# Derivation of a New Force Field for Crystal-Structure Prediction Using Global Optimization: Nonbonded Potential Parameters for Amines, Imidazoles, Amides, and Carboxylic Acids

**Anna Jagielska, Yelena A. Arnautova, and Harold A. Scheraga***

*Baker Laboratory of Chemistry and Chemical Biology, Cornell University, Ithaca, New York 14853-1301*

*Received: February 9, 2004; In Final Form: June 1, 2004*

New nonbonded parameters of a "6-exp-1" all-atom force field are presented for amines, imidazoles, amides, and carboxylic acids. These results are part of our work to derive a general force field for the prediction of crystal and protein structures (the nonbonded parameters for aliphatic and aromatic hydrocarbons and alcohols were published recently, *J. Phys. Chem. B* **2003**, *107*, 7143). The parameters were derived by using our global-optimization-based method (*J. Phys. Chem. B* **2003**, *107*, 712) which provides the best possible set of parameters for crystal structure prediction for a given form of the potential. The method consists of two steps. First, an initial set of parameters is derived from ab initio quantum mechanical interaction energies of dimers; then, the initial set is refined to satisfy the following criteria: the parameters should reproduce the observed crystal structures and sublimation enthalpies, and the experimental crystal structure should correspond to the global minimum of the potential. The refinement procedure uses information about the potential energy surface of the crystal obtained by global energy minimization. Vector Monte Carlo (*J. Phys. Chem. B* **2003**, *107*, 712) and force-minimization methods are used in the refinement step. The parameters presented in this paper reproduce experimental crystal structures with high accuracy. The experimental structures of all molecules tested were found as the global or one of the lowest-energy minima of the potential.

## 1. Introduction

An accurate force field is a necessary element in the prediction of conformations of biological molecules and crystal structures, which is currently at the center of interest in many fields of science. The accuracy of available potentials is still insufficient for successful predictions for many molecules, as shown by the results of blind predictions of the structures of proteins[1] and crystals.[2,3] Therefore, further improvement of potentials is required.

Since the structures of crystals and biomolecules are determined mainly by nonbonded interactions, the accuracy of the nonbonded part of the potential plays a major role in structure prediction. This work is a continuation of our previously published work[4] on the development and refinement of the nonbonded (6-exp) parameters of an all-atom force field. Our method combines ab initio quantum mechanical calculations of interaction energies and crystal-structure data and uses a global optimization procedure[5] to derive the atomic nonbonded parameters.

There are several possible reasons for the unsatisfactory accuracy of the nonbonded parameters of the existing force fields. One of them may be the too-simplified mathematical form of the potential used, which cannot describe interatomic interactions accurately. It was shown[6−10] that more sophisticated potentials, for example, those that include additional interaction sites[6,7] (besides those located on atoms) or a more accurate description of electrostatics and inclusion of polarization,[8−10] can improve the results of the calculations significantly. However, a complicated potential form greatly extends the time of the calculations and limits the possibility to search the space of structural parameters exhaustively, which is necessary for a successful prediction. As a result, the potentials used for a global search (in the whole range of structural parameters) are a compromise between accuracy and simplicity.

Another reason is the insufficient accuracy and insufficient amount of data used for deriving the parameters. The methods for deriving the nonbonded parameters are very diverse and can be divided into three broad categories according to the type of data used: methods based on ab initio quantum mechanical calculations,[7,8,11−13] methods based on experimental data,[14−35] and a combination of both.[4,9,36,37] In ab initio-derived potentials, the parameters are fitted to reproduce the calculated interaction energies of dimers or trimers of various molecules (or interaction energies between a molecule of interest and a probe (usually a rare-gas atom)[12]). The advantage of ab initio calculations is that interaction energies can be calculated for a wide range of intermolecular distances; therefore, it can be expected that the global shape of the potential is sufficiently accurate. This method can potentially provide much valuable information about interatomic interactions between any types of atoms. This is especially important in the case of molecules containing less-common atom types with much less (or no) experimental data available.

Much effort has been made to derive van der Waals parameters based completely on ab initio calculations.[7,8,11−13] However, the results of these efforts show that, in many cases, the van der Waals parameters derived solely from ab initio data yield condensed-phase properties in poor agreement with experiment.[7,12] Ab initio interaction energies are obtained for 0 K temperature and usually for dimers; thus, they are related to the gas phase, and the resulting potential parameters may not

* To whom correspondence should be addressed. Tel: (607)255-4034. Fax: (607)254-4700. E-mail: has5@cornell.edu.

10.1021/jp040115f CCC: $27.50 
Published on Web 07/10/2004

be well transferable to the condensed phase (liquids and crystals). Including multibody effects by introducing the interaction of at least trimers certainly can help to improve the transferability,[8,10] but it would also make the calculations extremely demanding in terms of computational time, especially for the higher-level quantum mechanical methods.

Another common source of inaccuracy of the ab initio calculated interaction energies is the limitation of low-level quantum mechanical methods to take electron correlation fully into account. Although the level of theory equivalent to the second-order Møller−Plesset perturbation method (MP2) usually includes the major part of electron correlation, for some systems, higher-order theory must be used to obtain the required accuracy for the interaction energies.[38,39] It was also shown[40] that the dispersion contribution to the total interaction energy is sensitive to the basis set used and that very large basis sets have to be used in order to evaluate the dispersion energies correctly. Mooij et al.[8,10] have shown that calculations carried out at a high level of theory and including an accurate description of electrostatic interactions, polarization, and anisotropy of interactions in ab initio-derived potentials (gas phase), as well as using interaction data for trimers, can give nonbonded parameters that provide a highly accurate description of the crystal state. However, these types of ab initio calculations are computationally expensive and make the derivation of van der Waals parameters *solely* from the quantum mechanical data unfeasible for larger molecules.

In the experimental nonbonded potentials, the parameters are fitted to reproduce experimental data available for the liquid and crystal states (e.g., liquid densities, crystal-structure parameters, sublimation enthalpies, and lattice frequencies).[14−35] For some force fields, such as OPLS,[29] AMBER,[26] and GROMOS,[35] simulations of liquids were carried out to obtain parameters for the van der Waals part of the potentials. Since it is reasonable to expect that some intermolecular interactions in crystals are essentially the same as the corresponding interactions in more flexible molecular aggregates (for example, nonbonded interactions within a protein molecule), the experimental structural and thermodynamic data for crystals have been used not only for the development of force fields for crystal calculations[14−25,27,28,30−34] but also for deriving van der Waals parameters of general all-atom potentials[17,18,27,28] for a wide range of molecules, including proteins. Since experimental data used for the derivation of nonbonded parameters in most cases pertain to the condensed state, the multibody effects are somehow taken into account.

The accuracy and transferability of the experimental potentials depend strongly on a sufficient amount of experimental data used for their development. A large number of experimental data is necessary to obtain an adequate description of interatomic interactions. Thus, fitting the parameters to reproduce experimental crystal structures of a very limited number of molecules may not provide enough information about the interactions at all ranges of distances (from very short to very long), since such potentials are optimized only for the ranges of distances that occur in the available experimental structures. If a sufficient amount of experimental data about interactions over a wide range of distances is not available, the fitted potential can have nonphysical minima in poorly described regions of the space of structural parameters, which can have energies lower than the minimum corresponding to the observed structure. Generally, the problem is even more difficult because parameters for several atom types are usually being fitted simultaneously, and a strong correlation exists between parameters of one potential function[19,32,33] as well as between parameters of potentials describing different types of atom−atom interaction.[4,37]

Ab initio and experimental approaches for deriving nonbonded parameters have their strong and weak points; hence, "hybrid" methods combining the advantages of both approaches and avoiding their drawbacks have been proposed.[4,37] Rare-gas atoms were used as probes of the van der Waals surfaces of model compounds in ab initio calculations.[37] The ab initio data were then used to determine the relative values of the van der Waals parameters, whereas the absolute values were determined by reproducing experimental condensed-phase properties (heats of vaporization and molecular volumes of pure liquids).

Recognizing the advantages of using both ab initio and experimental data for force-field development, we recently proposed[4] a new procedure for deriving nonbonded parameters. It takes into consideration ab initio interaction energies of dimers, as well as experimental data (crystal structures and sublimation enthalpies) for crystals of small organic molecules.

In the first part of the procedure, a starting parameter set is derived by fitting to ab initio interaction energies of dimers of selected molecules (the MP2/6-31G* method is chosen so that it can be applied to many medium-size molecules). The second part includes evaluation and refinement of the starting set of parameters by carrying out crystal calculations. At this stage, the method makes use not only of experimental information about the structure and energy of observed crystal forms but also information about the potential energy surface of the crystal obtained as a result of *global* energy minimization. Traditionally, the main criterion for evaluating potentials has been their ability to reproduce the experimental structure and other experimental data (e.g., enthalpies of sublimation).[14−25,27,28,30−34] This criterion is a necessary but not a sufficient one[10] because it shows only that the experimental structure corresponds to a minimum (not necessarily a low-energy minimum) of the potential energy surface. To be useful for prediction, potentials should be able to describe the observed structure as the global or one of the lowest-energy minima.[4,10,41] In other words, an accurate potential should satisfy the following requirements: it should reproduce experimental crystal structures and sublimation enthalpy with a certain accuracy, and the lowest minima of the potential should represent the observed crystal structures, with one of them (in an ideal case, the global minimum) corresponding to the most stable crystal polymorph. It was shown[5] that the results of a global search can be used not only for potential evaluation but also for parameter refinement, and the special refinement method called Vector Monte Carlo (VMC) was developed.[5] It uses the information about the potential energy surface of the crystal obtained from global energy optimization to refine the potential parameters so that they satisfy all the requirements.

In the second part of our procedure, *both* local and global lattice-energy minimizations are carried out for a set of molecules, for which experimental crystal data (crystal structure parameters and sublimation enthalpies) are available, for evaluation of the ab initio-derived parameters. If the parameters are found to provide satisfactory results, they are accepted as final; otherwise, the VMC method is used for their refinement. If the ab initio calculations fail to yield a good starting set of potential parameters, a force minimization method[5] is used to obtain a starting set for the VMC procedure.

The procedure described above was applied successfully to derive nonbonded parameters for aliphatic hydrocarbons, aromatic hydrocarbons, and alcohols[4] as part of our work on the development of van der Waals parameters for a general all-

atom force field applicable to different systems (such as crystals and biological molecules). As a continuation of this work, we present here the nonbonded parameters obtained for the following classes of molecules: amines, imidazoles, amides, and carboxylic acids.

## 2. Method

**2.1. Form of the Potential Energy Function.** The energies of the dimers, as well as the lattice energies of the crystal structures, were calculated as a sum of pairwise atom−atom interactions. The energy of interaction between atoms *i* and *j* was considered to be a sum of nonbonded, i.e., exchange repulsion, dispersion attraction, and electrostatic energies

$$E_{ij}^{\text{inter}} = E_{ij}^{\text{nbond}} + E_{ij}^{\text{electr}} \tag{1}$$

The nonbonded energy, $E_{ij}^{\text{nbond}}$, was modeled by using the 6-exp potential function

$$E_{ij}^{\text{nbond}} = -A_{ij} r_{ij}^{-6} + B_{ij} \exp(-C_{ij} r_{ij}) \tag{2}$$

where $r_{ij}$ is the distance between atoms *i* and *j* of different molecules (monomers); $A_{ij}$, $B_{ij}$, and $C_{ij}$ are parameters of the potential. The electrostatic interactions were modeled by using the Coulomb formula

$$E_{ij}^{\text{electr}} = \frac{q_i q_j}{\epsilon r_{ij}} \tag{3}$$

where $q_i$ and $q_j$ are point charges and the dielectric constant, $\epsilon$, was taken as unity. The electrostatic energy was calculated using the Ewald summation[42] without including a dipole moment correction term.[41]

The atom−atom contributions were summed in a special way to avoid small discontinuities of energy due to the fact that the nonbonded energy terms vanish only at infinity; these terms were smoothed ("feathered") to zero at a large but finite distance using a cubic spline, $f(r)$, and a cutoff,[43] chosen to ensure that the energy and its first derivative are continuous everywhere: $f(\omega_1 r_0) = E(\omega_1 r_0)$; $f'(\omega_1 A r_0) = E'(\omega_1 A r_0)$; $f(\omega_2 r_0) = f'(\omega_2 r_0) = 0$, where $r_0$ is an equilibrium distance for interaction between two atoms; $\omega_1$ and $\omega_2$ are cutoff parameters. In the present work, the cutoff parameter $\omega_2$, which specifies the distance at which the nonbonded energies and gradients become zero, was set to 8.5, and the parameter $\omega_1$, which specifies the start of the feather, was set to $\omega_2 - 0.5$. It was shown[43] that an almost negligible fraction of the total energy will be lost if the cutoff parameters are set to the values greater than 4.0 and 4.5, respectively.

The 6-exp nonbonded atom−atom potential used in the present work is described by

$$E(r) = -\frac{A}{r^6} + B \exp(-Cr), \quad r \le \omega_1 r_0$$

$$E(r) = f(r) = ar^3 + br^2 + cr + d, \quad \omega_1 r_0 \le r \le \omega_2 r_o$$

$$E(r) = 0, \quad r \ge \omega_2 r_0 \tag{4}$$

where *a*, *b*, *c*, and *d* are constants calculated for each pair of atoms (see ref 43).

To reduce the number of parameters in a fitting procedure, combination rules for the nonbonded parameters $A_{ij}$, $B_{ij}$, and $C_{ij}$ are used for describing heteroatomic interactions. The commonly used[3,19,24,32−34] combination rules:

$$A_{ij} = \sqrt{A_{ii} A_{jj}};\ B_{ij} = \sqrt{B_{ii} B_{jj}};\ C_{ij} = (C_{ii} + C_{jj})/2$$

were selected for the calculations, where $A_{ii}$, $B_{ii}$, and $C_{ii}$ are parameters of pairwise interaction between atoms of the same type, *i*. For some systems, the combination rules may provide a poor approximation for heteroatomic interactions; however, in our previous work[4] and in this work, use of the combination rules for hydrogen-bonded atoms led to satisfactory results.

A hydrogen-bonded atom was treated as a separate atom type, with different parameters, compared to the hydrogen atoms bonded to carbon. No special hydrogen-bond potential was used; we assume that the 6-exp-1 potential provides a correct description of the interactions involved in a hydrogen bond.

Point charges, $q_i$, were localized on atoms and were fitted to reproduce the electrostatic potential of the monomers, calculated with the Hartree−Fock wave function in the 6-31G* basis set. The RESP method[26a] implemented in the AMBER 6.0 program[26b] was used for fitting charges. However, for nitrogen atoms bonded to two other atoms, additional lone-pair electron charge sites were added (see section 4.1.2.). Point charges derived in this way were used to calculate the electrostatic interactions (eq 3) between monomers in the dimers and in crystals. The point charges for all the molecules considered in this work are available in Supporting Information.

**2.2. Derivation of Nonbonded Force-Field Parameters.** The procedure used in this work for deriving nonbonded potential parameters has been described in detail elsewhere.[4,5] It consists of the following steps: first, an initial set of force-field parameters is obtained by fitting to ab initio interaction energies of dimers; second, the accuracy of the parameters is evaluated by local and global energy minimizations for a set of small organic molecules containing the same types of atoms as the molecules in the dimers; third, if the accuracy of the resulting parameters is not sufficiently high, i.e., if the potential does not satisfy all three requirements listed in the Introduction, refinement is carried out with the VMC procedure using the ab initio parameters as a starting point. When ab initio parameters were not available or were proven to give unsatisfactory results, the force minimization procedure[5] was used to find a starting set of parameters. The following sections describe the methods used at each step of the procedure.

*2.2.1. Ab Initio Calculations.* Initial nonbonded parameters of the force field were derived by fitting the 6-exp potential functions (eq 2) to the ab initio van der Waals energy landscape. The van der Waals energies at each point were obtained by subtracting the electrostatic part of the interaction energy from the total ab initio interaction energy.

Because the goal of the ab initio calculations was to obtain a starting point for further VMC parameter optimization, we limited the ab initio calculations to interactions between the monomers within the dimers. The multibody interactions are included by subsequent VMC parameter optimization of the crystal structures. All calculations were carried out with the GAUSSIAN program.[44] The dimers consisted of two identical monomers. The geometry of the isolated monomers was fully optimized at the level of the density functional theory (DFT)[45] with the B3LYP functional and 6-311G** basis set and was kept frozen in the dimer geometries. The geometries of the dimers were generated on a grid of the Euler angles and spherical coordinates of the second monomer. For each conformation of a dimer, the total energy was calculated at the level of Møller−Plesset perturbation theory up to second order (MP2) with the 6-31G* basis set. This level of calculation should include the major part of the dispersion energy[39] and, at

relatively low computational cost, provide a sufficient level of accuracy for the interaction energy to derive an initial nonbonded parameter set. The ab initio interaction energy was calculated by subtracting the energies of the monomers from the energy of the dimer. The energies of the monomers in the conformation of the dimer were corrected for the basis set superposition error (BSSE) by using the Boys−Bernardi counterpoise method[46] in which the energy of the dimer and the energies of the monomers have to be calculated with the same basis set, the basis set of the dimer.

*2.2.2. Fitting Method.* The parameters of the 6-exp potential (eq 2) were fitted to reproduce the nonbonded part of the ab initio interaction energy by minimizing the least squares function:

$$F_{\mathrm{fit}}(\mathbf{a}) = \sum_k [E^{(k)}_{\mathrm{abinitio}}(\boldsymbol{r}_k) - E^{(k)}_{\mathrm{nbond}}(\mathbf{a};\boldsymbol{r}_k)]^2 \tag{5}$$

where $E^{(k)}_{\mathrm{abinitio}}(\boldsymbol{r}_k)$ is the nonbonded ab initio energy of the dimer in conformation $k$; $E^{(k)}_{\mathrm{nbond}}(\mathbf{a};\boldsymbol{r}_k)$ is the nonbonded energy of the dimer in conformation $k$ calculated with the potential (eq 2); $\boldsymbol{r}_k$ is the vector of atomic coordinates; and $\boldsymbol{a}$ is the vector of the parameters of the force field.

To improve the accuracy of the fit in the region of the minimum of the nonbonded energy, each component, $k$, in the sum (eq 5) was weighted, using exponential weights $w_k$:

$$w_k = \exp[-(E^{(k)}_{\mathrm{abinitio}}(\boldsymbol{r}_k) - E_{\mathrm{min}})/c] \tag{6}$$

where $E_{\mathrm{min}}$ is the minimal energy in the set of energy values for the given type of dimer, and $c$ is a constant. This type of weighting leads to better fitting in the vicinity of the energy minimum ($w_k \rightarrow 1$) and lower accuracy for the regions of the potential distant from the minimum ($w_k \rightarrow 0$). The value of the constant $c$ was adjusted within the range from 5 to 20 separately for two groups of molecules (viz., amines and imidazoles, and amides and carboxylic acids) to obtain the best fit to the ab initio interaction energy.

The function $F_{\mathrm{fit}}$ (eq 5) was minimized with the Secant Unconstrained Minimization Solver (SUMSL)[47] minimizing procedure. Potentially, the function $F_{\mathrm{fit}}$ can have more than one minimum in the multidimensional space of the force-field parameters. To explore a wide range of parameter space, many local minimizations were carried out, starting from the parameter sets generated on the grid of potential parameters. The initial search was carried out within the range from −50 to +50% of the values of the parameters $A$ and $B$ given by Williams[32−34] for the corresponding types of atoms and within the range from 3 to 5 Å$^{-1}$ for the parameter $C$ (for the parameters $A$ for the hydrogen atom connected to oxygen, having the value zero in Williams' force field, the range from 0 to 20 Å$^6$ kcal/mol was initially searched). Three or five equally displaced values were considered for each dimension in the parameter space (depending on the coarseness of the search and the dimensionality of the parameter space). In most cases, minimizations from different starting sets of parameters led to the same minimum of the function $F_{\mathrm{fit}}$ and the same final set of parameters, providing physically reasonable equilibrium distances of the 6-exp function. If more then one minimum was found in the parameter space, usually only one corresponded to a physically acceptable parameter set.

*2.2.3. Refinement of Potential Parameters: the VMC Method.* To obtain potentials satisfying the requirements discussed in the Introduction, a special VMC procedure for deriving potential parameters was designed.[5] The VMC method minimizes a target function $\boldsymbol{F}(\mathbf{a})$ (eq 7) consisting of the three main components optimized independently at different temperatures (if a set of $n$ molecules is used for parameter refinement, the vector function $\boldsymbol{F}(\mathbf{a})$ has $3n$ components).

$$\boldsymbol{F}(\mathbf{a}) = \begin{pmatrix} f_1 \\ f_2 \\ \dots \\ f_{3n} \end{pmatrix} = \begin{pmatrix} G(\mathbf{a}) \\ P(\mathbf{a}) \\ R(\mathbf{a}) \end{pmatrix} \tag{7}$$

where $G(\mathbf{a}) = (G_1,\dots,G_n)$; $P(\mathbf{a}) = (P_1,\dots,P_n)$; and $R(\mathbf{a}) = (R_1,\dots,R_n)$.

The first component $G(\mathbf{a})$ (eq 8) of the vector is the gap between the energy of the minimized experimental structure and the lowest-energy minimum, found by a global search for a given molecule. At the end of the refinement, the gap should be negative, i.e., the experimental structure should have the lowest energy among all minima found.

$$G_i(\mathbf{a}) = E^i_{\mathrm{m.e.}}(\mathbf{a}) - \min E^i_k(\mathbf{a}) \tag{8}$$

where $E^i_{\mathrm{m.e.}}$ is the energy of the minimized $i$th experimental structure; $E^i_k$ is the energy of minimum $k$ in the set of minima found by global optimization (the set does not include the minimum corresponding to the minimized experimental structure).

The second element $P(\mathbf{a})$ (eq 9) of the minimized vector function is a harmonic penalty function of the deviation of the lattice energy from the experimental enthalpy of sublimation $\Delta H^i_{\mathrm{subl}}$,

$$P_i = (E^i_{\mathrm{m.e.}}(\mathbf{a}) - \Delta H^i_{\mathrm{subl}})^2 \tag{9}$$

The third element $R(\mathbf{a})$ (eq 10) is the function minimizing the structural deviations between the minimized experimental and the experimental structures:

$$R_i(\mathbf{a}) = w_c \cdot diff_c(\mathbf{a}) + w_a \cdot diff_a(\mathbf{a}) + w_t \cdot diff_t(\mathbf{a}) \tag{10}$$

where $w_c$, $w_a$, and $w_t$ are empirical weights.

$$diff_c = (|a - a_{\mathrm{e}}| + |b - b_{\mathrm{e}}| + |c - c_{\mathrm{e}}| + |\alpha - \alpha_{\mathrm{e}}| + |\beta - \beta_{\mathrm{e}}| + |\gamma - \gamma_{\mathrm{e}}|) \tag{11}$$

with $a_{\mathrm{e}}$, $b_{\mathrm{e}}$, $c_{\mathrm{e}}$, $\alpha_{\mathrm{e}}$, $\beta_{\mathrm{e}}$, and $\gamma_{\mathrm{e}}$ being the experimental unit cell parameters;

$$diff_a = \sum_{i=1}^{Z} \Omega_i / Z \tag{12}$$

$$diff_t = \sum_{i=1}^{Z} (fr_i - fr_i^e)^{2/Z} \tag{13}$$

where $Z$ is the number of molecule in the unit cell. The quantities $diff_c$, $diff_a$, and $diff_t$ correspond to unit cell parameters, orientations of all molecules in the unit cell, and translational positions (in fractional coordinates $fr_i$), respectively, and describe deviations between experimental (e) and minimized experimental structures. Rotational angles, $\Omega_i$, characterize similarities of molecular orientations in the experimental and minimized experimental structures. $\Omega_i$ is found as

$$\Omega_i = \arccos(w_{11} + w_{22} + w_{33} - 1)/2 \tag{14}$$

where $w_{11}$, $w_{22}$, and $w_{33}$ are diagonal elements of the matrix $\boldsymbol{W} = R_1^{-1}R_2$, relating rotation matrices, $R_1$ and $R_2$, of corresponding molecules in two structures.

The target function $\boldsymbol{F}(\mathbf{a})$ of eq 7 is minimized using the VMC method.[5] The refinement of potential parameters using the VMC method requires the following information: 1) a starting set of potential parameters; 2) parameters of the experimental crystal structures for a given set of molecules; 3) sublimation enthalpies (if available); and 4) a set of minima found for the molecules by a global search using the selected starting set of potential parameters. This set of minima was restricted to energies that were 2 kcal/mol above the global minimum. If the experimental structure was not the global minimum, then the structures with energies of 2 kcal/mol above the energy of the minimized experimental structure plus all minima below this energy were included in the set.

In principle, refinement of the potential parameters is an iterative process and includes two steps. The first step is refinement of the parameters carried out by optimization of the target function $\boldsymbol{F}(\mathbf{a})$ by using the VMC method. It involves local energy minimizations carried out at each step of the refinement for a set of minima found by the initial global search. The second step is evaluation of the resulting potential model. To evaluate new parameters, local and global energy minimizations are carried out for molecules used in the parameter refinement. The whole procedure was to be iterated, if necessary, by using the new set of minima and the new potential parameters as a starting point. After a good set of parameters (i.e., those satisfying all the requirements listed in the Introduction) is found, global optimizations are also carried out for similar molecules to check the transferability of the potential.

To find a good starting set (or sets) of parameters for the VMC refinement, ab initio calculations are used. In some cases, they may not be able to provide a reasonable starting set of parameters. In this case, the force minimization procedure[5] is used. First, a wide-range random search is carried out to find sets of potential parameters, **a**, corresponding to low values of the function:

$$FF(\mathbf{a}) = \sum_k \sum_i \mathrm{w}_i^k [F_i^k(x_i,\mathbf{a})]^2 + \mathrm{w}' \sum_k (E_\mathrm{e}^k(\mathbf{a}) - \Delta H_\mathrm{subl})^2 \quad (15)$$

where

$$F_i^k(x_i,\mathbf{a}) = -\frac{\partial E_\mathrm{e}^k}{\partial x_i}$$

is the equilibrium force due to a change of the energy $E_\mathrm{e}^k$ of the $k$th experimental crystal structure with respect to the $i$th structural parameter; $x_i$ is a structural parameter; and $w_i^k$ and $w'$ are empirical weights. The requirement for $FF(\mathbf{a})$ to be as close to zero as possible means that all forces should vanish at the observed structures, i.e., these structures are local minima of the 6-exp-1 potential with parameters **a**. Second, we check if any of the resulting sets correspond to low values of the target function $\boldsymbol{F}(\mathbf{a})$ (eq 7). This is carried out by local energy minimizations for all structures found for the molecules under consideration. Again, a few (usually 2−3) best sets of parameters are selected and used as starting points for VMC parameter optimization.

**2.3. Local and Global Energy Minimization.** To check if a given set of nonbonded parameters satisfies the requirements discussed in the Introduction, local (i.e., starting from the experimental crystal structure) and global energy minimizations have to be carried out for the selected molecules. The deviations between the unit cell parameters of the experimental structure and those of the structure obtained as a result of the local energy minimization with a given set of parameters (the minimized experimental structure) should be small (<5%). No symmetry constraints were used during the local minimizations carried out as part of this work. The molecules in the unit cell were treated as independent rigid bodies. Energy minimizations were carried out with the translational and rotational (Euler angles) coordinates of all molecules in the unit cell and the six lattice parameters as independent variables.

Global energy minimizations are necessary to check if a given set of parameters is able to predict the minimized experimental structure as one of the lowest-energy minima. In the present work, global optimization runs were carried out by using the Conformation-Family Monte Carlo (CFMC) method for crystals.[48] The method does not rely on any information about space group symmetry in order to locate the lowest-energy minima. All global optimization calculations in this paper were carried out for the same number of molecules, $Z$, in the unit cell as in the experimental crystal structures. Two global optimization runs were carried out for each chiral molecule: one with the L (or *R*) enantiomer and the other one with the racemic mixture. These two types of runs were usually repeated three times each, using different search parameters (temperatures, perturbation parameters, etc.) to ensure an adequate search for low-energy structures. A detailed description of the method used for potential-energy calculations has been published.[48] The results of the global optimization (i.e., the set of low-energy structures found for each molecule) were used later for the parameter refinement.

## 3. Experimental Data Used for Deriving and Testing Parameters

The molecules used for potential-parameter derivation and evaluation are shown in Figures 1−3. The corresponding experimental crystal structures were taken from the Cambridge Structural Database (CSD).[49] The molecules were selected according to the following criteria: 1) the molecules should contain only C, H (or D), N, and O atoms; molecules with D were used only for geometry, but the potential parameters for D were taken as those for H (the atom types are listed in Table 1); 2) the observed structures should have no disorder. In the case of carboxylic acids, the structures with almost equal C−O and C═O bond lengths (difference less than 0.05 Å) that may be due to orientational disorder of the carboxylic group in the crystal[50] were discarded; 3) in the case of X-ray diffraction data, if several structures are available for a given molecule, the one with lowest discrepancy factor, $R$, was used; 4) if several structures for a given molecule obtained at different temperatures are available, the one corresponding to the lowest temperature is used in order to minimize the errors due to the neglect of temperature effects; 5) the structure obtained by neutron diffraction is preferred, if available; and 6) only structures with one molecule in the asymmetric unit were considered.

Ideally, the discrepancy factor, $R$, should be less than 5% for the molecules used for parameter refinement; however, some structures with larger $R$ factors were used for evaluation of the parameters.

Since the goal of this work is to obtain nonbonded potential parameters, all molecules were considered to be rigid, i.e., with no variation of internal rotational degrees of freedom. Molecular structures, i.e., values of bond lengths, valence, and torsional angles, were taken from the original crystal structure determinations, except that C−H, N−H, and O−H bond lengths were

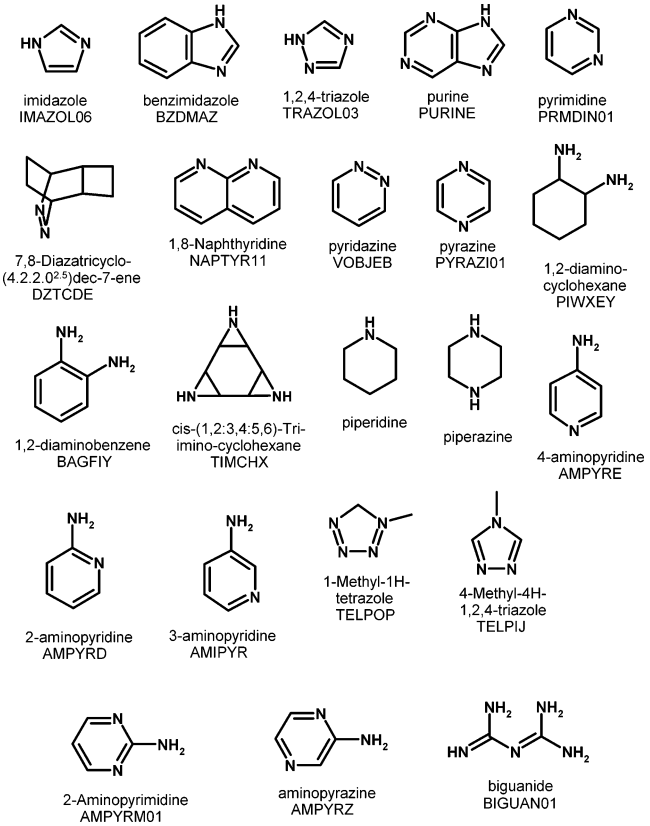



**Figure 1.** Structures of amines and imidazoles. The Cambridge Structural Database reference codes[49] are also shown.

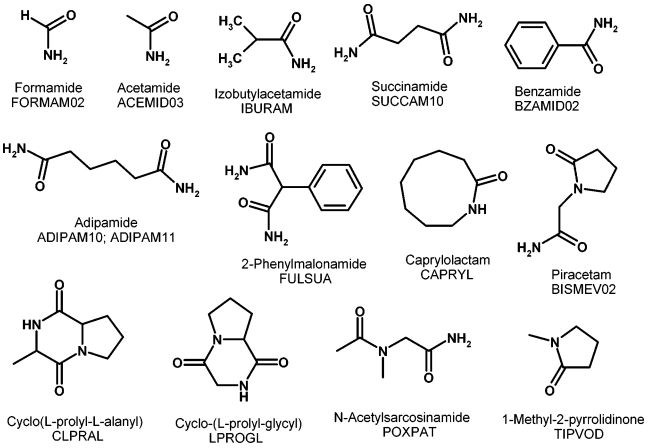



**Figure 2.** Structures of amides. The Cambridge Structural Database reference codes[49] are also shown.

adjusted to the average experimental values of 1.083, 1.009, and 0.983 Å, respectively, on the basis of neutron diffraction data.[51]

Many of the molecules considered here can take more than one polymorphic form. Polymorphism is especially common among amides and carboxylic acids, whereas most of the amine and imidazole molecules considered in this work exist in only one crystal form.

Among the amides, acetamide, benzamide, and adipamide form more than one crystal polymorph. Acetamide crystallizes in two modifications at room temperature: a stable trigonal form[52] and a metastable orthorhombic form (with two molecules in the asymmetric unit).[53] Two crystalline forms were reported for benzamide at room temperature. The stable form, I,[54] consists of large blocks, whereas crystals of the metastable form, II,[55]

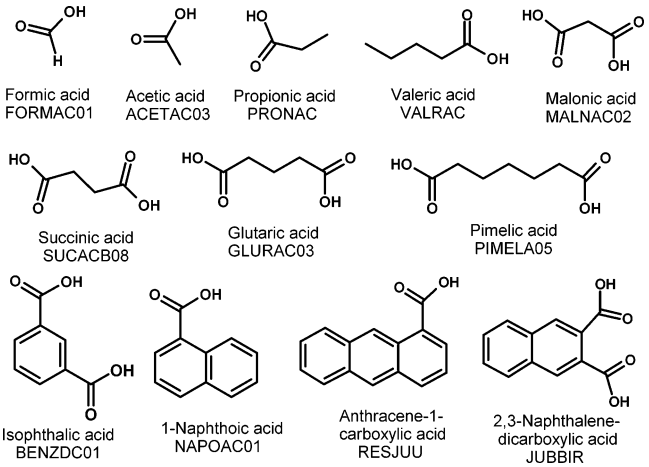
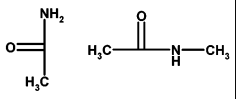
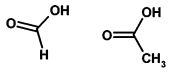



**Figure 3.** Structures of carboxylic acids. The Cambridge Structural Database reference codes[49] are also shown.

**TABLE 1: Atom Types and Molecules Used in ab Initio Calculations of Interaction Energies of Dimers**

| Atom type | Ab initio energies of dimers |
|---|---|
| Aliphatic carbon (C1)<br>Hydrogen bonded to aliphatic carbon (HC) | $CH_4$ $H_3C—CH_3$ |
| Alcoholic oxygen (OH1)<br>Hydrogen in alcohol group (HO1) | $HO-CH_3$ |
| Aliphatic nitrogen (N1)<br>Hydrogen bonded to amine or $sp^2$ nitrogen (HN1) | $H_2N-CH_3$ |
| $sp^2$ nitrogen (NA)<br>Hydrogen bonded to amine or $sp^2$ nitrogen (HN1) |  |
| Aromatic carbon (CA)<br>Hydrogen bonded to aromatic carbon (HA) |  |
| Hydrogen in amide group (HN2) (parameters were taken the same as for HN1)<br>Amide nitrogen (N2)<br>Amide oxygen (O2)<br>Carbon in amide group (C3) |  |
| Hydrogen in carboxyl group (HO2)<br>Carbon in carboxyl group (C2)<br>Carboxyl oxygen bonded to one atom as in C=O group (O1)<br>Carboxyl oxygen bonded to two atoms as in C-O-H (OH2) |  |

have the shapes of needles and plates. Only form I has been well studied experimentally.[54]

The presence of two polymorphs was originally reported for adipamide by identification of distinct crystal morphologies.[56] A crystal structure of the monoclinic form was solved[56] from single-crystal X-ray data, whereas the triclinic form was solved[57] from X-ray powder diffraction data using a direct space global optimization method. No information about the relative stabilities of the forms is available. However, a theoretical study,[58] including modeling of hypothetical crystal packings with different hydrogen-bond networks and calculation of their lattice energies using different potentials, showed that the experimental structure of the monoclinic form was the most stable one among the structures considered.

Carboxylic acids are the best studied, both experimentally and theoretically, of the class of molecules considered in this work. Much effort[50,59−61] has been made to understand the relationship between the functional groups in a carboxylic acid molecule and the crystal structure that it forms, with special

attention paid to the type of hydrogen-bond network formed by a particular molecule. Carboxylic acids are commonly observed to crystallize by forming a dimer motif; however, there are exceptions, such as formic and acetic acids, in which the crystal structure does not include this motif.

Both mono- and dicarboxylic acids have a strong propensity for crystallizing in more than one crystal modification. Formic and acetic acids can exist in two crystal forms: low-temperature and high-pressure (above 4.5 and 0.2 GPa for formic[62] and acetic[63] acids, respectively). The dicarboxylic acids are known to crystallize in two crystal forms, α and $\beta$ (the dimorphism is prevalent in the acids with odd numbers of carbon atoms), the $\beta$-form being stable at room temperature. In the $\beta$-form, the molecules with an even number of carbon atoms exhibit inversion symmetry, while the odd-numbered ones, except malonic acid, have a 2-fold rotational axis. Transformation from $\beta$ to α occurs at around 350 K. Malonic acid is an exception and does not possess any molecular symmetry in either form and exists as a third polymorph, $\gamma$.

Among all the carboxylic acids considered in this paper, propanoic, valeric, and isophthalic acids are the molecules known to form only one polymorphic modification.

Since the VMC method is based on the idea that the observed thermodynamically most stable polymorph corresponds to the global (or one of the lowest) lattice-energy minimum, special care had to be taken to make sure that the experimental structure used in the refinement is in fact the most stable one among all polymorphs known for a given molecule.

More than one polymorph could potentially be used for the refinement if accurate experimental data including information about the relative stabilities of the forms are available. A transition between different polymorphs is often accompanied by changes in molecular geometry (conformational polymorphism), which cannot be reproduced with the rigid-body approximation used in this work. The energies of the crystal packings formed by rigid molecules with experimental geometries corresponding to different polymorphs cannot be compared directly. Therefore, when the experimental data for more than one polymorph of a particular molecule were available, the data for only one, the most stable form, was used for the parameter refinement and evaluation. The only exceptions were the crystal structure of the $\beta$-form of malonic acid (MALNAC02) used for the refinement and the crystal structure of form II (ADIPAM11) of adipamide used for parameter evaluation. For both structures, only local energy minimizations were carried out, and the quality of different potentials was assessed solely on their ability to reproduce the experimental structure and sublimation enthalpy accurately.

The available enthalpies of sublimation ($\Delta H_{\mathrm{subl}}$)[64] (values given in Tables 3, 4, 6, and 8, column *E*) were used for energy scaling of the intermolecular potentials (by $P_i$ of eq 9). Taking into account that there are many uncertainties in comparing lattice energies and corresponding sublimation enthalpies and that an average experimental error is about 2 kcal/mol or more, a deviation of a few kcal/mol between the lattice energy and the corresponding sublimation enthalpy was considered to be acceptable.

## 4. Results and Discussion

**4.1. Amines and Imidazoles.** *4.1.1. Derivation of Initial Values of Potential Parameters by ab Initio Calculations.* Initial parameters, based on ab initio calculations described in Section 2.2, were derived sequentially for small sets of molecules containing a few atom types at each step of the derivation. The molecules chosen for ab initio calculations of the interaction energy in dimers are shown in Table 1 (second column). We assumed that the same atom type can be used for the hydrogen atom in an amine group and the hydrogen atom connected to nitrogen in imidazole. To obtain the parameters for $sp^2$ nitrogen (NA), the hydrogen connected to nitrogen (HN1) and aliphatic nitrogen (N1), ab initio interaction energies of dimers of imidazole molecules and dimers of methylamine molecules were used. Since these two molecules share the HN1 atom type, the interaction energy surfaces of their dimers were used together in the fitting procedure. The values of the parameters for aliphatic and aromatic carbon (C1, CA) and hydrogen connected to aliphatic and aromatic carbon (HC, HA) atom types were taken as those obtained previously[4] for hydrocarbons and were frozen during all stages of the fitting procedure for all molecules containing these atom types. The resulting parameters, obtained by fitting all *A*, *B*, and *C* parameters (eq 2) for HN1, NA, and N1 atom types without restrictions, reproduced the ab initio nonbonded interaction energies for dimers of imidazole and methylamine with a standard deviation of 0.028 kcal/mol.

**TABLE 2: Potential Parameters Obtained by ab Initio Calculations and the VMC Refinement for Amines and Imidazoles**

| atom type | | *A*, Å$^6$ kcal/mol | *B*, kcal/mol | *C*, Å$^{-1}$ |
|---|---|---|---|---|
| hydrogen bonded to nitrogen (amines and imidazole) (HN1) | | 4.56 | 2376.8 | 4.75 |
| aliphatic amine nitrogen (N1) | ab initio | 769.3 | 54694.5 | 3.48 |
| | ab initio+VMC | 769.3 | 51218.6 | 3.48 |
| $sp^2$ nitrogen (NA) | | 475.3 | 25836.2 | 3.48 |

However, we obtained relatively large values of the *A* parameters, 1075 and 5053 Å$^6$ kcal/mol for NA and N1 respectively, compared to the value of 760 Å$^6$ kcal/mol estimated from the Slater−Kirkwood formula.[65] This is the result of a strong correlation between the potential parameters.[32,33] The correlations between the *A*, *B*, and *C* parameters for a given atom type, as well as between the parameters describing different types of interactions (for example, H···H and C···C), mean that more than one set of potential parameters can be obtained using the same data. Even though some of these sets may contain parameters with rather unphysical values, each of them, when used as a consistent set, may provide the desired accuracy of crystal calculations. It is generally difficult to avoid the parameter correlations in unrestricted fitting, even using a large set of ab initio or experimental data.[32,33] To decrease the correlation between parameters and obtain more physical values for the *A* parameter for N1 and NA atom types, we refitted the parameters, keeping the parameter *C* for N1 and NA frozen at the value of 3.48 Å$^{-1}$, calculated by Williams,[34] and for HN1 frozen at the value of 4.75 Å$^{-1}$, obtained in the first stage of the fitting procedure. The new values of the potential parameters are given in Table 2. Use of the restrictions on the values of *C* led to a slightly higher standard deviation (0.030 kcal/mol) compared to the one obtained for the unrestricted fitting. However, it enabled us to obtain more physical values of the *A* parameters, i.e., closer to those calculated from the Slater−Kirkwood formula.

*4.1.2. Evaluation and Refinement of Potential Parameters.* Molecules selected for evaluation of the potential parameters obtained for amines and imidazoles are shown in Figure 1. We divided these molecules into four groups: 1) imidazoles; 2) molecules containing aromatic nitrogen or the nitrogen bonded to two other non-hydrogen atoms; 3) amines; and 4) molecules

**TABLE 3: Results of Local and Global Energy Minimizations Carried out for Amine and Imidazole Molecules (Figure 1) with the Potential Parameters Obtained by ab Initio Calculations and with Final Parameters After the VMC Refinement**

| compound[a] | parameter set | $a^b$ | $b^b$ | $c^b$ | $\alpha^c$ | $\beta^c$ | $\gamma^c$ | $\Delta t^d$ | $\Delta\Omega$ [e] | $E^f$ | $\Delta E^g$ | rank[h] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| IMAZOL06 | exptl | 7.57 | 5.37 | 9.79 | 90.0 | 119.1 | 90.0 | | | −18.02 | | |
| $P2_1/c$, $Z=4$ | ab initio | 2.4 | −2.7 | 1.6 | | −0.5 | | 0.014 | 4.31 | −17.83 | 0.0 | 1 |
| BZDMAZ | exptl | 6.94 | 13.49 | 6.81 | 90.0 | 90.0 | 90.0 | | | −24.33 | | |
| $Pna2_1$, $Z=4$ | ab initio | 1.7 | 0.7 | −1.7 | | | | 0.005 | 5.86 | −21.07 | 0.46 | >5 |
| TRAZOL03 | exptl | 9.76 | 9.35 | 6.99 | 90.0 | 90.0 | 90.0 | | | −20.08 | | |
| *Pbca*, $Z=8$ | ab initio | 1.7 | −3.1 | 1.3 | | | | 0.012 | 6.92 | −19.87 | 0.31 | >5 |
| PURINE | exptl | 15.55 | 9.37 | 3.66 | 90.0 | 90.0 | 90.0 | | | | | |
| $Pna2_1$, $Z=4$ | ab initio | −4.2 | 1.0 | 2.9 | | | | 0.014 | 6.89 | −28.26 | 0.28 | 2 |
| PRMDIN01 | exptl | 11.56 | 9.46 | 3.69 | 90.0 | 90.0 | 90.0 | | | −11.93 | | |
| $Pna2_1$, $Z=4$ | ab initio | −1.3 | 0.9 | 2.9 | | | | 0.001 | 1.49 | −13.66 | 0.00 | 1 |
| DZTCDE | exptl | 5.89 | 10.04 | 6.58 | 90.0 | 113.8 | 90.0 | | | | | |
| $P2_1/m$, $Z=2$ | ab initio | 0.7 | 2.3 | 0.2 | | −0.2 | | 0.020 | 0.06 | −16.47 | 0.69 | 3 |
| NAPTYR11 | exptl | 6.14 | 10.41 | 11.26 | 90.0 | 117.8 | 90.0 | | | | | |
| $P2_1/c$, $Z=4$ | ab initio | 0.6 | −2.2 | 7.5 | | 0.5 | | 0.028 | 3.87 | −19.66 | 0.0 | 1 |
| VOBJEB | exptl | 3.79 | 10.74 | 9.72 | 90.0 | 91.4 | 90.0 | | | | | |
| $P2_1/n$, $Z=4$ | ab initio | 2.8 | 1.7 | 1.1 | | −0.9 | | 0.028 | 7.09 | −14.54 | 0.52 | >5 |
| PYRAZI01 | exptl | 9.33 | 5.85 | 3.73 | 90.0 | 90.0 | 90.0 | | | −13.50 | | |
| *Pmnn*, $Z=2$ | ab initio | 0.5 | 0.3 | 2.8 | | 1.9 | | 0.008 | 6.33 | −12.30 | 0.82 | >5 |
| PIWXEY | exptl | 5.25 | 8.47 | 7.69 | 90.0 | 90.0 | 90.0 | | | | | |
| $P2_12_12$, $Z=2$ | ab initio | 2.7 | −3.5 | 1.8 | | | | 0.006 | 3.77 | −18.29 | 0.94 | 4 |
| | ab initio + VMC | 2.5 | −3.9 | 2.0 | | | | 0.007 | 3.86 | −18.59 | 0.21 | 4 |
| BAGFIY | exptl | 10.32 | 7.54 | 7.72 | 90.0 | 99.9 | 90.0 | | | | | |
| $P2_1/c$, $Z=4$ | ab initio | −2.6 | −0.4 | 2.9 | | 1.0 | | 0.020 | 3.30 | −19.47 | 0.0 | 1 |
| | ab initio + VMC | −2.7 | −0.7 | 2.8 | | 1.1 | | 0.019 | 3.27 | −19.80 | 0.0 | 1 |
| TIMCHX | exptl | 8.56 | 10.66 | 6.57 | 90.0 | 107.7 | 90.0 | | | | | |
| $P2_1/c$, $Z=4$ | ab initio | 4.7 | 0.9 | 0.5 | | −0.2 | | 0.009 | 3.05 | −21.91 | 0.12 | 3 |
| | ab initio + VMC | 4.4 | 0.7 | 0.3 | | −0.1 | | 0.008 | 3.00 | −22.28 | 0.0 | 1 |
| piperidine | exptl | 8.69 | 5.26 | 12.01 | 90.0 | 96.8 | 90.0 | | | | | |
| $P2_1/c$, $Z=4$ | ab initio | −0.3 | 1.9 | 0.2 | | 0.1 | | 0.008 | 1.04 | −11.44 | 0.0 | 1 |
| | ab initio + VMC | −0.4 | 1.7 | 0.3 | | 0.8 | | 0.010 | 1.11 | −11.57 | 0.0 | 1 |
| piperazine | exptl | 6.01 | 5.19 | 8.41 | 90.0 | 108.3 | 90.0 | | | −17.70 | | |
| $P2_1/n$, $Z=2$ | ab initio | −0.9 | 2.0 | −3.1 | | 1.3 | | 0.006 | 0.52 | −16.03 | 0.0 | 1 |
| | ab initio + VMC | −0.6 | 1.7 | −3.4 | | 1.8 | | 0.007 | 0.86 | −16.32 | 0.0 | 1 |
| AMIPYR | exptl | 6.19 | 15.29 | 5.71 | 90.0 | 110.5 | 90.0 | | | −20.10 | | |
| *Cc*, $Z=4$ | ab initio + VMC | −0.2 | 1.5 | 1.5 | −0.2 | 5.1 | 0.3 | 0.034 | 2.86 | −17.90 | 0.0 | 2 |
| AMPYRE | exptl | 5.57 | 7.32 | 12.12 | 90.0 | 90.0 | 90.0 | | | −21.10 | | |
| $P2_12_12_1$, $Z=4$ | ab initio + VMC | 0.0 | −1.1 | 0.8 | | | | 0.012 | 1.82 | −20.06 | 0.21 | 2 |
| AMPYRD | exptl | 11.71 | 5.67 | 7.59 | 90.0 | 95.6 | 90.0 | | | −18.83 | | |
| $P2_1/c$, $Z=4$ | ab initio + VMC | 0.1 | −0.2 | 2.1 | | 3.9 | | 0.024 | 1.74 | −17.34 | 0.49 | >5 |
| BIGUAN01 | exptl | 9.55 | 5.06 | 9.99 | 90.0 | 102.9 | 90.0 | | | | | |
| $P2_1/n$, $Z=4$ | ab initio + VMC | −2.1 | −2.8 | −0.4 | | −0.4 | | 0.005 | 2.22 | −31.53 | 0.0 | 1 |
| TELPOP | exptl | 5.58 | 6.32 | 10.79 | 90.0 | 90.0 | 90.0 | | | −21.20 | | |
| *Pnma*, $Z=4$ | ab initio + VMC | 1.8 | 0.5 | −3.3 | | | | 0.003 | 2.56 | −18.46 | 0.0 | 1 |
| TELPIJ | exptl | 10.93 | 6.46 | 5.69 | 90.0 | 90.0 | 90.0 | | | | | |
| *Pnma*, $Z=4$ | ab initio + VMC | 1.9 | −0.4 | 1.9 | | | | 0.002 | 0.55 | −19.59 | 0.0 | 1 |
| AMPYRM01 | exptl | 14.86 | 10.88 | 5.63 | 90.0 | 90.0 | 90.0 | | | | | |
| *Pbca*, $Z=8$ | ab initio + VMC | 1.9 | 0.4 | 0.2 | | | | 0.009 | 4.19 | −20.82 | 0.39 | >5 |
| AMPYRZ | exptl | 6.2p | 10.24 | 7.37 | 90.0 | 93.1 | 90.0 | | | | | |
| $P2_1/c$, $Z=4$ | ab initio + VMC | −3.0 | −0.2 | 1.4 | | −1.3 | | 0.011 | 4.02 | −20.49 | 0.0 | 1 |

[a] Cambridge Structural Database reference code.[49] [b] In exptl lines, lattice constants in Å; in all other lines, percent deviation from the experimental values. [c] In exptl lines, unit cell angles in degrees; in all other lines, percent deviation from the experimental values. Since, the original space group symmetry of the experimental structure was preserved after the energy minimization for all structures considered in this work, we report deviations only for the unit cell angles that are not fixed by the space group. [d] Molecular translations from the experimental positions, in fractional coordinates. [e] Molecular rotations from the experimental positions, in degrees. [f] In exptl lines, the experimental sublimation enthalpy. In all other lines, lattice energy of the minimized experimental structure in kcal/mol; $E \approx -\Delta H_{\mathrm{subl}}$. [g] Difference ($E_{\mathrm{m.e.}} - E_{\mathrm{min}}$) between the energy of the minimized experimental structure and the energy of the lowest minimum found (kcal/mol) by global energy minimization. [h] Rank refers to the minimized experimental structure with respect to the global energy minimum.

containing both $sp^2$ and amine nitrogens. The molecules from the last group were used only for the final testing of the parameters.

To evaluate the ab initio-derived set of parameters, we carried out local energy minimizations for the first three groups of molecules 1−3 (14 molecules from IMAZOL06 to piperazine in Table 3). Results of these calculations showed that the experimental crystal structures of amines (group 3) are stable upon energy minimization (with small changes in the unit cell parameters), whereas very large deviations were obtained for the molecules from groups 1 and 2. This was not unexpected since it was shown[66] that electrostatic potentials and crystal structures of the molecules containing nitrogen with $sp^2$ hybridization and with no bonded hydrogens (for example, in azabenzenes) cannot be reproduced accurately using only partial charges located on atomic sites. It was also demonstrated[66] that addition of charges located close to the nitrogen atoms, and intended to model their lone-pair electrons, dramatically improves the accuracy of the results. Therefore, we included these additional lone-pair electron charge sites into our electrostatic model. The additional charges were located 0.4 Å from the corresponding nitrogen with quasitrigonal symmetry (the value

suggested by Williams[34]). Since the results of the local energy minimizations carried out using only the atomic charges were satisfactory in the case of amines, we did not include any additional charge sites near the amine nitrogen in our model. For the first two groups of molecules, local energy minimizations were repeated using the additional charge sites.

The results of these local energy minimizations of the experimental structures from IMAZOL06 to piperazine carried out using the ab initio-derived potentials are given in Table 3. Addition of the lone-pair charge sites significantly improved the accuracy of the results for the molecules from groups 1 and 2. All the experimental structures, except PURINE, NAPTYR11, and TIMCHX were reproduced very well. On average, the deviations of the unit cell parameters from the experimental values did not exceed 4%. The deviations were slightly larger for PURINE (4.2%) and TIMCHX (4.4%). A much larger deviation was obtained for the unit cell parameter *c* for NAPTYR11 (7.5%). The naptyridine molecule is nonplanar in the crystal, with the largest out-of-plane distances observed for hydrogens and nitrogen lone electron pairs.[67] The NAPTYR11 crystal structure is very sensitive to small changes in the molecular geometry, which may be a reason for the large deviation obtained from our calculations for this unit cell parameter.

Changes in the molecular orientations $\Delta\Omega$ (column 10, Table 3) were larger in the case of the molecules from groups 1 and 2 compared to those from the group 3.

In the case of PYRAZI01, the experimental crystal symmetry was partially lost as a result of the energy minimization (the unit cell angle, $\beta$, changed from 90 to 91.7°). The experimental electron density map available for the pyrazine crystal[68] shows that the maximum of the electron density corresponding to the lone pair electrons of nitrogen is located at 0.7 Å from the nitrogen. The experimental space group symmetry of the PYRAZI01 structure was retained after the local energy minimization was carried out using lone-pair charges located 0.7 Å, instead of 0.4 Å, from the nitrogen which had been used in all other cases.

The ab initio parameters provide very good agreement between the observed sublimation enthalpies and the calculated lattice energies. The largest discrepancy of 1.44 kcal/mol was found for PRMDIN. The good agreement with the experimental solid-state data found for amines and imidazoles using the ab initio parameters is in contrast with the results obtained for aliphatic and aromatic hydrocarbons,[4] for which the use of the ab initio-derived parameters led to quite significant underestimations of the lattice energies. It was shown previously[40] that medium level ab initio quantum mechanical methods usually underestimate the dispersion energy. On the other hand, our calculations showed that the relative dispersion energy contribution to the total lattice energy is much smaller in the case of more polar molecules when comparing hydrocarbons to amines and imidazoles. This may be the reason for the better agreement with the experimental data obtained for these classes of molecules.

As the next step, global energy optimizations were carried out for the first three groups of molecules. The ranks of the minimized experimental structures and the differences between their energies and the energies of the lowest minima found are shown in Table 3 (last two columns).

The minimized experimental structures of IMAZOL06, PURINE, PRMDIN01, DZTCDE, NAPTYR11, BAGFIY, piperidine, and piperazine were found as one of the four lowest-energy minima. For several molecules from the first two groups, namely BZDMAZ, TRAZOL03, VOBJEB, and PYRAZI01, many hypothetical structures with energies lower than the one of the minimized experimental structure were found during the global search. There may be several reasons for these results. First of all, ab initio parameters may need some refinement. Second, the potential form (6-exp-1) may be inadequate to predict the observed crystal structures of these molecules correctly. Last but not least, other, more stable crystal polymorphs may exist.

To check the first of these hypotheses, we carried out the VMC parameter refinement starting from the ab initio-derived values of the parameters. To decrease the correlation between the potential parameters, groups 2 and 3 were considered separately. First, we tried to optimize the potential parameters for the $sp^2$ and amine nitrogens separately. The molecules from groups 2 (PRMDIN01, DZTCDE, NAPTYR11, VOBJEB, and PIRAZI01) and 3 (PIWXEY, BAGFIY, TIMCHX, piperidine, and piperazine) were considered. The values of the parameters for all types of carbon and hydrogen were fixed at the values obtained earlier[4] and in this work (ab initio parameters for the amine hydrogen). The VMC run carried out using the molecules from the second group (and intended to find better parameters for $sp^2$ nitrogen) did not result in any improvement of the rank of the minimized experimental structures considered. On the other hand, the VMC refinement of the parameters for amine nitrogen (molecules from group 3) changed the rank of the minimized TIMCHX structure from 3 to 1. In fact, only the repulsion parameter, *B*, changed (it decreased from 54 694.5 to 51 218.6 kcal/mol) as a result of the optimization. The local and global energy minimizations carried out for the amine molecules using the refined parameters (Table 3, lines ab initio + VMC) showed that the parameter optimization indeed led to the improvement of the rank of the minimized TIMCHX structure without causing any negative changes in the results obtained for other amine molecules.

Next, we attempted to optimize the parameters of the hydrogen that is bonded to nitrogen (in amines and imidazoles). All molecules from group 3, plus IMAZOL06, BZDMAZ, and PURINE, were considered. Ab initio potential parameters for nitrogen with $sp^2$ hybridization and the optimized parameters for amine nitrogen were used. The only parameters allowed to vary during the refinement were *A* and *B* for the hydrogen that is bonded to nitrogen. The VMC refinement did not find any values of the hydrogen parameters that performed better than the ab initio values.

We also carried out a VMC run using only IMAZOL06, BZDMAZ, TRAZOL03, and PURINE to check if two types of nitrogens (with and without bonded hydrogen) instead of one are necessary to describe these crystals. Again, only the *A* and *B* parameters for nitrogens were allowed to vary. These calculations demonstrated that addition of another atom type did not lead to improvement of the results; therefore, the same atom type can be used for all nitrogens present in these molecules.

The results of the VMC refinement show that, despite our attempts to refine potential parameters, the experimental structures of BZDMAZ, TRAZOL03, VOBJEB, and PYRAZI01still have quite high rank. There are some indications[69] that different potential models, in particular, including better description of the electrostatic interactions, are necessary to reproduce these structures accurately. However, crystal structure prediction calculations have to be carried out to make sure that such models are able to predict the observed crystals' structures as the ones with lowest lattice energy.

The energy differences between the minimized experimental structures of BZDMAZ, TRAZOL03, VOBJEB, and PYRAZI01

and the corresponding lowest-energy minima found by global search are within the range of possible polymorphism (the energy differences between polymorphic structures are usually less than 1−2 kcal/mol[70]). Although there is no evidence in the literature of other polymorphs existing for these compounds, the possibility of polymorphism cannot be ruled out since there is no evidence that a systematic experimental study has been carried out in order to find new forms.

Small changes in the unit cell parameters and in the orientations ($\Delta\Omega$) and positions ($\Delta t$) of molecules obtained as a result of local energy minimizations for the first 14 molecules in Table 3 suggest that the new set of parameters (including ab initio-derived parameters plus the VMC refined parameters for amine nitrogen) is able to reproduce the experimental structures fairly well. The results of the global energy minimizations (Table 3, columns 11−13) also suggest that these parameters are accurate enough to be used for crystal structure prediction for the majority of amine and imidazole molecules.

To check if this conclusion holds for the molecules containing both the amine group and $sp^2$ nitrogen, we carried out local and global energy optimizations for the last 7 molecules from Table 3 (lines ab initio + VMC). The experimental structures were reproduced with good accuracy (the unit cell parameter deviations were less than 4%). All the minimized experimental structures, except AMPYRD and AMPYRM01) were found as the first or second lowest minimum of the lattice energy.

The set of parameters, including ab initio-derived parameters for $sp^2$ nitrogen and amine hydrogen plus the refined parameters for amine nitrogen, was considered as final. The final parameter values are given in Table 2.

**4.2. Amides and Carboxylic Acids.** *4.2.1. Derivation of Initial Values of Potential Parameters by ab Initio Calculations.* The parameters for atoms in amide and carboxylic groups were determined together in the next stage of the fitting procedure. The parameters for the following atom types were derived in this stage: amide nitrogen (N2), amide oxygen (O2), carboxyl oxygen connected to two atoms (OH2) and carboxyl carbon (C2). Initially, we assumed that the same atom type can be used for oxygen bonded to one atom in amide and carboxyl groups. Parameters for hydrogen (HO2) connected to oxygen (OH2) in a carboxylic group were frozen and set to values obtained previously for alcohols.[4] Parameters for hydrogen connected to an amide nitrogen (N2) were also frozen at this stage and set to values obtained earlier (see previous section) for methylamine and imidazole (HN1). The dimers of acetamide, *N*-methyl-acetamide, formic acid, and acetic acid molecules were used to obtain ab initio interaction energy surfaces, which were used together in the fitting procedure. The best resulting parameter set reproduced the ab initio interaction energy surfaces with standard deviations of 0.016 kcal/mol for *N*-methyl-acetamide, formic acid, and acetic acid and 0.032 kcal/mol for acetamide. The initial test of the resulting parameters on crystal structures of amides (*N*-methyl-acetamide and isopropylamide) gave acceptable reproducibility (average deviations below 7%) of the experimental crystal structures after local minimization. However, the resulting parameters, which were satisfactory enough for amides, failed completely in the test for carboxylic acids. None of the crystal structures of carboxylic acids (formic, acetic, and oxalic acids) used for the test was reproduced in local minimization of the experimental structure. Further trials to fit the parameters for atoms in the carboxylic group separately, using only the surfaces for acetic and formic acids, did not result in any good parameter sets (the minimized experimental crystal structures deviated strongly from the starting geometry).

As for amines and imidazoles, a strong correlation of parameters occurred also for amides and carboxylic acids, resulting in a value of 0.0 Å$^6$ kcal/mol for the parameter *A* for carbon C2. The shift of electron density away from the carbon atom, reflected in a large positive charge on atom C2 in amide and carboxyl groups can partly explain the tendency of the parameter *A*, which is supposed to describe mostly the dispersion and the induction part of the interaction energy, to have a small value. However, the value of 0.0 for the parameter *A*, implying an absence of electron density on the atom, is rather unphysical and is caused by the correlation with other parameters. We tried to refit the parameters in the same way as for amines and imidazoles, i.e., by freezing the *C* parameters at values calculated independently by Williams,[34] as well as by starting the fitting procedure from the values of *A* calculated from the Slater−Kirkwood formula. Despite the imposed restrictions, the *A* parameter of carbon still converged to the value 0.0. It is worth noting that the resulting parameter set with obviously unphysical values performed well as a complete consistent set in a test consisting of local and global energy minimizations carried out for a set of amides. We decided, however, to seek a more physical set of parameters for amides and carboxylic acids using the force minimization and the VMC methods.

*4.2.2. Evaluation and Refinement of Potential Parameters for Amides.* Thirteen amide molecules, shown in Figure 2, with known experimental crystal structures (Table 4, column 1) were selected for evaluation and refinement of potential parameters for amides. Since fitting ab initio interaction energies of amide dimers did not yield a physical set of potential parameters, we decided to use the force minimization and the VMC procedures. We considered a small set of amide molecules, including formamide (FORMAM02), izobutylacetamide (IBURAM) and succinamide (SUCCAM10), and benzamide (BZAMID02) to find a set of parameters which could be used as a good starting point for the VMC refinement. To reduce the effect of correlation between potential parameters, the *C* parameters for amide carbon, hydrogen, nitrogen, and oxygen were fixed at the values suggested by Williams.[34] The search for corresponding *A* values was restricted to the range of $A_{SK} \pm 100$ Å$^6$ kcal/mol, where $A_{SK}$ values were calculated[17] from the Slater−Kirkwood formula. Only the *A* and *B* parameters of amide hydrogen (HN2), carbon (C3), nitrogen (N2), and oxygen (O2) were allowed to vary during the calculations, with the rest of the parameters (*A*, *B*, and *C* for aliphatic and aromatic carbon and hydrogen) being fixed at the values obtained earlier.[4]

Several reasonable sets of parameters were found as a result of the force minimization. Local energy minimizations carried out using the best set of parameters found by force minimization [corresponding to the lowest value of the function *FF* (eq 15)] showed that the set of parameters can reproduce the experimental structures of isobutylamide, succinamide, and benzamide with good accuracy. The minimized experimental structures of all the molecules except succinamide were also the lowest-energy structures found by the global search. On the other hand, use of this parameter set for the local energy minimization led to significant changes of the experimental crystal structure of formamide. The minimized experimental structure of succinamide was found to have an energy 0.12 kcal/mol higher than the global minimum. Therefore, we decided to carry out the VMC refinement of the amide potential parameter starting from the values obtained by the force minimization. Our goal was to improve the parameters in such a way that the minimized experimental structure of SUCCAM10 would become the lowest-energy structure and the deviations of the unit cell

**TABLE 4: Results of Local and Global energy Minimizations Carried out for Amide Molecules (Figure 2) with the Potential Parameters Obtained by Force Minimization and the Subsequent VMC Refinement[a]**

| compound | parameter set | *a* | *b* | *c* | α | *β* | *γ* | Δ*t* | ΔΩ | *E* | Δ*E* | rank |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| FORMAM02 | exptl | 3.60 | 9.04 | 6.99 | 90.0 | 100.5 | 90.0 | | | 17.30 | | |
| $P2_1/n$, $Z=4$ | | −0.2 | 9.1 | −10.8 | | −2.8 | | 0.042 | 15.5 | −16.55 | 0.0 | 1 |
| ACEMID03 | exptl | 7.91 | 7.91 | 7.91 | 93.3 | 93.3 | 93.3 | | | 17.0 | | |
| $R3c$, $Z=18$ | | 0.5 | 0.5 | 0.5 | 0.4 | 0.4 | 0.4 | 0.019 | 2.59 | −17.84 | | |
| IBURAM | exptl | 10.36 | 5.99 | 9.66 | 90.0 | 108.1 | 90.0 | | | | | |
| $P2_1/c$, $Z=4$ | | −1.7 | −0.6 | 1.3 | | 4.7 | | 0.044 | 3.16 | −17.48 | 0.0 | 1 |
| SUCCAM10 | exptl | 6.93 | 7.99 | 9.88 | 90.0 | 102.5 | 90.0 | | | | | |
| $C2/c$, $Z=4$ | | −1.9 | −0.4 | 3.7 | | 0.4 | | 0.009 | 2.39 | −33.23 | 0.09 | >5 |
| ADIPAM11 | exptl | 5.11 | 5.57 | 7.05 | 69.6 | 87.1 | 75.5 | | | | | |
| $P\bar{1}$, $Z=1$ | | 0.5 | −0.2 | 0.9 | 0.1 | 3.2 | 6.2 | 0.0 | 4.48 | −34.49 | 0.43 | >5 |
| ADIPAM10 | exptl | 6.89 | 5.15 | 10.67 | 90.0 | 111.0 | 90.0 | | | | | |
| $P2_1/c$, $Z=2$ | | 1.8 | −0.5 | 0.0 | | 0.6 | | 0.005 | 4.16 | −36.34 | 0.0 | 1 |
| BZAMID02 | exptl | 5.53 | 5.03 | 21.34 | 90.0 | 88.7 | 90.0 | | | 23.20 | | |
| $P2_1/c$, $Z=4$ | | −1.9 | 1.5 | 6.7 | | 0.0 | | 0.037 | 4.03 | −21.93 | 0.0 | 1 |
| FULSUA | exptl | 9.08 | 7.56 | 13.02 | 90.0 | 104.1 | 90.0 | | | | | |
| $P2_1/n$, $Z=4$ | | 1.0 | −1.1 | 1.2 | | 0.7 | | 0.005 | 1.24 | −32.94 | 0.0 | 1 |
| CAPRYL | exptl | 5.00 | 23.15 | 7.21 | 90.0 | 104.76 | 90.0 | | | | | |
| $Cc$, $Z=4$ | | 0.0 | 0.4 | 2.0 | | 2.5 | | 0.026 | 4.47 | −19.64 | 0.0 | 1 |
| BISMEV | exptl | 16.40 | 6.41 | 6.50 | 90.0 | 92.05 | 90.0 | | | | | |
| $P2_1/n$, $Z=4$ | | 0.3 | 3.7 | −2.4 | | 1.6 | | 0.018 | 3.92 | −28.09 | 1.07 | 9 |
| CLPRAL | exptl | 9.26 | 6.55 | 7.26 | 90.0 | 111.40 | 90.0 | | | | | |
| $P2_1$, $Z=2$ | | −1.0 | −0.6 | 2.02 | | −1.7 | | 0.037 | 2.26 | −26.39 | 0.0 | 1 |
| LPROGL | exptl | 9.66 | 5.87 | 13.06 | 90.0 | 90.0 | 90.0 | | | | | |
| $P2_12_12_1$, $Z=4$ | | 2.3 | −1.1 | −1.9 | | | | 0.012 | 1.94 | −27.66 | 0.0 | 1 |
| POXPAT | exptl | 6.15 | 13.50 | 8.20 | 90.0 | 96.32 | 90.0 | | | −27.50 | | |
| $P2_1/n$, $Z=4$ | | 1.1 | 2.2 | −1.6 | | −2.1 | | 0.041 | 4.13 | −26.26 | 0.0 | 1 |
| TIPVOD | exptl | 6.22 | 12.07 | 7.52 | 90.0 | 111.03 | 90.0 | | | | | |
| $P2_1/c$, $Z=4$ | | 0.8 | 1.5 | 3.0 | | 0.8 | | 0.019 | 1.35 | −14.49 | 0.0 | 1 |

[a] For symbols and units, see footnotes to Table 3.

**TABLE 5: Potential Parameters Obtained by Force Minimization and the Sebsequent VMC Refinement for Amides**

| atom type | *A*, Å$^6$ kcal/mol | *B*, kcal/mol | *C*, Å$^{-1}$ |
|---|---|---|---|
| amide hydrogen (HN2) | 11.31 | 269.3 | 3.50 |
| amide carbon (C3) | 728.45 | 67230.2 | 3.60 |
| amide nitrogen (N2) | 434.78 | 56066.0 | 3.48 |
| amide oxygen (O2) | 310.15 | 36409.9 | 3.80 |

parameters for FORMAM02 would decrease. Sets of low energy minima found for formamide, isobutylamide, succinamide, and benzamide by the global energy optimization, were used during the refinement.

The VMC refinement runs carried out at different temperatures produced similar results, i.e., the only noticeable improvement was achieved for succinamide. For this molecule, the energy difference between the experimental structure and the global minimum found for the initial parameter set (from the force minimization) became negative, that is the experimental structure of succinamide minimized with the new set of parameters had the lowest energy compared to the all hypothetical structures considered. No significant decrease of the function $R_i$ (eq 10) for formamide was achieved. The final values of the parameters for the amide group are presented in Table 5.

To make sure that the VMC refinement indeed produced a better set of parameters for the atom types present in the amide group and that these parameters are transferable to other amide molecules, the refined set was used during global optimization runs carried out for all molecules from Table 4. We also carried out local energy minimizations of the corresponding experimental structures. The results of these calculations are given in Table 4. The local minimizations of the lattice energies, starting from the experimental structures, showed that the new parameters are able to reproduce not only the observed structural parameters well (e.g., unit cell parameters and rotations and translations of the molecules within the unit cell) but also the experimental values of the sublimation enthalpies. For all crystal structures except FORMAM02 and BZAMID02, the deviations of the unit cell parameters were less than 4%. The deviations of the lattice energies from the corresponding sublimation enthalpies did not exceed 1.3 kcal/mol. The data from Table 4 (columns Δ*t* and ΔΩ) show that the positions and orientations of the molecules in the unit cells, obtained by energy minimization, do not deviate significantly from the initial experimental values. The only molecule for which the results were unsatisfactory was formamide. For this molecule, local energy minimization led to deviations in the unit cell parameters as large as 10.8% and caused a significant change (15.5°) in the orientations of the molecules in the unit cell.

The results of the global energy minimizations demonstrated that the refined parameters are quite accurate since the minimized experimental structures for all the amide molecules considered in this work (the only exceptions being the triclinic crystal form of adipamide, ADIPAM11, SUCCAM10, and BISMEV) were found as global minima. In the case of succinamide, the VMC refinement slightly decreased the energy gap between the minimized experimental structure and the global energy minimum, but there are still several hypothetical structures with lower energies located by the global search.

On the whole, the set of parameters obtained for amides provides satisfactory results. The main deficiency of the parameter set was the low accuracy in reproducing the crystal structure of formamide. The experimental structure of formamide is not very stable upon energy minimization, i.e., the deviations in the unit cell edges are large (up to 10.8%). There may be several different reasons for these results. First of all, in our calculations we assumed that the same nonbonded parameters can be used for the hydrogen atom bonded to carbon in formamide and the hydrogen atoms bonded to aliphatic carbon (for example in the methyl group of acetamide). However, those two types of carbon are chemically very different; thus, an

**TABLE 6: Results of Local and Global Energy Minimizations Carried out for the Carboxylic Acid Molecules (Figure 3) Used for Derivation of Potential Parameters[a]**

| compound | parameter set | *a* | *b* | *c* | α | *β* | *γ* | Δ*t* | ΔΩ | ***E*** | Δ***E*** | rank |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| FORMAC01 | exptl | 10.24 | 3.54 | 5.36 | 90.0 | 90.0 | 90.0 | | | 14.80 | | |
| *Pna*2$_1$, $Z = 4$ | set1 | 6.3 | −1.5 | −3.8 | | | | 0.021 | 10.8 | −14.21 | 0.24 | 4 |
| ACETAC03 | exptl | 13.15 | 3.92 | 5.76 | 90.0 | 90.0 | 90.0 | | | 16.10 | | |
| *Pna*2$_1$, $Z = 4$ | set1 | 8.6 | −0.9 | −6.5 | | | | 0.049 | 5.75 | −14.19 | 0.00 | 1 |
| MALNAC02 | exptl | 5.16 | 5.34 | 8.41 | 71.5 | 76.1 | 85.1 | | | 25.10 | | |
| (*β* form) P1, $Z = 2$ | set1 | 8.1 | −1.3 | −3.3 | −5.7 | −4.3 | −1.7 | 0.066 | 7.84 | −27.69 | | |
| SUCACB08 | exptl | 5.48 | 8.79 | 5.03 | 90.00 | 92.91 | 90.00 | | | 28.80 | | |
| (*β* form)*P*2$_1$/*c*, $Z = 2$ | set1 | −1.9 | 6.4 | 0.0 | | 0.5 | | 0.016 | 4.72 | −29.18 | 0.0 | 1 |
| GLURAC03 | exptl | 12.98 | 4.75 | 9.69 | 90.0 | 98.30 | 90.0 | | | 29.00 | | |
| *C*2/*c*, $Z = 4$ | set1 | −1.9 | 2.3 | 2.8 | | −0.5 | | 0.029 | 0.37 | −29.39 | 0.0 | 1 |
| BENZDC01 | exptl | 3.76 | 16.36 | 11.70 | 90.0 | 90.3 | 90.0 | | | | | |
| *P*2$_1$/*c*, $Z = 4$ | set1 | 4.6 | −2.5 | −0.1 | | −4.9 | | 0.120 | 4.35 | −32.05 | 0.0 | 1 |

[a] For symbols and units, see footnotes to Table 3.

additional type of hydrogen with different nonbonded parameters may be necessary to reproduce the experimental crystal structure of formamide with good accuracy. Second, comparison of our results with the results obtained using some popular all-atom potentials,[41] such as AMBER and DISCOVER, showed that our potentials reproduce the crystal structure of formamide with comparable accuracy. On the other hand, the significantly better results[30,61] obtained using a distributed multipole analysis (DMA) electrostatic model suggest that a more accurate description of electrostatic interactions may be required to reproduce the formamide crystal structure accurately.

Higher than average deviations in cell edges were also obtained in the case of benzamide. The minimized structure was expanded along the c-direction by 6.7%.

It should be mentioned that, initially, we had access to the experimental data only for the triclinic crystal form of adipamide (ADIPAM11). Global optimization runs carried out for the adipamide molecule using the initial and the final refined parameter sets always predicted the ADIPAM11 crystal structure as a higher-energy minimum (>5) with many hypothetical low-energy structures laying within a quite narrow energy window. These results suggested that adipamide may be able to form another polymorph which is more stable than the triclinic form. A search in the CSD revealed that such a form actually exists (the monoclinic form, ADIPAM10). The experimental molecular geometry from the ADIPAM10 crystal structure was used for the local and global energy minimizations with the refined parameters and was found as the global minimum. Unfortunately, the different molecular geometries used for prediction of ADIPAM10 and ADIPAM11 structures make it impossible to compare the energies of the two polymorphs. (In our calculations, we used the molecular geometries from the original paper.[56,57] The reported molecular conformations are very similar for both polymorphs. However, the standard bond lengths and valence angles used for solving the ADIPAM11 crystal structure from X-ray powder diffraction data are quite different (up to 5° for valence angles) from the corresponding experimental values for ADIPAM10.) Although there is no experimental information regarding the relative stabilities of the two crystal forms of adipamide, the fact that the experimental structure of the triclinic polymorph was always found as a higher-energy minimum with both potentials may suggest that this form is thermodynamically less stable than the monoclinic one. Additional calculations carried out with a flexible adipamide molecule could help to clarify this matter.

The new set of parameters reproduced the crystal structures of 13 amide molecules with very good accuracy. Moreover, it predicted the experimental crystal structures of 10 molecules as global minima of the lattice energy. We consider the parameter set obtained by the VMC refinement as final for the atom types present in the amide group (Table 5).

**TABLE 7: Potential Parameters for Carboxylic Acids**

| atom type | *A*, Å$^6$ kcal/mol | *B*, kcal/mol | *C*, Å$^{-1}$ |
|---|---|---|---|
| carboxyl hydrogen (HO2) | 13.43 | 360.2 | 3.73 |
| carboxyl carbon (C2) | 661.02 | 39788.8 | 3.41 |
| carboxyl oxygen bonded to two atoms (OH2) | 374.99 | 40979.5 | 3.54 |
| carboxyl oxygen bonded to one atom (O1) | 309.44 | 51363.0 | 3.99 |

*4.2.3. Evaluation and Refinement of Potential Parameters for Carboxylic Acids.* Since no reasonable sets of parameters for the atom types in the carboxyl group were obtained by using ab initio calculations, we decided to use the force minimization and the VMC optimization procedures. During both procedures, the parameters for all the atom types [namely, carboxyl hydrogen (HO2), carboxyl oxygens (O1 and OH2), carboxyl carbon (C2)] present in the carboxyl group were allowed to vary. First, the force minimization procedure was applied to find a reasonably good starting set of values of the parameters. The experimental crystal structures of six molecules (Table 6, Figure 3) were used to find a set of parameters corresponding to a lowest value of the function *FF* (eq 15). As a result, several sets, providing relatively small deviations of the unit cell parameters of the minimized experimental and experimental structures and lattice energies reasonably close to the experimental sublimation enthalpies, were found. The results of the local energy minimizations carried out with the best parameter set found (Table 7) are given in Table 6 ('set1' lines). The lattice energies were predicted correctly by the set1 potentials to within the expected experimental uncertainty. On the average, the unit cell parameters were reproduced reasonably close to the experimental values; however, the deviations in the unit cell parameters were larger than in the case of amines and amides. The largest deviations in the unit cell egdes were found in the case of formic, acetic, malonic, and succinic acids (6.3, 6.8, 8.1, and 6.4%, respectively). The local energy minimizations also led to a significant change (10.8°) in the orientation of the molecules in the crystal structures of formic acid.

To assess the accuracy of the parameters from set1 further, global search runs were carried out for all molecules in Table 6, except malonic acid. The global search results are presented in Table 6 (columns 11−13). The minimized experimental structures of acetic, succinic, glutaric, and isophthalic acids were found as global minima of the potential, while the minimized experimental structure of formic acid was minimum number 4.

Transferability of the nonbonded parameters from set1 to other carboxylic acids was assessed by carrying out local and

**TABLE 8: Results of Local and Global Energy Minimizations Carried out for the Carboxylic Acid Molecules (Figure 3) Used for Transferability Tests[a]**

| compound | parameter set | *a* | *b* | *c* | α | *β* | *γ* | Δ*t* | ΔΩ | *E* | Δ*E* | rank |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PRONAC | exptl | 4.04 | 9.06 | 11.00 | 90.0 | 91.3 | 90.0 | | | 17.70 | | |
| $P2_1/c$, $Z=4$ | set1 | 0.4 | 0.1 | 2.1 | | 1.8 | | 0.039 | 6.69 | −15.24 | 0.66 | 10 |
| VALRAC | exptl | 5.55 | 9.66 | 11.34 | 90.0 | 101.8 | 90.0 | | | | | |
| $P2_1/c$, $Z=4$ | set1 | 1.9 | 0.5 | 0.0 | | −0.9 | | 0.030 | 5.61 | −18.18 | 0.00 | 1 |
| PIMELA05 | exptl | 5.65 | 9.65 | 16.02 | 90.0 | 108.0 | 90.0 | | | | | |
| $P2_1/c$, $Z=4$ | set1 | 0.1 | −0.2 | −1.5 | | 1.8 | | 0.030 | 2.99 | −33.34 | 0.00 | 1 |
| NAPOAC01 | exptl | 6.91 | 3.84 | 30.96 | 90.0 | 92.0 | 90.0 | | | 26.39 | | |
| $P2_1/c$, $Z=4$ | set1 | 1.4 | 4.6 | −3.1 | | 2.3 | | 0.130 | 7.8 | −23.22 | 0.65 | 5 |
| RESJUU | exptl | 20.53 | 6.14 | 8.60 | 90.0 | 90.1 | 90.0 | | | | | |
| $P2_1/n$, $Z=4$ | set1 | 0.3 | 3.2 | −0.8 | | 0.7 | | 0.018 | 2.59 | −28.81 | 0.00 | 1 |
| JUBBIR | exptl | 5.09 | 19.22 | 9.55 | 90.0 | 93.8 | 90.0 | | | | | |
| $C2/c$, $Z=4$ | set1 | 2.8 | −0.1 | −1.2 | | 1.3 | | 0.015 | 3.11 | −36.41 | 0.00 | 1 |

[a] For symbols and units, see footnotes to Table 3.

**TABLE 9: Results Obtained for Acetic Acid by Different Authors**

| reference | force field | Δ*a*[a] | Δ*b*[a] | Δ*c*[a] | Δ*E*[b] | rank[c] |
|---|---|---|---|---|---|---|
| Mooij et al.[72] | GROMOS | 8.5 | 0.4 | −17.7 | 0.22 | 11 |
| | AMBER | 8.4 | −2.4 | −8.7 | 0.22 | 3 |
| | DMA | 1.6 | −0.6 | 3.9 | 0.24 | 7 |
| | Dreiding | 9.9 | 3.7 | −3.8 | 0.24 | 11 |
| Payne et al.[71] | Dreiding | 8.9 | 1.9 | −7.3 | 0.22 | 6 |
| van Eijck[73] | OPLS-AC | | 8.7 | | 0.14 | 17 |
| Hagler et al.[59] | DISCOVER (1979) | | | | | |
| | 9−6−1 | 8.1 | −1.1 | −4.5 | − | − |
| | 12−6−1 | 11.1 | 2.2 | −13.1 | − | − |
| Williams [33] | W99 | 7.1 | −3.0 | −7.1 | − | − |
| this work | | 8.6 | −0.9 | −6.5 | 0.0 | 1 |

[a] Percent deviations of the lattice constants from the experimental values. [b] Difference ($E_{m.e.} - E_{min}$) between the energy of the minimized experimental structure and the energy of the lowest minimum found (kcal/mol); the dashes indicate that global energy minimizations were not carried out by these authors. [c] Rank refers to the minimized experimental structure with respect to the global energy minimum.

global lattice-energy minimizations for the six additional molecules (Figure 3) given in Table 8. In general, the potential parameters perform well. Average deviations of unit cell parameters from the experimental values did not exceed 2.9% (NAPOAC01). set1 predicted the lattice energies to be slightly lower than the experimental sublimation enthalpies. The minimized experimental crystal structures of all molecules, except propionic (PRONAC) and 1-naphthoic (NAPOAC01) acids, were found as global minima of the potential. In the case of propionic and 1-naphthoic acids, the energy differences between the minimized experimental structures and the corresponding global minima are not very large (0.4−0.6 kcal/mol). Nevertheless, several hypothetical low-energy structures were found in this narrow energy range for both potentials, which may suggest that there is a strong tendency to polymorphism for these molecules.

*4.2.4. Comparison with Other Works.* A number of works[60,71−74] has appeared which focus on predicting (or generating) possible crystal structures for carboxylic acids. These works provide valuable information about the types of crystal packings preferred by different molecules. They also enable one to evaluate the quality of different force fields on the basis of not only the accuracy with which they reproduce the experimental structures but also their ability to predict the experimental structure as one of the lowest-energy minima. In Tables 9 and 10, we present the compilation of the results of local and global (when available) energy minimizations obtained for some carboxylic acids by different authors[60,71−74] along with the results reported in this work.

**TABLE 10: Results Obtained for Some Carboxylic Acids by Different Authors**

| reference | force field | Δ*a*[a] | Δ*b*[a] | Δ*c*[a] | Δ*θ*[b] | Δ*E*[c] | rank[d] |
|---|---|---|---|---|---|---|---|
| | | | Formic Acid | | | | |
| Pillardy et al.[41] | DISCOVER | 4.4 | −7.2 | 4.2 | − | 0.32 | − |
| | AMBER | 2.9 | −1.2 | −3.5 | − | 0.31 | − |
| van Eijck[73] | OPLS-AC | | 5.4 | | − | 1.55 | 218 |
| Beyer et al.[74] | DMA | −0.3 | 0.7 | 0.5 | − | 0.07 | |
| Hagler et al.[59] | DISCOVER 9−6−1 | 3.1 | 1.3 | 3.1 | − | − | − |
| | DISCOVER 12−6−1 | −2.5 | 3.6 | 0.3 | − | − | − |
| Williams[33] | W99 | 4.9 | 0.7 | −4.2 | − | − | − |
| This work | | 6.3 | −1.5 | −3.8 | − | 0.949 | 4 |
| | | | Malonic Acid (*β* form) | | | | |
| van Eijck[73] | OPLS-AC | | 2.1 | | 4.7 | 1.51 | 50 |
| Williams[33] | W99 | 3.3 | −2.2 | −4.7 | 1.0 | − | − |
| This work | | 8.1 | −1.3 | −3.2 | 3.7 | − | − |
| | | | Succinic Acid (*β* form) | | | | |
| van Eijck[73] | OPLS-AC | | 2.5 | | 3.1 | 0.26 | 5 |
| This work | | −1.9 | 6.4 | 0.0 | 0.5 | 0.0 | 1 |
| | | | Isophthalic Acid | | | | |
| Williams[33] | W99 | 2.9 | −3.8 | −1.7 | −5.4 | − | − |
| This work | | 4.6 | −2.5 | −0.1 | −4.9 | 0.0 | 1 |
| | | | Propionic Acid | | | | |
| van Eijck[73] | OPLS-AC | | 1.7 | | 0.4 | 0.74 | 142 |
| Hagler et al.[59] | DISCOVER 9−6−1 | −13.4 | −0.4 | 0.2 | −0.5 | − | − |
| | DISCOVER 12−6−1 | −0.2 | −0.4 | 2.5 | −0.3 | − | − |
| This work | | 0.4 | 0.1 | 2.1 | 1.8 | 0.66 | 10 |
| | | | Valeric Acid | | | | |
| Hagler et al.[59] | DISCOVER 9−6−1 | 0.0 | −3.0 | −0.7 | 0.7 | − | − |
| | DISCOVER 12−6−1 | 0.1 | −3.0 | −0.6 | −1.2 | − | − |
| This work | | 1.9 | 0.5 | 0.0 | −0.9 | 0.00 | 1 |

[a] Percent deviations of the lattice constants from the experimental values. [b] Average percent deviation of the unit cell angles from the experimental values. [c] Difference ($E_{m.e.} - E_{min}$) between the energy of the minimized experimental structure and the energy of the lowest minimum found (kcal/mol); the dashes indicate that global energy minimizations were not carried out by these authors. [d] Rank refers to the minimized experimental structure with respect to the global energy minimum.

The crystal structure of acetic acid has probably been the most popular object of theoretical study among all the carboxylic acids. Possible crystal structures of acetic acid were generated[72] in eight space groups using three types of force fields: a united-atom (GROMOS), an all-atom (AMBER and Dreiding), and a potential[30,72] that makes use of a DMA besides the regular 6-exp dispersion−repulsion terms to describe the electrostatic interactions. Payne at el.[71] generated possible crystal structures of acetic

acid using the Dreiding force field. Crystal structure prediction, based on the use of space groups with two independent molecules in the asymmetric unit, was also carried out[73] with the OPLS force field. The authors of the DISCOVER[20,31] and W99[32] force fields used local energy minimization of the crystal structure of acetic acid to assess the quality of their potentials.

All together, the results of these works demonstrate that all the force fields, except DMA, give similar results when used for local energy minimization of the experimental crystal structure of acetic acid, i.e., energy minimization always led to significant changes in the unit cell parameters. On the other hand, the DMA force field is superior to the other force fields in maintaining the experimentally correct structure. For all the force fields used for predicting possible crystal structures, the experimental structure was found among a few lowest-energy minima not more than 0.5 kcal/mol above the global minimum. A large number of energetically feasible structures were found for acetic acid, and the qualitative packing types (dimers and catemers) are all stable for the all force fields considered.

The results presented in this work are in agreement with above conclusions. The final set of parameters for carboxylic acids (Table 7) and the popular force fields (AMBER, Dreiding, and OPLS) reproduce the experimental structure of acetic acid with comparable accuracy. However, our potential found the experimental structure of acetic acid as a global minimum. Unfortunately, this result alone does not allow us to draw any conclusions regarding the accuracy of our potential, since the energy differences reported in Table 9 are of the same order of magnitude as the expected uncertainties of the potentials.

In general, the data in Table 9 show that the force fields that include a point-charge electrostatic model reproduce the experimental crystal structure of acetic acid significantly worse than the DMA-based potential. This suggests that an accurate description of the anisotropy of the electrostatic interactions is very important in predicting the crystal structure of this molecule.

The crystal structure of formic acid has been used by several authors[33,41,59,73,74] to develop and evaluate the quality of potentials. The results of local energy minimizations carried out with different force fields are given in Table 10. Although the deviations of the unit cell parameters are smaller than in the case of acetic acid, the difference between the results obtained with a simple point charge model (DISCOVER, AMBER, OPLS-AC, W99, and set1 of this work) and the more accurate DMA electrostatic model is clearly visible. None of the force fields predicted the experimental structure of formic acid as a global minimum; however, it was only 0.07 kcal/mol higher in the case of the DMA potential. The experimental structure had an especially high rank (218)[73] in the case of the OPLS force field. set1 potentials predicted the experimental structure as minimum number 4.

The crystal structures of other carboxylic acids have been less well studied; however, results of local and global energy minimizations for propionic, malonic, isophthalic, succinic, and valeric acids can be found in the literature.[33,59,73] Our potentials and other force fields, such as OPLS, W99, and DISCOVER, reproduce the experimental crystal structures of these molecules fairly well. The results of crystal structure predictions carried out for malonic, succinic, and propionic acids, based on the use of space groups with two independent molecules in the asymmetric unit, were reported.[73] For the OPLS force field used,[73] as well as for our potentials (this work), the experimental crystal structure of propionic acid was quite high in energy (0.74 and 0.66 kcal/mol above the global minimum, respectively) and many hypothetical low-energy structures were found. In the case of the $\beta$-form of succinic acid, the experimental crystal structure was found as minimum number 5 for the OPLS force field, while it is the global minimum for our potential.

**TABLE 11: Final Values of the Potential Parameters**

| atom type | description | $A$, Å$^6$ kcal/mol | $B$, kcal/mol | $C$, Å$^{-1}$ |
|---|---|---|---|---|
| HC | hydrogen bonded to aliphatic carbon | 13.89 | 1567.7 | 3.81 |
| HA | hydrogen bonded to aromatic carbon | 41.75 | 3392.1 | 3.88 |
| HO1 | hydrogen in alcohol group | 3.25 | 137.8 | 3.69 |
| HO2 | hydrogen in carboxyl group | 13.43 | 360.2 | 3.73 |
| HN1 | hydrogen bonded to amine or sp$^2$ nitrogen | 4.56 | 2376.8 | 4.75 |
| HN2 | hydrogen in amide group | 11.31 | 269.3 | 3.50 |
| C1 | aliphatic carbon | 650.65 | 30869.1 | 3.16 |
| CA | aromatic carbon | 445.65 | 50360.5 | 3.39 |
| C2 | carbon in carboxyl group | 661.02 | 39788.8 | 3.41 |
| C3 | carbon in amide group | 728.45 | 67230.2 | 3.60 |
| N1 | aliphatic nitrogen (amines) | 769.30 | 51218.6 | 3.48 |
| NA | sp$^2$ nitrogen | 475.30 | 25836.2 | 3.48 |
| N2 | amide nitrogen | 434.80 | 56066.0 | 3.48 |
| OH1 | alcoholic oxygen | 332.72 | 20231.9 | 3.46 |
| OH2 | carboxyl oxygen bonded to two atoms | 374.99 | 40979.5 | 3.54 |
| O1 | carboxyl oxygen bonded to one atom | 309.44 | 51363.0 | 3.99 |
| O2 | amide oxygen | 310.15 | 36409.9 | 3.80 |

A complete comparison of the results obtained for carboxylic acids in this paper with those of other authors is not always possible because we assumed the molecules to be rigid, whereas flexible molecules were used in the other works[71−73] for generating possible crystal structures. The observed crystal structure and molecular conformation are the result of a balance between intra- and intermolecular interactions; therefore, to evaluate the accuracy of a given potential, the intramolecular degrees of freedom (especially the torsional angles) have to be taken into account. We are currently working on deriving parameters of a torsional potential and are planning to carry out crystal structure prediction calculations for the molecules considered in this work using flexible molecular geometries. Since, our goal is to develop a general all-atom force field for crystal and protein structure prediction, the future work will also include calculations of crystal structures of amino acids and small peptides to evaluate the applicability of the resulting nonbonded parameters to biomolecular systems.

## Conclusions

This paper presents the continuation of our work on the development of nonbonded parameters for a general all-atom force field for crystal and protein structure prediction. Here, the procedure for deriving nonbonded potential parameters developed earlier[4,5] was employed to obtain parameters for the atom types present in amine, imidazole, amide, and carboxyl groups. The method was successful in finding accurate sets of parameters for all groups of molecules considered, i.e., those satisfying all the requirements discussed in the Introduction. The final parameters for all the atom types considered in this and the previous work[4] are given in Table 11.

For amines and imidazoles, accurate parameters were obtained after the first step of the procedure, i.e., as a result of fitting to ab initio interaction energies of dimers. In the case of amides and carboxylic acids, force minimization was used to find a starting set of parameters for the VMC refinement.

On average, the experimental structures are well reproduced by the final potentials. Average deviations of the unit cell parameters were less than 4%, whereas the lattice energies agreed with the experimental sublimation enthalpies within the experimental error (2 kcal/mol). The crystal structures of the smaller molecules in the series of amides and carboxylic acids (namely, formamide, formic acid, and acetic acid) were not very stable upon energy minimization carried out with the new potential parameters. This result, considered together with the data reported in refs 33, 41, 71, 72, and 74 (see the discussion in the last subsection of Results and Discussion), suggests that the molecular packings of these molecules are very sensitive to the potentials used, and more accurate potential models, for example, those employing the distributed multipole analysis to describe electrostatic interactions, are required to reproduce observed crystal structures accurately.

The final nonbonded parameters were able to predict the most stable experimental structures for all groups of molecules as the global or one of the three lowest-energy minima. The higher rank of the experimental structure and large number of hypothetical low-energy structures found for some molecules may indicate that a more accurate potential model should be used or/and other low-energy polymorphs can exist for these molecules.

The potential parameters are shown to be well transferable to other molecules from the same families.

Comparison with the results obtained by other authors for the same molecules using some popular force fields (including the 6-exp or "6−12" potential functions with a point-charge electrostatic model) demonstrated that our parameters provide comparable or better accuracy in reproducing the experimental structures.

It was shown that the nonbonded parameters reported in this and our previous paper[4] are accurate and can be used for predicting crystal structures of small organic molecules on the basis of a rigid-body approximation. Ab initio prediction of crystal structures for flexible molecules would require taking into account at least torsional degrees of freedom. We are currently working on deriving parameters of torsional potentials. To assess the accuracy of the final force field, including electrostatic, nonbonded, and torsional terms, we plan to carry out crystal structure calculations using flexible molecular geometries for the molecules considered in this work, as well as for some amino acids, di-, tri-, and tetrapeptides.

**Acknowledgment.** This research was supported by grants from the National Institutes of Health (GM-14312) and the National Science Foundation (MCB00-03722). This work was carried out using computational resources provided in part by: (a) the Cornell Theory Center, which receives funding from Cornell University, New York State, and members of the Theory Center's Corporate Partnership Program, (b) the National Science Foundation Terascale Computing System at the Pittsburgh Supercomputer Center, and (c) with our own array of 193 dual-processor PC computers.

**Supporting Information Available:** Atomic charges used in crystal calculations. This supplement contains all molecules used in this work and the atomic charges on each of them. This material is available free of charge via the Internet at http://pubs.acs.org.

## References and Notes

(1) Fifth Meeting on the Critical Assessment of Techniques for Protein Structure Prediction. *Proteins* **2003**, *53*, S6, 333.

(2) Lommerse, J. P. M.; Motherwell, W. D. S.; Ammon, H. L.; Dunitz, J. D.; Gavezzotti, A.; Hofmann, D. W. M.; Leusen, F. J. J.; Mooij, W. T. M.; Price, S. L.; Schweizer, B.; Schmidt, M. U.; van Eijck, B. P.; Verwer, P.; Williams, D. E. *Acta Crystallogr., Sect. B: Struct. Sci.* **2000** *56*, 697.

(3) Motherwell, W. D. S.; Ammon, H. L.; Dunitz, J. D.; Dzyabchenko A.; Erk, P.; Gavezzotti, A.; Hofmann, D. W. M.; Leusen, F. J. J.; Lommerse, J. P. M.; Mooij, W. T. M.; Price, S. L.; Scheraga, H.; Schweizer, B.; Schmidt, M. U.; van Eijck, B. P.; Verwer, P.; Williams, D. E. *Acta Crystallogr., Sect. B: Struct. Sci.* **2002**, *58*, 647.

(4) Arnautova, Y. A.; Jagielska, A.; Pillardy, J.; Scheraga, H. A. *J. Phys. Chem. B* **2003**, *107*, 7143.

(5) Arnautova, Y. A.; Pillardy, J.; Czaplewski, C.; Scheraga, H. A. *J. Phys. Chem. B* **2003**, *107*, 712.

(6) Williams, D. E.; Weller, R. R. *J. Am. Chem. Soc.* **1983**, *105*, 4143.

(7) Tsuzuki, S.; Uchimaru, T.; Tanabe, K. *J. Phys. Chem.* **1994**, *98*, 1830.

(8) Mooij, W. T. M.; van Duijneveldt, F. B.; van Duijneveldt-van de Rijdt, J. G. C. M.; van Eijck, B. P. *J. Phys. Chem.* A **1999**, *103*, 9872.

(9) Tsui, H. H. Y.; Price, S. L. *CrystEngComm.* **1999**, *1*, 24.

(10) Mooij, W. T. M.; van Eijck, B. P.; Kroon, J. *J. Phys. Chem.* A **1999**, *103*, 9883.

(11) (a) Maple, J. R.; Dinur, U.; Hagler, A. T. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 5350. (b) 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. (c) Hagler, A. T.; Ewig, C. S. *Comput. Phys. Commun.* **1994**, *84*, 131.

(12) Hill, J.-R. *J. Comput. Chem.* **1997**, *18*, 211.

(13) Bordner, A. J.; Cavasotto, C. N.; Abagyan, R. A. *J. Phys. Chem. B* **2003**, *107*, 9601.

(14) Scheraga, H. A. *Adv. Phys. Org. Chem.* **1968**, *6*, 103.

(15) Warshel, A.; Lifson, S. *J. Chem. Phys.* **1970**, *53*, 582.

(16) Hagler, A. T.; Huler, E.; Lifson, S. *J. Am. Chem. Soc.* **1974**, *96*, 5319.

(17) Momany, F. A.; Carruthers, L. M.; McGuire, R. F.; Scheraga, H. A. *J. Phys. Chem.* **1974**, *78*, 1595.

(18) Momany, F. A.; McGuire, R. F.; Burgess, A. W.; Scheraga, H. A. *J. Phys. Chem.* **1975**, *79*, 2361.

(19) Williams, D. E.; Starr, T. L. *Comput. Chem. (Oxford)* **1977**, *1*, 173.

(20) Lifson, S.; Hagler, A. T.; Dauber, P. *J. Am. Chem. Soc.* **1979**, *101*, 5111.

(21) Cox, S. R.; Hsu, L.-Y.; Williams, D. E. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **1981**, *37*, 293.

(22) Williams, D. E.; Cox, S. R. *Acta Crystallogr., Sect. B: Struct. Sci.* **1984**, *40*, 404.

(23) Williams, D. E.; Houpt, D. J. *Acta Crystallogr., Sect. B: Struct. Sci.* **1986**, *42*, 286.

(24) Rappé, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A., III; Skiff, W. M. *J. Am. Chem. Soc.* **1992**, *114*, 10024.

(25) Filippini, G.; Gavezzotti, A. *Acta Crystallogr., Sect. B: Struct. Sci.* **1993**, *49*, 868.

(26) (a) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269. (b) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E., III; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Radmer, R. J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. *AMBER 6*; University of California, San Francisco: San Francisco, 1999.

(27) Hwang, M. J.; Stockfisch, T. P.; Hagler, A. T. *J. Am. Chem. Soc.* **1994**, *116*, 2515.

(28) Sun. H.; Mumby, S. J.; Maple, J. R.; Hagler, A. T. *J. Am. Chem. Soc.* **1994**, *116*, 2978.

(29) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225.

(30) Coombes, D. S.; Price, S. L.; Willock, D. J.; Leslie, M. *J. Phys. Chem.* **1996**, *100*, 7352.

(31) Ewig, C. S.; Thacher, T. S.; Hagler, A. T. *J. Phys. Chem. B* **1999**, *103*, 6998.

(32) Williams, D. E. *J. Mol. Struct.* **1999**, *485*−*486*, 321.

(33) Williams, D. E. *J. Comput. Chem.* **2001**, *22*, 1.

(34) Williams, D. E. *J. Comput. Chem.* **2001**, *22*, 1154.

(35) Schuler, L. D.; Daura, X.; van Gunsteren, W. F. *J. Comput. Chem.* **2001**, *22*, 1205.

(36) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote. R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586.

(37) Yin, D.; MacKerell, A. D., Jr. *J. Comput. Chem.* **1998**, *19*, 334.

(38) Chałasiński, G.; Szczęśniak M. M. *Chem. Rev.* **1994**, *94*, 1723.

(39) Jeziorski, B.; Moszynski, R.; Szalewicz, K. *Chem. Rev.* **1994**, *94*, 1887.

(40) Tsuzuki, S.; Tanabe, K. *J. Phys. Chem.* **1991**, *95*, 2272.
(41) Pillardy, J.; Wawak, R. J.; Arnautova, Y. A.; Czaplewski, C.; Scheraga, H. A. *J. Am. Chem. Soc.* **2000**, *122*, 907.
(42) Ewald, P. *Ann. Phys.* **1921**, *64*, 253.
(43) Gibson, K. D.; Scheraga, H. A. *J. Phys. Chem.* **1995**, *99*, 3752.
(44) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D., Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J., Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.7, Gaussian, Inc.: Pittsburgh, PA, 1998.
(45) (a) Hohenberg, P.; Kohn, W. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1964**, *136*, 864. (b) Kohn, W.; Sham, L. J. *Phys. Rev. A: At., Mol., Opt. Phys.* **1965**, *140*, 1133.
(46) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.
(47) Gay, D. M. *ACM Trans. Math. Software* **1983**, *9*, 503.
(48) Pillardy, J.; Arnautova, Y. A.; Czaplewski, C.; Gibson, K. D.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 12351.
(49) Allen, F. H.; Kennard, O.; Taylor, R. *Acc. Chem. Res.* **1983**, *16*, 146.
(50) Leiserowitz, L. *Acta Crystallogr., Sect. B: Struct. Sci.* **1976**, *32*, 775.
(51) *The Weak Hydrogen Bond in Structural Chemistry and Biology*; Desiraju G. R., Steiner, T., Eds.; Oxford University Press: Oxford, **1999**; p 507.
(52) (a) Jeffrey, G. A.; Ruble, J. R. *Acta Crystallogr., Sect. B: Struct. Sci.* **1980**, *36*, 2292. (b)Zobel, D.; Luger, P.; Dreissig, W. *Acta Crystallogr., Sect. B: Struct. Sci.* **1992**, *48*, 837.
(53) Hamilton, W. C. *Acta Crystallogr.* **1965**, *18*, 866.
(54) (a) Penfold, B. R.; White, J. C. B. *Acta Crystallogr.* **1959**, *12*, 130. (b) Blake, C. C. F.; Small, R. W. H. *Acta Crystallogr., Sect. B: Struct. Sci.* **1972**, *28*, 2201. (c) Gao, Q.; Jeffrey, G. A.; Ruble, J. R.; McMullan, R. K. *Acta Crystallogr., Sect. B: Struct. Sci.* **1991**, *47*, 742.
(55) Müller, A. H. R. *Z. Phys. Chem.* **1914**, *86*, 177.
(56) Hospital, M.; Housty, J. *Acta Crystallogr.* **1966**, *20*, 626.
(57) Seaton, C. C.; Tremayne, M. *Chem. Commun.* **2002**, 880.
(58) Hagler, A. T.; Leiserowitz, L. *J. Am. Chem. Soc.* **1978**, *100*, 5879.
(59) Hagler, A. T.; Dauber, P.; Lifson, S. *J. Am. Chem. Soc.* **1979**, *101*, 5131.
(60) Berkovitch-Yellin, Z.; Leiserowitz, L. *J. Am. Chem. Soc.* **1982**, *104*, 4052.
(61) Mitchell, J. B. O.; Price, S. L. *J. Phys. Chem. A* **2000**, *104*, 10958.
(62) Allan, D. R.; Clark, S. J. *Phys. Rev. Lett.* **1999**, *82*, 3464.
(63) Allan, D. R., Clark, S. J. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1999**, *60*, 6328.
(64) Reference for the experimental heats of sublimation: NIST Chemistry Webbook. http://webbook.nist.gov/chemistry.
(65) Slater, J.; Kirkwood, J. G. *Phys. Rev.* **1931**, *37*, 682.
(66) Williams, D. E.; Cox, S. R. *Acta Crystallogr., Sect. B: Struct. Sci.* **1984**, *40*, 404.
(67) Dapporto, P.; Ghilardi, C. A.; Mealli, C.; Orlandini, A.; Pacinotti, S. *Acta Crystallogr., Sect C: Cryst. Struct. Commun.* **1984**, *40*, 891.
(68) de With, G.; Harkema, S.; Feil, D. *Acta Crystallogr., Sect. B: Struct. Sci.* **1976**, *32*, 3178.
(69) Day, G. M.; Price, S. L.; Leslie, M. *J. Phys. Chem. B* **2003**, *107*, 10919.
(70) Bernstein, J. *Polymorphism in Molecular Crystals*; Clarendon Press: Oxford, 2002; p 32.
(71) Payne, R. S.; Roberts, R. J.; Rowe, R. C.; Docherty, R. *J. Comput. Chem.* **1998**, *19*, 1.
(72) Mooij, W. T. M.; van Eijck, B. P.; Price, S. L.; Verwer, P.; Kroon, J. *J. Comput. Chem.* **1998**, *19*, 459.
(73) van Eijck, B. P. *J. Comput. Chem.* **2002**, *23*, 456.
(74) Beyer, T.; Price, S. L. *J. Phys. Chem. B* **2000**, *104*, 2647.