vibrational Frequencies (in cm^{-1}) for MOF-5

mode	FF, harm approx (this work)	FF, MD (this work)	DFT, harm approx[80]	exp[83]
C_6H_4 torsion	45		43	
"breathing"	55		52	
$\nu_s(Zn-O_{cent})$	256		253	
$\nu_{as}(Zn-O_{cent})$	530	517	512	530
$\nu_{as}(Zn-O_{carb})$	602	565	606	585
wag C_{carb} + wag H($-C_{ph}$)	750	733	749	745
out-of-phase $O_{carb}-C_{carb}-O_{carb}$ bend	840	839	830	830
$\nu_s(C_{carb}-O_{carb})$	1480	1469	1460	1380
$\nu_{as}(C_{carb}-O_{carb})$	1690	1682	1650	1575

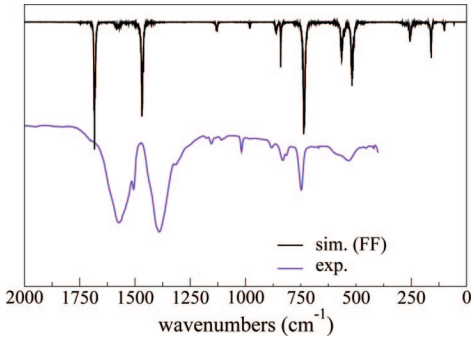

Figure 6. Comparison of the simulated (top) with experimental (bottom)[83] IR spectrum of MOF-5.

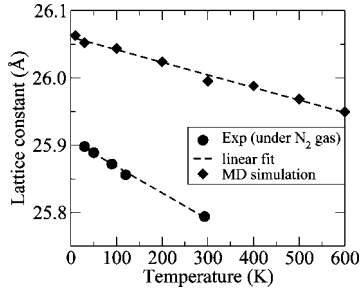

Figure 7. Temperature dependence of the simulated lattice constant compared with experimental data.[78]

under N_2 gas atmosphere[78] are shown for comparison. Our simulated volumetric thermal-expansion coefficient of MOF-5 (-22×10^{-6} K^{-1}), determined by a linear fit of the unit-cell length as a function of temperature (dashed line in Figure 7), is in agreement with other simulated values obtained using different force fields (-24×10^{-6} K^{-1},[87] -39×10^{-6} K^{-1} [76] and -40×10^{-6} K^{-1} [11]). An experimental value of -40×10^{-6} K^{-1} is obtained from a linear fit of the data points plotted in Figure 7 taken from ref 78. This value agrees with a recent direct experimental measurement of the thermal expansion of MOF-5 using neutron powder diffraction in the temperature range 4−600 K (-48×10^{-6} K^{-1}).[88]

Three independent elastic stiffnesses, C_{11}, C_{12} and C_{44}, which are sufficient to describe the elastic properties of a cubic lattice, are calculated by applying three suitable distortions as described by Mehl.[89] The first distortion consists of a uniform compression (or expansion), leading to the evaluation of the bulk modulus B, which is equal to $(C_{11} + 2C_{12})/3$, the second consists of an orthorhombic volume-conserving strain leading to the evaluation of the shear modulus, $(C_{11}-C_{12})$, and the third distortion consists of a monoclinic volume-conserving strain leading to the evaluation of C_{44}. The first two distortions enable C_{11} and C_{12} to be determined separately. Because the last two distortions are volume-conserving, no pressure or stress term needs to be evaluated. The applied distortions (δ) are at most 1%. The bulk modulus is derived as a second derivative of the total energy with respect to the volume of a strained unit cell (V). For the evaluation of strain modulus ($C_{11} - C_{12}$) and C_{44} the following equations were used (for more details, see ref 89):

TABLE 8: Elastic Constants (in GPa) of MOF-5 (at $T = 0$ K)

	B	C_{11}	C_{12}	C_{44}
this work (FF)	14.4	25.3	8.9	2.3
Bahr et al.[82] (DFT)	16.3	27.8	10.6	3.6
Greathouse et al.[76] (FF)	20.0			
Han et al.[87] (FF)	19.4	44.5	6.7	1.8

$$E(\delta) = E(0) + (C_{11} - C_{12})V\delta^2 + \ldots$$

$$E(\delta) = E(0) + \frac{1}{2}C_{44}V\delta^2 + \ldots \quad (6)$$

where terms of order δ^4 and higher in the energy expansion, $E(\delta)$, can safely be neglected. The energy $E(\delta)$ was calculated as a function of $V\delta^2$ with δ varied over the range 0−1% and a linear fit yields $(C_{11} - C_{12})$ or C_{44}. Table 8 compares the elastic constants of MOF-5 derived in the present work with those available in the literature. The bulk modulus (B) and elastic stiffnesses (C_{11}, C_{12}, and C_{44}), calculated directly from the change of the stress tensor applying small strains varied over the range from −1% to 1% for the energy-minimized structure ($T = 0$ K), have quantitatively similar values as compared to those obtained from the second derivative of the energy (see Table 8).

Our simulated elastic constants are in good agreement with the DFT predictions,[82] further confirming that our FF reproduces the DFT potential energy surface. The Young's modulus, $E = (C_{11} - C_{12})(C_{11} + 2C_{12})/(C_{11} + C_{12})$, of 20.7 GPa and the Poisson's ratio, $\nu = C_{12}/(C_{11} + C_{12})$, of 0.26 are also in accord with the DFT results ($E = 21.9$ GPa and $\nu = 0.28$).[82] Overall, our FF results are in better agreement with the DFT calculated values as those from other FF calculations, also shown in Table 8. The elastic stiffnesses, C_{11} and C_{12}, at room temperature were obtained from an ensemble average of the diagonal elements of the directly calculated stress tensor in a constant volume simulation (NVT, $T = 300$ K, 500 ps). An uniaxial compression of 0.5% was applied. The elastic stiffnesses, C_{11} and C_{12}, determined separately, were found to decrease by about 15% and 40%, respectively, giving rise to a decrease of the bulk modulus by about 25% from the value calculated at 0 K. This trend is in agreement with other MD simulations based on different force fields.[76,87]

4. Conclusion

In this contribution we introduce a new, efficient and straightforward approach to parametrize a molecular mechan-

TABLE 9: Force Field Parameters for MOF-5[a]

Bond Stretches

	reference distance, Å	force parameter, mdyn/Å
Zn−O	1.991 (1.992)	1.330 (1.315)
C_{carb}−O_{carb}	1.263 (1.263)	9.826 (9.571)
C_{carb}−C_{ph}	1.503	4.445
C_{ph}−C_{ph}	1.38[b]	6.56[b]

In-Plane Angle Bending

	reference angle, deg	force parameter, (mdyn Å)/rad^2
Zn−O_{cent}−Zn	109.6 (111.2)	0.347 (0.488)
O−Zn−O	105.6 (100.4)	0.246 (0.359)
Zn−O_{carb}−C_{carb}	135.3 (136.2)	0.130 (0.174)
O_{carb}−C_{carb}−O_{carb}	128.3 (127.0)	1.507 (1.574)
O_{carb}−C_{carb}−C_{ph}	117.5	0.999
C_{ph}−C_{ph}−C_{carb}	115.3	0.723

Out-of-Plane Angle Bending

	reference angle, deg	force parameter, (mdyn Å)/rad^2
C_{carb}[c]−O_{carb}	0.0	2.5

Stretch−Stretch (mdyn/Å)

Zn−O_{cent}/Zn−O_{cent}	0.356 (0.0)
C_{carb}−O_{carb}/C_{carb}−O_{carb}	2.697 (0.0)
C_{carb}−O_{carb}/C_{ph}−C_{carb}	1.014 (0.0)
C_{ph}−C_{ph}/C_{ph}−C_{ph}	1.061 (0.0)
C_{ph}−C_{carb}/C_{ph}−C_{ph}	0.293 (0.0)

Stretch−Bend (mdyn/rad)

Zn−O_{cent}/Zn−O_{cent}−Zn	0.035 (0.0)
C_{carb}−O_{carb}/O_{carb}−C_{carb}−O_{carb}	0.413 (0.0)
C_{carb}−O_{carb}/O_{carb}−C_{carb}−C_{ph}	0.509 (0.0)
C_{ph}−C_{carb}/O_{carb}−C_{carb}−C_{ph}	0.418 (0.0)
C_{ph}−C_{ph}/C_{ph}−C_{ph}−C_{ph}	0.124 (0.130)
C_{ph}−C_{carb}/C_{carb}−C_{ph}−C_{ph}	0.353 (0.0)
C_{ph}−C_{ph}/C_{carb}−C_{ph}−C_{ph}	0.092 (0.0)

Torsional Parameters (kcal/mol)

	Fourier term	V_0
Zn−O_{carb}−C_{carb}−O_{carb}	$1/2V_0(1 - \cos 2\varphi)$	4.11 (4.20)
C_{ph}−C_{ph}−C_{carb}−O_{carb}	$1/2V_0(1 - \cos 2\varphi)$	2.12
C_{ph}−C_{ph}−C_{ph}−C_{carb}	$1/2V_0(1 - \cos 2\varphi)$	5.983[b]
C_{carb}−C_{ph}−C_{ph}−H	$1/2V_0(1 - \cos 2\varphi)$	5.983[b]
C_{ph}−C_{ph}−C_{ph}−C_{ph}	$1/2V_0(1 - \cos 2\varphi)$	5.983[b]
C_{ph}−C_{ph}−C_{ph}−H	$1/2V_0(1 - \cos 2\varphi)$	5.385[b]
H−C_{ph}−C_{ph}−H	$1/2V_0(1 - \cos 2\varphi)$	6.881[b]

Van der Waals Parameters

	radius, Å	ε, kcal/mol
Zn	2.29	0.276
O	1.82	0.059
C_{ph} (C_{carb})	1.96 (1.94)	0.056
H	1.62	0.020

[a] Parameters in parentheses correspond to the model without cross terms. [b] These values were adjusted by performing a SCF-MO calculation for the π-system and fixed (see text). [c] Trigonal center.

ics force field from first principles reference data by optimizing a novel objective function with a Genetic Algorithm. The method is highly automated and allows us to derive the force field parameters in a very consistent way, reproducing both structure and normal modes of the reference system accurately. The GA optimization is shown to be very efficient for the strongly coupled, multiminima optimization problem at hand. The crucial point of our method is the objective function, which is based on a representation of both structure and Hessian matrix in a redundant internal coordinate set, which can in principle be defined automatically by the connectivity of the molecular mechanics FF. It can be used in combination with any quantum-mechanical package that is able to provide the optimized structure and the Hessian matrix in Cartesian coordinates. The approach is particularly interesting for systems where experimental data is not available or scarce and a consistent FF parameter set for a broad range of atom types in different chemical environments is needed. A prototypical field of application is the promising class of metal−organic frameworks with their inorganic secondary building units, which was the primary target for the development of our method. As a proof

of concept we reparameterized our previously published MM3 type force field for the Zn_4O based MOF-5.[21,22] In contrast to our previous work we adopted a "building block" strategy using the minimal basic zinc formate $(Zn_4O)(O_2CH)_6$ as a molecular model system for the inorganic SBU and dilithium terephthalate as a model for the linker. Despite the merging of the parameter set from the two smaller reference systems, we find both structure and vibrational frequencies of both the previous larger reference system $Zn_4O(O_2CPh)_6$, as well as the fully periodic model of MOF-5, are well reproduced. This is particularly true for the low normal modes, which are most relevant for the lattice dynamics of the framework. This has been confirmed by investigating the negative thermal expansion and the elastic constants of MOF-5, which are in good accord with experimental data and first principles predictions.

With our strategy we have already derived a first principles based FF for the covalent organic framework COF-102[90] and could confirm the experimentally observed preference for the network topology.[91] Because no experimental data were used in the parametrization (apart from the tabulated van der Waals parameters)[43] the results of the present work can be regarded as being predictive in nature and could be employed also for not yet synthesized systems. In addition, such a first principles derived FF is also ideally suited for a combined QM/MM approach, using a consistent QM level, for the investigation of reactive processes within MOFs.

In the future, we plan to investigate some other alternatives to GAs, such as Particle Swarm Optimization, to improve the parametrization for large parameter sets with an extended number of cross terms. In addition, we intend to extend the objective function with relative energy deviations for isomeric structures. However, our main focus for further work is to apply the proposed methodology to various classes of relevant metal−organic or covalent framework materials to derive a fully consistent and comparable FF. We envision a database of force field parameters for different types of SBUs and linkers, parametrized with reference to the same level of first principles theory, suitable for molecular simulations of coordination polymers taking the framework flexibility fully into account.

Acknowledgment. The Alfried Krupp von Bohlen und Halbach Stiftung and the Deutsche Forschungsgemeinschaft (SFB-558) are acknowledged for their financial support of this project.

References and Notes

(1) Kitagawa, S.; Noro, S.-I.; Nakamura, T. *Chem. Commun.* **2006**, 701.
(2) James, S. L. *Chem. Soc. Rev.* **2003**, *32*, 276.
(3) Janiak, C. *Dalton Trans.* **2003**, 2781.
(4) Rowsell, J. L. C.; Yaghi, O. M. *Microporous Mesoporous Mater.* **2004**, *73*, 3.
(5) Demontis, P.; Suffritti, G. B. *Chem. Rev.* **1997**, *97*, 2845 and references therein.
(6) Chui, S. S. Y.; Lo, S. M. F.; Charmant, J. P. H.; Orpen, A. G.; Williams, I. D. *Science* **1999**, *283*, 1148.
(7) Park, K. S.; Ni, Z.; Cote, A. P.; Choi, J. Y.; Huang, R. D.; Uribe-Romo, F. J.; Chae, H. K.; O'Keeffe, M.; Yaghi, O. M. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 10186.
(8) Serre, C.; Millange, F.; Thouvenot, C.; Nogues, M.; Marsolier, G.; Louer, D.; Ferey, G. *J. Am. Chem. Soc.* **2002**, *124*, 13519.
(9) Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. *J. Am. Chem. Soc.* **1992**, *114*, 10024.
(10) Mayo, S. L.; Olafson, B. D.; Goddard, W. A. *J. Phys. Chem.* **1990**, *94*, 8897.
(11) Dubbeldam, D.; Walton, K. S.; Ellis, D. E.; Snurr, R. Q. *Angew. Chem. Int. Ed.* **2007**, *46*, 4496.
(12) Huang, B. L.; McGaughey, A. J. H.; Kaviany, M. *Int. J. Heat Mass Transf.* **2007**, *50*, 393.
(13) Dybtsev, D. N.; Chun, H.; Kim, K. *Angew. Chem. Int. Ed.* **2004**, *43*, 5033.
(14) Uemura, K.; Yamasaki, Y.; Komagawa, Y.; Tanaka, K.; Kita, H. *Angew. Chem. Int. Ed.* **2007**, *46*, 6662.
(15) Millange, F.; Guillou, N.; Walton, R.; Greneche, J.-M.; Margiolaki, I.; Ferey, G. *Chem. Commun.* **2008**, 4732.
(16) Brenner, D. W. *Phys. Status Solidi B* **2000**, *217*, 23.
(17) van Duin, A. C. T.; Dasgupta, S.; Lorant, F.; Goddard, W. A. *J. Phys. Chem. A* **2001**, *105*, 9396.
(18) Greathouse, J. A.; Allendorf, M. D. *J. Am. Chem. Soc.* **2006**, *128*, 10678.
(19) Greathouse, J. A.; Allendorf, M. D. *J. Am. Chem. Soc.* **2006**, *128*, 13312.
(20) Norrby, P. O.; Brandt, P. *Coord. Chem. Rev.* **2001**, *212*, 79.
(21) Tafipolsky, M.; Amirjalayer, S.; Schmid, R. *J. Comput. Chem.* **2007**, *28*, 1169.
(22) Amirjalayer, S.; Tafipolsky, M.; Schmid, R. *Angew. Chem. Int. Ed.* **2007**, *46*, 463.
(23) Yang, Q. Y.; Zhong, C. L. *J. Phys. Chem. B* **2005**, *109*, 11862.
(24) Hoops, S. C.; Anderson, K. W.; Merz, K. M. J. *J. Am. Chem. Soc.* **1991**, *113*, 8262.
(25) Ryde, U. *Proteins* **1995**, *21*, 40.
(26) Erhart, P.; Juslin, N.; Goy, O.; Nordlund, K.; Mueller, R.; Albe, K. *J. Phys. Condens. Matter* **2006**, *18*, 6585.
(27) Salles, F.; Ghoufi, A.; Maurin, G.; Bell, R. G.; Mellot-Draznieks, C.; Ferey, G. *Angew. Chem. Int. Ed.* **2008**, *47*, 8487.
(28) Furche, F.; Perdew, J. P. *J. Chem. Phys.* **2006**, *124*, 044103.
(29) Allinger, N. L.; Yuh, Y. H.; Lii, J.-H. *J. Am. Chem. Soc.* **1989**, *111*, 8551.
(30) Mattesini, M.; Soler, J. M.; Yndurain, F. *Phys. Rev. B* **2006**, *73*, 094111.
(31) Hubner, O.; Gloss, A.; Fichtner, M.; Klopper, W. *J. Phys. Chem. A* **2004**, *108*, 3019.
(32) Sagara, T.; Klassen, J.; Ganz, E. *J. Chem. Phys.* **2004**, *121*, 12543.
(33) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
(34) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
(35) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
(36) Rayon, V. M.; Valdes, H.; Diaz, N.; Suarez, D. *J. Chem. Theory Comput.* **2008**, *4*, 243.
(37) Wilson, A. K.; Woon, D. E.; Peterson, K. A.; Dunning, T. H. *J. Chem. Phys.* **1999**, *110*, 7667.
(38) Figgen, D.; Rauhut, G.; Dolg, M.; Stoll, H. *Chem. Phys.* **2005**, *311*, 227.
(39) Peterson, K. A.; Puzzarini, C. *Theor. Chem. Acc.* **2005**, *114*, 283.
(40) Balabanov, N. B.; Peterson, K. A. *J. Chem. Phys.* **2005**, *123*, 064107.
(41) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J.C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; iashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision B.04; Gaussian, Inc.: Pittsburgh PA, 2003.
(42) Besler, B. H.; Merz, K. M.; Kollman, P. A. *J. Comput. Chem.* **1990**, *11*, 431.
(43) Allinger, N. L.; Zhou, X. F.; Bergsma, J. *THEOCHEM* **1994**, *118*, 69.
(44) Ponder, J. W.; Richards, F. M. *J. Comput. Chem.* **1987**, *8*, 1016; TINKER version 4.2, June 2004; http://dasher.wustl.edu/tinker/.
(45) Allinger, N. L.; Li, F.; Yan, L.; Tai, J. C. *J. Comput. Chem.* **1990**, *11*, 868.
(46) Maple, J. R.; Hwang, M. J.; Stockfisch, T. P.; Dinur, U.; Waldman, M.; Ewig, C. S.; Hagler, A. T. *J. Comput. Chem.* **1994**, *15*, 162.
(47) Nevins, N.; Chen, K.-S.; Allinger, N. L. *J. Comput. Chem.* **1996**, *17*, 669.
(48) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 490.
(49) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225.
(50) Maple, J. R.; Dinur, U.; Hagler, A. T. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 5350.
(51) Leonard, J. M.; Ashman, W. P. *J. Comput. Chem.* **1990**, *11*, 952.

(52) Dasgupta, S.; Yamasaki, T.; Goddard, W. A. *J. Chem. Phys.* **1996**, *104*, 2898.
(53) Seminario, J. M. *Int. J. Quantum Chem.* **1996**, *60*, 1271.
(54) Palmo, K.; Pietila, L. O.; Krimm, S. *J. Comput. Chem.* **1991**, *12*, 385.
(55) Wang, J. M.; Kollman, P. A. *J. Comput. Chem.* **2001**, *22*, 1219.
(56) Goldberg, D. E. *Genetic Algorithms in Search, Optimization, and Machine Learning*; Addison-Wesley: New York, 1989.
(57) Hunger, J.; Beyreuther, S.; Huttner, G.; Allinger, K.; Radelof, U.; Zsolnai, L. *Eur. J. Inorg. Chem.* **1998**, 693.
(58) Cundari, T. R.; Fu, W. T. *Inorg. Chim. Acta* **2000**, *300*, 113.
(59) Strassner, T.; Busold, M.; Radrich, H. *J. Mol. Model.* **2001**, *7*, 374.
(60) Mostaghim, S.; Hoffmann, M.; Konig, P.; Frauenheim, T.; Teich, J. *Congr. Evolutionary Comput.* **2004**, *1*, 212.
(61) Globus, A.; Menon, M.; Srivastava, D. *Cmes-Comput. Model. Eng. Sci.* **2002**, *3*, 557.
(62) Bukkapatnam, S.; Malshe, M.; Agrawal, P. M.; Raff, L. M.; Komanduri, R. *Phys. Rev. B* **2006**, *74*, 224102.
(63) Brutovsky, B.; Ulicny, J.; Miskovsky, P.; Lisy, V.; Chinsky, L. *Biospectroscopy* **1997**, *3*, 61.
(64) Metcalfe, T. S.; Charbonneau, P. *J. Comput. Phys.* **2003**, *185*, 176.
(65) Mallik, A.; Runge, K.; Cheng, H.-P.; Dufty, J. *Mol. Simul.* **2005**, *31*, 695.
(66) Charbonneau, P., Release Notes for PIKAIA 1.2, NCAR Technical Note TN-451+STR (April 2002); http://www.hao.ucar.edu/Public/models/pikaia/pikaia.html.
(67) Vaiana, A. C.; Cournia, Z.; Costescu, I. B.; Smith, J. C. *Comput. Phys. Commun.* **2005**, *167*, 34.
(68) Sipachev, V. A. *THEOCHEM* **1985**, *121*, 143. We have used slightly modified subroutines from the program SHRINK written by Sipachev.
(69) Wilson, E. B.; Decius, J. C.; Cross, P. C. *Molecular Vibrations - The Theory of Infrared and Raman Vibrational Spectra*; Dover, NY, 1980.
(70) Hattig, C. *J. Chem. Phys.* **2003**, *118*, 7751.
(71) Hrenar, T.; Rauhut, G.; Werner, H. J. *J. Phys. Chem. A* **2006**, *110*, 2060.
(72) Subotnik, J. E.; Head-Gordon, M. *J. Phys. Condens. Matter* **2008**, *20*, 294211.
(73) Han, S. S.; Goddard, W. A. *J. Am. Chem. Soc.* **2007**, *129*, 8422.
(74) Yin, M.-C.; Wang, C.-W.; Al, C.-C.; Yuan, L.-J.; Sun, J.-T. *Wuhan Univ. J. Nat. Sci.* **2004**, *9*, 939.
(75) Maurin, G. Personal communication, 2008.
(76) Greathouse, J. A.; Allendorf, M. D. *J. Phys. Chem. C* **2008**, *112*, 5795.
(77) Gould, S. L.; Tranchemontagne, D.; Yaghi, O. M.; Garcia-Garibay, M. A. *J. Am. Chem. Soc.* **2008**, *130*, 3246.
(78) Rowsell, J. L. C.; Spencer, E. C.; Eckert, J.; Howard, J. A. K.; Yaghi, O. M. *Science* **2005**, *309*, 1350.
(79) Yildirim, T.; Hartman, M. R. *Phys. Rev. Lett.* **2005**, *95*, 215504.
(80) Civalleri, B.; Napoli, F.; Noel, Y.; Roetti, C.; Dovesi, R. *Cryst. Eng. Commun.* **2006**, *8*, 364.
(81) Blomqvist, A.; Araujo, C. M.; Srepusharawoot, P.; Ahuja, R. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 20173.
(82) Bahr, D. F.; Reid, J. A.; Mook, W. M.; Bauer, C. A.; Stumpf, R.; Skulan, A. J.; Moody, N. R.; Simmons, B. A.; Shindel, M. M.; Allendorf, M. D. *Phys. Rev. B* **2007**, *76*, 184106.
(83) Hermes, S.; Schroeder, F.; Amirjalayer, S.; Schmid, R.; Fischer, R. A. *J. Mater. Chem.* **2006**, *16*, 2464.
(84) In principle, if experimental vibrational frequencies are available, the DFT Hessian matrix can be modified prior to the parameterization procedure by direct scaling of primitive valence force constants to match experimental spectra according to a procedure developed by Baker et al. in ref 85.
(85) Baker, J.; Jarzecki, A. A.; Pulay, P. *J. Phys. Chem. A* **1998**, *102*, 1412.
(86) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.
(87) Han, S. S.; Goddard, W. A. *J. Phys. Chem. C* **2007**, *111*, 15185.
(88) Zhou, W.; Wu, H.; Yildirim, T.; Simpson, J. R.; Walker, A. R. H. *Phys. Rev. B* **2008**, *78*, 054114.
(89) Mehl, M. J. *Phys. Rev. B* **1993**, *47*, 2493.
(90) El-Kaderi, H. M.; Hunt, J. R.; Mendoza-Cortes, J. L.; Cote, A. P.; Taylor, R. E.; O'Keeffe, M.; Yaghi, O. M. *Science* **2007**, *316*, 268.
(91) Schmid, R.; Tafipolsky, M. *J. Am. Chem. Soc.* **2008**, *130*, 12600.

JP807487F