

The Impact of Atomic Dipoles and Quadrupoles on Calculated Crystal Structures and Sublimation Energies of Model Amide Compounds

Michael J. Dudek
Biomolecule Objects LLC
5411 Linda Vista Rd. #12
San Diego, CA 92110, U.S.A.

Manuscript prepared August 1995, never submitted for publication.

Abstract

The level of success of a molecular mechanics based method for predicting protein structure is critically dependent on the accuracy with which interaction energies are represented between the ubiquitous amide functional groups. The growing collection of experimental determinations of crystal structures and sublimation energies of model amide compounds provides a database with which the accuracies of amide potential energy functions can be evaluated. For some model amide compounds (for example formamide and acetamide), predictions of crystal structures and sublimation energies to within the accuracy of the experiments have not yet been achieved. This failure to achieve accurate predictions indicates errors in the representations of intermolecular energies. Components of the intermolecular energy include electrostatic, polarization, repulsion, and dispersion energies. Each of these components is a possible source of error. Most previous studies of amide crystal packing have used simple electrostatic models in which the charge density of a molecule is represented by partial atomic point charges located at the nuclei. These models are limited by the restriction that the molecular charge density be a superposition of spherically symmetric atomic charge densities. More accurate electrostatic models can be achieved through the use of anisotropic atomic charge densities and a truncated multipole expansion representation of electrostatic interaction energies between pairs of atomic charge densities. However, the significance of this increased accuracy in terms of its effect on differences between predicted and experimentally observed quantities remains a subject of controversy. For the purpose of contributing to this discussion, we have explored the sensitivity of the predicted crystal structures and sublimation energies of formamide, oxamide, and urea to the addition of atomic dipoles and quadrupoles. Several possible methods have been suggested for calculating atomic multipoles from molecular wavefunctions. To gain confidence that our conclusions concerning the significance of atomic dipoles and quadrupoles are not method dependent, results are presented and compared for three distinct methods of calculation: Force-Related, DMA, and Hirshfeld. For all three of these methods, the addition of atomic dipoles and quadrupoles has a significant impact on the calculated structure of this collection of systems, the largest observed change in unit cell lengths being greater than 1 Å for all methods. Also, for all of the methods, the addition of atomic dipoles and quadrupoles causes large adjustments to calculated electrostatic energies, the largest observed change being greater than 6.6 kcal/mol for all methods. The results of this study suggest that atomic dipoles and quadrupoles may be a necessary component for accurate prediction of structure and energies in biomolecular systems.

1 INTRODUCTION

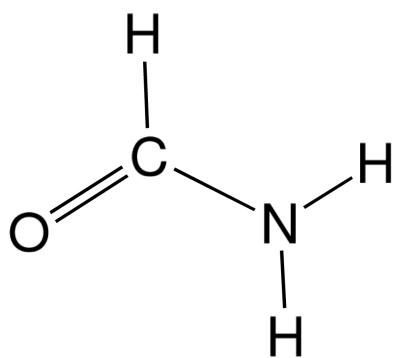
The level of success of a molecular mechanics based method for predicting protein structure is critically dependent on the accuracy with which interaction energies are represented between the ubiquitous

amide functional groups. Historically, the growing database of experimental determinations of crystal structures and sublimation energies of model amide compounds has been used, in combination with quantum mechanical and spectroscopic data, as a primary source of information in the construction of molecular mechanics models of amide functional groups.⁽¹⁻⁶⁾ Typically, some measure of distance between predicted and experimental crystal structures and sublimation energies is minimized with respect to some subset of model potential energy function parameters. The small number of independent degrees of freedom that specify structure in these systems (unit cell vectors, position and orientation of rigid molecules within the central unit cell) constitute a valuable property of the crystal data. As a consequence, the calculations that are needed to translate model energy functions into predicted quantities are simple and direct. Other valuable properties of the crystal data (considering as an alternative either liquid phase experimental data or gas phase *ab initio* data) include the high accuracy, the condensed phase nature, and the relatively larger ratio of data to independently adjustable intermolecular potential energy function parameters. Both a great strength and ultimately the primary limitation of the crystal data lies in the observed sensitivity of predicted structures and energies to the choice of model energy function parameters and forms. For some of the model amide compounds; for example formamide, acetamide, and glutaramide; predictions of crystal structures and sublimation energies to within the accuracy of the experiments have not yet been achieved. This failure to achieve accurate and dependable predictions indicates errors in the functions that have been used to model intermolecular energies. Most importantly, at the current stage in the evolution of functional forms and parameters, this failure provides us with a measure of accuracy that is sensitive to errors of the type that remain. Such a measure is critical to the development of more accurate models.

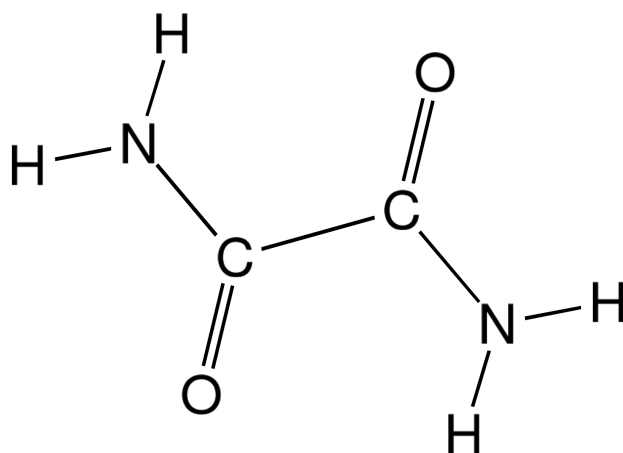
Components of the intermolecular energy include electrostatic, polarization, repulsion, and dispersion energies. Each of these components is a possible source of error. In this paper, we focus on the electrostatic component.

In essentially all energy functions that have been commonly used in simulations of biomolecular systems, the charge density of a molecule has been represented by partial atomic point charges located at the nuclei.⁽⁷⁾ Apart from effects related to the overlap of the electron clouds of different molecules, the use of partial atomic point charges at nuclei is equivalent to the assumption of spherically symmetric atomic charge densities. From both experiment and theory, it is known that the charge density of a molecule is more complex than a simple superposition of many spherically symmetric atomic charge densities.⁽⁸⁻¹³⁾ In particular, quantum mechanics predicts a buildup of electron density at the center of each covalent bond relative to a best-fit superposition of isotropic atomic charge densities. A balance between loss of accuracy and computational intractability can be achieved through the use of anisotropic atomic charge densities and a truncated multipole expansion representation of electrostatic interaction energies between pairs of atomic charge densities.⁽¹²⁾ The atomic multipoles of this more sophisticated electrostatic model can be chosen to reflect the charge density of a good *ab initio* wavefunction. Such models have been used previously in crystal packing calculations.⁽¹⁴⁻¹⁸⁾ Willock et al. have used Stone's distributed multipole analysis⁽¹⁹⁾ (DMA) to calculate atomic multipoles through hexadecapole, and have truncated their multipole expansion at the inverse fifth power of distance. Ritchie et al. have used Hirshfeld's method⁽²⁰⁾ to calculate atomic multipoles out to the quadrupole moment.

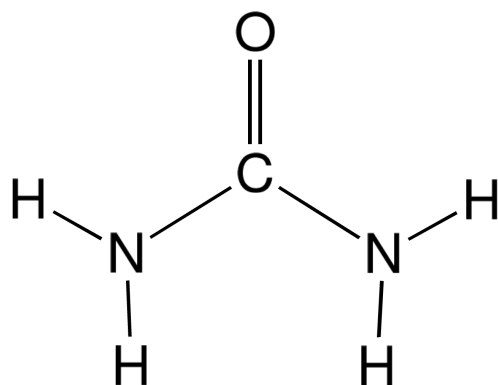
While it is generally accepted that more accurate electrostatic models can be achieved through the introduction of atomic dipoles, quadrupoles, and higher-order moments, the significance of this increased accuracy in terms of its effect on differences between predicted and experimentally observed quantities remains a subject of discussion. For the purpose of contributing to this discussion, we have explored the sensitivity of the predicted crystal structures and sublimation energies of model amide compounds to the addition of atomic dipoles and quadrupoles. In this paper, results are presented for three small planar molecules: formamide, oxamide, and urea. The structures of these molecules are shown in Figure 1. We note that the interactions that occur in these crystals are similar in nature to those that occur between amide groups in proteins. For each of these molecules, crystal structures and sublimation energies have been determined.⁽²¹⁻²⁹⁾ In addition, minimum energy structures have been calculated using a range of potential energy functions.^(2,3,5,6,18)



formamide



oxamide



urea

Figure 1: Molecular structures of formamide, oxamide, and urea.

Several possible methods have been suggested for calculating atomic multipoles from molecular wavefunctions.^(19,20,30-33) To increase confidence that our conclusions concerning the significance of atomic dipoles and quadrupoles are not dependent on the method by which the atomic multipoles were calculated, results are presented and compared for three distinct methods: Force-Related,⁽³⁰⁻³³⁾ DMA,⁽¹⁹⁾ and Hirshfeld.⁽²⁰⁾ Central to the logic of the conclusion forming process is the following convergence property of the Force-Related method. Accuracy at the point monopole level (before addition of point dipoles, quadrupoles, ...) has been shown to be as good as that of optimal potential-derived partial atomic point charge models.⁽³¹⁾ Offsetting this and other highly desirable properties of the Force-Related partitioning of the molecular charge density, rigorous application is limited to planar molecules.

The primary goal of this work was to generate data relevant to a decision concerning the necessity of including higher-order atomic multipoles for the attainment of accurate predictions of structure and energetics in amide crystals, and by association in biomolecular systems. This was accomplished by observing the magnitudes of changes to predicted structures and energies that result from the addition of atomic dipoles and quadrupoles, as opposed to the movement of these quantities relative to their experimental values. Because of the possibility of large errors in the functions that were used to represent the repulsion and polarization components of the total energy, the substitution of a more accurate representation of the electrostatic component is not guaranteed to cause predicted quantities to move closer to the experimental target. The attainment of more accurate amide potential energy functions, a larger goal that is expected to require simultaneous substitutions of more accurate representations of repulsion and possibly polarization, is beyond the scope of this work.

2 METHODS

The following sequence of calculations was carried out for formamide, oxamide, and urea.

A molecular geometry was selected. The positions of the heavy atoms were taken from the experimental crystal structure.^(21,25,27) The positions of the hydrogen atoms were obtained by initial placement in a standard position, followed by refinement by energy minimization, holding the positions of the heavy atoms fixed. This geometry was held fixed throughout all remaining calculations.

A wavefunction was calculated at the HF/6-31G* level for an isolated molecule using the *ab initio* molecular orbital program GAUSSIAN90.⁽³⁴⁾ Three sets of atomic multipoles were calculated from the wavefunction using the Force-Related,⁽³⁰⁻³³⁾ DMA,⁽¹⁹⁾ and Hirshfeld⁽²⁰⁾ methods. DMA atomic multipoles were calculated out to the seventh moment, and both Force-Related and Hirshfeld atomic multipoles were calculated out to quadrupole.

For each set of atomic multipoles, the energy of the crystal was minimized, starting from the experimental structure; using atomic monopoles, dipoles, and quadrupoles; and using monopoles alone. The repulsion +dispersion component of the total energy was represented by a 9-6 functional form, with parameters taken from the **cff93** force field. The energy of the crystal was minimized with respect to the 9 cartesian components of the three unit cell vectors and, for each molecule in the unit cell other than the first molecule, 6 parameters that specify the position and orientation of that molecule relative to the first molecule. Thus the experimental symmetry of the crystal was not imposed. The individual molecules were treated as rigid bodies during the minimization. Interaction energies were calculated between a central unit cell and all unit cells within 16 Å of the central unit cell. Energy minimizations were accomplished with the use of a crystal packing program that was developed by the author. This program enables the use of atomic multipoles out to the seventh moment.

For each set of atomic multipoles, the electrostatic energy of the crystal, calculated at the experimentally observed geometry using atomic monopoles, dipoles, and quadrupoles, was decomposed into its moment-moment contributions. As a point of reference, the electrostatic energy of the experimental crystal structure was also calculated by carrying out the DMA expansion through the seventh moment where the convergence is almost complete. Again, all unit cells within 16 Å of the central unit cell were included.

Table I: Comparison of Predicted Crystal Structures ^{a, b, c} of Formamide Obtained Using Atomic Monopoles, Dipoles, and Quadrupoles from the Force-Related, DMA, and Hirshfeld Methods to Predicted Structures Obtained Using Monopoles Alone.

	Unit Cell Vectors					
	Force-Related		DMA		Hirshfeld	
	q	q, μ, Θ	q	q, μ, Θ	q	q, μ, Θ
a	3.56	3.66(.10)	4.00	3.64(-.36)	3.66	3.59(-.07)
b	9.70	9.84(.14)	9.45	9.54(.09)	9.06	9.69(.63)
c	7.62	7.29(-.33)	7.31	7.41(.10)	9.63	7.41(-2.22)
α	90.00	90.00(.00)	90.00	89.98(-.02)	90.00	90.00(.00)
β	124.66	125.74(1.08)	131.89	124.91(-6.98)	133.67	124.46(-9.21)
γ	90.11	90.02(-.09)	90.01	90.32(.31)	90.01	90.06(.05)

^aUnits of cell lengths and angles are angstroms and degrees, respectively.

^bThe repulsion +dispersion component of the energy is represented by the functional form and parameters of cff93.

^cAll unit cells within 16 Å of the central unit cell are included in the calculation.

3 RESULTS AND CONCLUSIONS

3.1 Changes in Predicted Structures Caused by Addition of Atomic Dipoles and Quadrupoles.

A comparison of the predicted crystal structures of formamide obtained using atomic monopoles, dipoles, and quadrupoles to predicted structures obtained using monopoles alone is given in Table I. Here, the unit cell vectors are specified by the lattice parameters a, b, c and α, β, γ .

At the monopole level, we expect that the Force-Related model should be more accurate than the DMA or Hirshfeld models. This is expected because, of the three methods, only the Force-Related method has the property that the molecular dipole is reproduced by the partial charges alone. In other words, the vector sum of the point dipoles is zero. Alternatively, using the language of convergence, this statement can be expressed as follows. At the monopole level, we expect that the Force-Related model should be more fully converged than the DMA or Hirshfeld models.

For all three of the methods, the addition of atomic dipoles and quadrupoles causes the predicted structure to change. For the Force-Related method, the observed shifts are about .3 Å in unit cell lengths and about 1° in unit cell angles. Consistent with our expectation that the Force-Related model should be more fully converged at the monopole level, the movements obtained using the DMA and Hirshfeld methods are larger.

The same data for oxamide is presented in Table II. For the Force-Related method, the observed changes are about .1 Å in unit cell lengths and about 1° in unit cell angles. Again the displacements are larger for the DMA and Hirshfeld methods.

The same data for urea is shown in Table III. Here, the changes of greater than 1 Å in unit cell lengths that result from use of the Force-Related method are large in comparison to those obtained for formamide and oxamide. Consistent with this result are two possible origins in terms of changes to the energy surface. One possibility is that the addition of atomic dipoles and quadrupoles causes some large shift in the position of the global minimum. Another possibility is that the energy surface of the monopole model has a local minimum near the position of the crystal structure. In this scenario, the addition of atomic dipoles and quadrupoles causes the disappearance of some small barrier that separates the local and global minima. To distinguish between these two possibilities, a second energy minimization was carried out for the Force-Related model using monopoles alone, this time starting from the minimum-energy structure that was obtained using atomic monopoles, dipoles, and quadrupoles. The final energy-minimized structure was unaffected by this alternative choice of the starting point. In addition, structures

Table II: Comparison of Predicted Crystal Structures ^a of Oxamide Obtained Using Atomic Monopoles, Dipoles, and Quadrupoles from the Force-Related, DMA, and Hirshfeld Methods to Predicted Structures Obtained Using Monopoles Alone.

	Unit Cell Vectors					
	Force-Related		DMA		Hirshfeld	
	q	q, μ, Θ	q	q, μ, Θ	q	q, μ, Θ
a	3.55	3.56(.01)	3.57	3.64(.07)	3.56	3.64(.08)
b	5.19	5.16(-.03)	5.00	5.13(.13)	5.39	5.14(-.25)
c	5.70	5.60(-.10)	5.53	5.51(-.02)	5.89	5.61(-.28)
α	81.81	82.83(1.02)	81.12	86.16(5.04)	79.34	82.54(3.20)
β	113.95	113.91(-.04)	114.33	113.28(-1.05)	116.88	113.82(-3.06)
γ	114.20	115.39(1.19)	117.31	115.41(-1.90)	115.08	115.77(.69)

^aSee footnotes of Table I.

Table III: Comparison of Predicted Crystal Structures ^a of Urea Obtained Using Atomic Monopoles, Dipoles, and Quadrupoles from the Force-Related, DMA, and Hirshfeld Methods to Predicted Structures Obtained Using Monopoles Alone.

	Unit Cell Vectors					
	Force-Related		DMA		Hirshfeld	
	q	q, μ, Θ	q	q, μ, Θ	q	q, μ, Θ
a	5.53	6.72(1.19)	5.32	6.81(1.49)	7.71	5.47(-2.24)
b	5.53	5.10(-.43)	5.32	5.01(-.31)	3.68	5.45(1.77)
c	4.79	3.67(-1.12)	4.69	3.68(-1.01)	5.01	4.78(-.23)
α	89.98	90.01(.03)	90.01	90.09(.08)	89.99	89.99(.00)
β	90.04	89.97(-.07)	89.91	89.96(.05)	90.02	90.02(.00)
γ	89.99	89.62(-.37)	89.98	89.81(-.17)	90.01	90.00(-.01)

^aSee footnotes of Table I.

that differ by greater than 1 Å in unit cell lengths were found to differ by only a few tenths of a kcal/mol in energy. These observations are consistent with a wide flat potential energy well that has only one minimum. Therefore, the first possibility, that the addition of atomic dipoles and quadrupoles has caused a large shift in the position of the global minimum, appears to be the correct interpretation.

3.2 Changes in Predicted Energies Caused by Addition of Atomic Dipoles and Quadrupoles.

The intermolecular electrostatic energy of the formamide crystal; calculated at the experimentally observed geometry using atomic monopoles, dipoles, and quadrupoles from the Force-Related, DMA, and Hirshfeld methods; and a decomposition of this energy into its moment–moment contributions, are presented in Table IV. For all three of the methods, the addition of atomic dipoles and quadrupoles causes large adjustments to the predicted energy. For the Force-Related method, dipole interactions contribute 2.6 kcal/mol to the electrostatic energy, which decreases from -14.2 kcal/mol to -16.8 kcal/mol, a change of roughly 20%. Again, consistent with our expectation that the Force-Related model should be more fully converged at the monopole level, the changes obtained using the DMA and Hirshfeld methods are even larger, with the latter method producing charge–dipole and dipole–dipole interactions as large as the coulomb term. Also clear from Table IV, the three electrostatic models do not agree with one another, yielding differences in predicted energies as large as 1.9 kcal/mol, although total energy varies

Table IV: Intermolecular Electrostatic Energy of the Formamide Crystal ^a Calculated at the Experimentally Observed Geometry Using Atomic Monopoles, Dipoles, and Quadrupoles from the Force-Related, DMA, and Hirshfeld Methods.

	Electrostatic Energy (kcal/mol) and Components		
	Force-Related	DMA	Hirshfeld
E_{qq}	-14.2705	-27.1490	-6.6592
$E_{q\mu}$	-2.5602	11.3479	-6.8949
$E_{\mu\mu}$	-.0580	-1.1934	-1.4389
$E_{q\Theta}$	.1932	-.9913	-.6669
$E_{\mu\Theta}$	-.1336	-.2193	-.5347
$E_{\Theta\Theta}$	-.0410	.0420	-.0346
total	-16.8701	-18.1631	-16.2292
converged		-16.8	

^aAll unit cells within 16 Å of the central unit cell are included in the calculation.

less than individual components. This implies that some or all of the models have not fully converged at the quadrupole level. Also included in Table IV is an estimate of the intermolecular electrostatic energy of the crystal calculated at the experimentally observed geometry using the charge density of the wavefunction. This estimate was obtained by carrying out the DMA expansion through the seventh moment where it has converged to at least one decimal place. Apart from lack of complete convergence, it is exact to the extent that the energetic effects of penetration (the overlap of the electron densities of two distinct molecules) can be neglected.^(35,36) If errors due to penetration are either negligible or independent of method, then the different multipole expansions should converge to the same limit, independent of the method used to calculate atomic multipoles.

The same data is presented for oxamide and urea in Tables V and VI. For all three of the methods, quadrupole interactions can become substantial. From Table VI, it can be seen that the cumulative contribution of the Force-Related quadrupoles, which amounts to only a few tenths of a kcal/mol for the formamide and oxamide crystals, is 3.1 kcal/mol for the urea crystal, although this is largely balanced by the dipole contributions. Comparisons of electrostatic energies calculated using atomic monopoles, dipoles, and quadrupoles to the limiting values of the DMA expansions suggest that, at the quadrupole level, none of the models is fully converged. In addition, these comparisons suggest that, at the quadrupole level, the extent of convergence is similar for the three models. For the Force-Related, DMA, and Hirshfeld models, errors introduced by the neglect of octopoles and beyond are as large as 1.2, 1.9, and 1.0 kcal/mol, respectively. For the DMA model, as is illustrated in Table VII for the formamide crystal, most of this error results from the neglect of monopole–octopole contributions.

3.3 Method Dependence of Atomic Multipoles.

A comparison of atomic monopoles, dipoles, and quadrupoles calculated for the amide functional groups of formamide, oxamide, and urea using the Force-Related, DMA, and Hirshfeld methods is given in Table VIII. The units of charge and distance are 1 e_cu and 1 Å, respectively. To facilitate a comparison of the atomic dipoles and quadrupoles of different molecules, all moments have been expressed with respect to the local coordinate systems of Figure 2. The primary result of Table VIII is the very large method dependence that was observed for the atomic multipoles of all three molecules. The existence of a wide range of possible choices for the atomic multipoles, all having similar convergence properties at large distances, is a consequence of the existence of more than one expansion site.

Figure 3 shows the atomic monopoles and dipoles of formamide calculated using the Force-Related, DMA, and Hirshfeld methods. These three sets of atomic multipoles differ from one another by large amounts. Yet, as presented in Table IX for formamide, at points just a few (9 to 15) tenths of an angstrom

Table V: Intermolecular Electrostatic Energy of the Oxamide Crystal ^a Calculated at the Experimentally Observed Geometry Using Atomic Monopoles, Dipoles, and Quadrupoles from the Force-Related, DMA, and Hirshfeld Methods.

	Electrostatic Energy (kcal/mol) and Components		
	Force-Related	DMA	Hirshfeld
E_{qq}	-19.0838	-40.3545	-8.6340
$E_{q\mu}$	-6.3137	15.4519	-10.2398
$E_{\mu\mu}$	-.4645	-1.2299	-2.7637
$E_{q\ominus}$	-.0623	-.6587	-1.6557
$E_{\mu\ominus}$	.0422	-.6228	-.6696
$E_{\ominus\ominus}$	.1350	.5480	.1193
total	-25.7472	-26.8661	-23.8435
converged		-24.9	

^aSee footnotes of Table IV.

Table VI: Intermolecular Electrostatic Energy of the Urea Crystal ^a Calculated at the Experimentally Observed Geometry Using Atomic Monopoles, Dipoles, and Quadrupoles from the Force-Related, DMA, and Hirshfeld Methods.

	Electrostatic Energy (kcal/mol) and Components		
	Force-Related	DMA	Hirshfeld
E_{qq}	-19.3983	-32.3033	-8.1209
$E_{q\mu}$	-3.6203	10.0665	-10.2471
$E_{\mu\mu}$	.0207	-.7159	-1.8527
$E_{q\ominus}$	3.0463	.6399	-1.2814
$E_{\mu\ominus}$	.2330	.4928	-.1216
$E_{\ominus\ominus}$	-.1628	.0117	.1727
total	-19.8815	-21.8082	-21.4510
converged		-21.1	

^aSee footnotes of Table IV.

Table VII: Moment–Moment Intermolecular Electrostatic Energy Contributions (kcal/mol) Calculated for Crystalline Formamide at Experimental Geometry Using DMA Atomic Moments Through Order Seven. ^{a, b}

Atomic Moment	Atomic Moment							
	0	1	2	3	4	5	6	7
0	-27.14	11.34	-.99	1.66	-.11	-.01	.03	-.02
1		-1.19	-.21	-.18	-.03	-.00	.00	-.00
2			.04	.04	-.01	.00	.00	-.00
3				-.01	.01	.00	-.00	.00
4					.00	-.00	.00	-.00
5						.00	-.00	-.00
6							-.00	.00
7								.00

^aAll unit cells within 16 Å of the central unit cell are included in the calculation.

^bTotal intermolecular electrostatic energy per molecule is -16.81 kcal/mol.

Table VIII: Comparison of Atomic Monopoles, Dipoles, and Quadrupoles ^{a, b, c} for the Amide Functional Groups of Formamide, Oxamide, and Urea Calculated Using the Force-Related, DMA, and Hirshfeld Methods.

	Force-Related			DMA			Hirshfeld		
	form	ox	urea	form	ox	urea	form	ox	urea
amide N									
q	-.79	-.75	-.84	-.84	-.85	-.97	-.13	-.13	-.16
μ_z	-.15	-.16	-.18	-.14	-.12	-.04	.01	.01	.00
μ_x	-.02	.00	-.02	.02	.01	-.05	.00	.00	.00
Θ_{zz}	.42	.31	.41	.25	.23	.12	.07	.08	.10
Θ_{xz}	.02	.01	.02	-.07	-.03	.10	-.02	.00	-.02
Θ_{xx-yy}	.10	.16	.16	.17	.16	.20	.19	.17	.22
amide H (trans to O)									
q	.40	.39	.39	.40	.43	.38	.13	.14	.12
μ_z	.05	.05	.05	.03	.03	.02	.11	.10	.10
μ_x	.00	-.03	-.02	.00	.00	.00	.00	-.02	.00
Θ_{zz}	-.05	-.03	-.07	.00	.00	.01	.00	.01	.01
Θ_{xz}	-.02	-.05	-.02	.00	.00	.00	.00	-.05	.00
Θ_{xx-yy}	-.01	-.07	-.03	.00	.00	.00	.05	.02	.05
amide H (cis to O)									
q	.39	.38	.40	.41	.41	.42	.14	.14	.14
μ_z	.06	.07	.07	.03	.03	.03	.11	.11	.11
μ_x	-.02	-.02	-.03	.00	.00	.00	-.01	-.01	.00
Θ_{zz}	-.01	-.08	-.07	.00	.00	.00	.01	.01	.02
Θ_{xz}	.04	.05	.05	.00	.00	.00	.00	.00	-.01
Θ_{xx-yy}	-.14	-.06	-.06	.00	.00	.00	.03	.04	.04
amide C									
q	.48	.49	.68	.92	.93	1.38	.20	.22	.24
μ_z	-.07	-.05	.07	-.04	-.03	-.07	-.02	-.01	.05
μ_x	.00	-.06	.00	.32	.24	.00	-.05	-.05	.00
Θ_{zz}	-.22	-.28	-.26	.07	.03	.05	-.16	-.16	-.16
Θ_{xz}	-.20	-.29	.00	.07	.14	.00	-.10	-.10	.00
Θ_{xx-yy}	-.27	-.12	-.51	-.15	-.07	-.04	-.22	-.23	-.19
amide O									
q	-.55	-.51	-.58	-.92	-.92	-1.04	-.40	-.38	-.44
μ_z	-.04	-.06	-.05	.15	.14	.20	-.10	-.10	-.11
μ_x	.00	-.01	.00	-.04	.00	.00	.01	-.01	.00
Θ_{zz}	.17	.14	.20	.12	.12	.09	.08	.05	.09
Θ_{xz}	-.03	-.06	.00	.11	.07	.00	-.01	-.03	.00
Θ_{xx-yy}	.01	.02	.04	-.14	-.18	-.17	-.05	-.02	-.02

^aUnits of charge and distance are e_cu and Å, respectively.

^bMoments are expressed with respect to the local coordinate systems of Figure 2.

^cThe symmetries of the quadrupole components are those of the corresponding d orbitals in quantum chemistry.

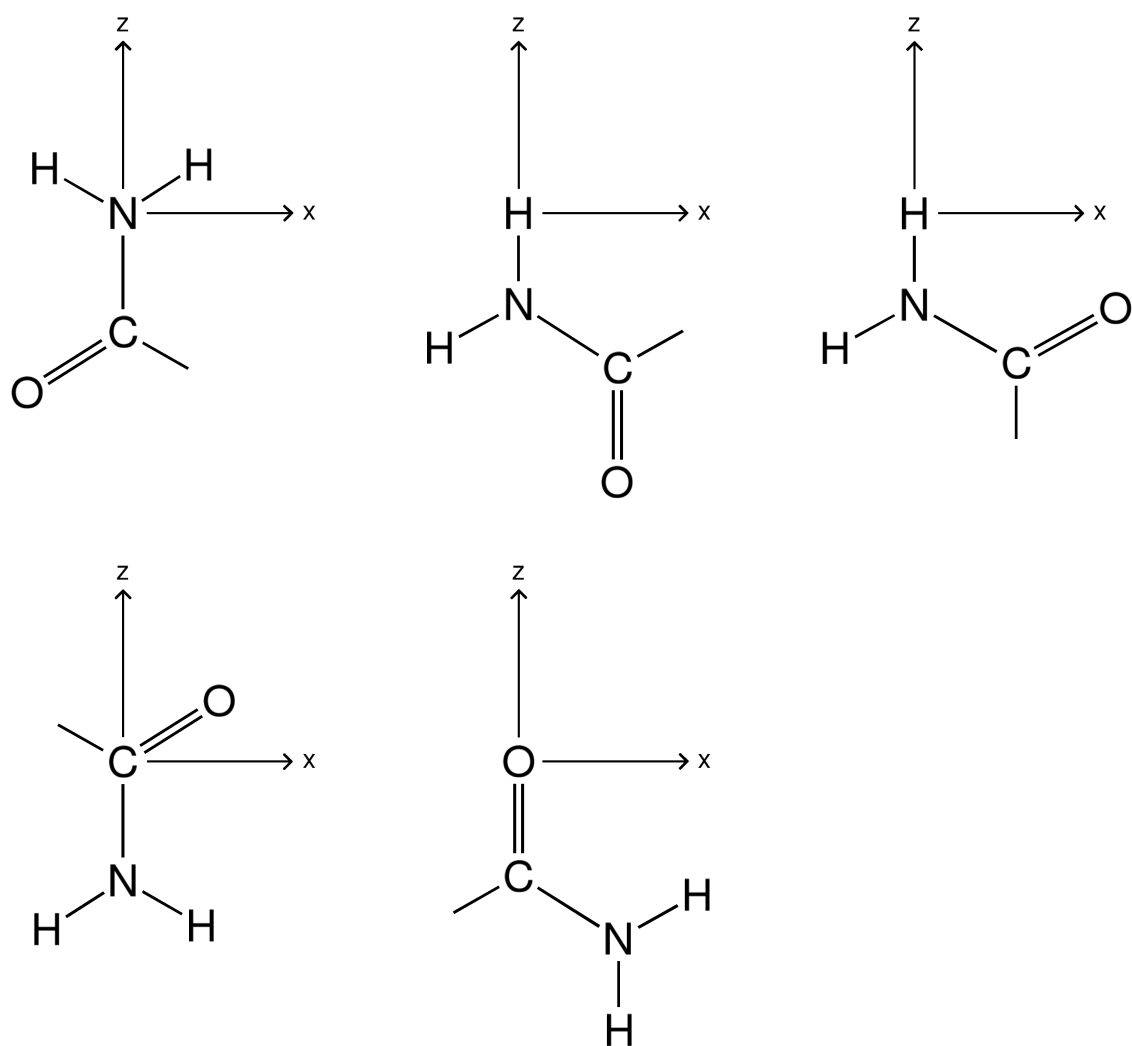


Figure 2: Local coordinate systems for atoms of the amide functional group.

Table IX: Average Absolute Error ^a (and Maximum Absolute Error) of Alternative Multipole Expansion Representations of the Electrostatic Potential of the Formamide Wavefunction at Points on the Molecular Surface.

Surf	Electrostatic Models				
	FR-o(0) ^b	DMA-o(2) ^c	FR-o(2)	H-o(2) ^d	DMA-o(7) ^e
vdw+0.3 ^f	2.30(8.09)	1.15(5.37)	1.78(8.42)	1.22(4.42)	0.47(1.96)
vdw+0.6	1.54(4.67)	0.60(2.64)	0.97(4.83)	0.59(2.01)	0.08(0.39)
vdw+0.9	1.05(3.02)	0.35(1.35)	0.61(3.10)	0.32(0.95)	0.01(0.08)
vdw+1.2	0.79(2.09)	0.23(0.79)	0.44(2.10)	0.19(0.56)	0.00(0.01)
vdw+1.5	0.60(1.52)	0.15(0.50)	0.31(1.48)	0.12(0.35)	0.00(0.00)
vdw+1.8	0.46(1.13)	0.10(0.33)	0.24(1.05)	0.08(0.23)	0.00(0.00)

^aThe unit of electrostatic potential is 1 kcal/mol ecu.

^bForce-Related atomic moments. Truncation at monopole term. Resulted in higher overall accuracy than potential-derived partial atomic charge models tested.

^cDMA atomic moments. Truncation at quadrupole terms.

^dHirshfeld atomic moments. Truncation at quadrupole terms.

^eTruncation at 7th moment terms.

^fConnolly molecular surface defined by atomic van der Waals radii of {1.4, 0.8, 1.9, 1.75, 1.6} Å, respectively, for {H, polar H, C, N, O}, plus an outward displacement of .3 Å and a probe radius of .8 Å.

Table X: Comparison of Predicted Crystal Structures ^a of Formamide Obtained Using Atomic Monopoles, Dipoles, and Quadrupoles from the Force-Related, DMA, and Hirshfeld Methods to the Experimental Structure.

	Unit Cell Vectors			
	Experimental ⁽²⁷⁾	Calculated		
		Force-Related	DMA	Hirshfeld
a	3.61	3.66(.05)	3.64(.03)	3.59(-.02)
b	9.05	9.84(.79)	9.54(.49)	9.69(.64)
c	8.42	7.29(-1.13)	7.41(-1.01)	7.41(-1.01)
α	90.00	90.00(.00)	89.98(-.02)	90.00(.00)
β	125.39	125.74(.35)	124.91(-.48)	124.46(-.93)
γ	90.00	90.02(.02)	90.32(.32)	90.06(.06)

^aSee footnotes of Table I.

outside of the surface of the molecule, the electrostatic potential of the wavefunction is reproduced almost perfectly by all three sets. Of the three methods, atomic monopoles tend to be largest for DMA, smaller for Force-Related, and smallest for Hirshfeld. For the DMA method, the point dipoles tend to oppose the electric field created by the point monopoles, with hydrogen being an exception to this rule. In contrast, the Hirshfeld dipoles tend to reinforce the electric field created by the Hirshfeld monopoles. The Force-Related dipoles, the vector sum being zero, lie somewhere in between.

3.4 Comparison of Predicted Structures to Experiment.

A comparison of the predicted crystal structures of formamide obtained using atomic monopoles, dipoles, and quadrupoles to the experimental structure is given in Table X. For all of the methods, the movement away from the experimental structure is about 1 Å in unit cell lengths. As a point of reference, the three predicted structures differ from one another by only a few tenths of an angstrom in unit cell lengths. This result is consistent with the previous observation that, at the quadrupole level, the three models are roughly similar in their extent of convergence. It suggests that differences between the three models may be small when measured by changes in predicted crystal structures.

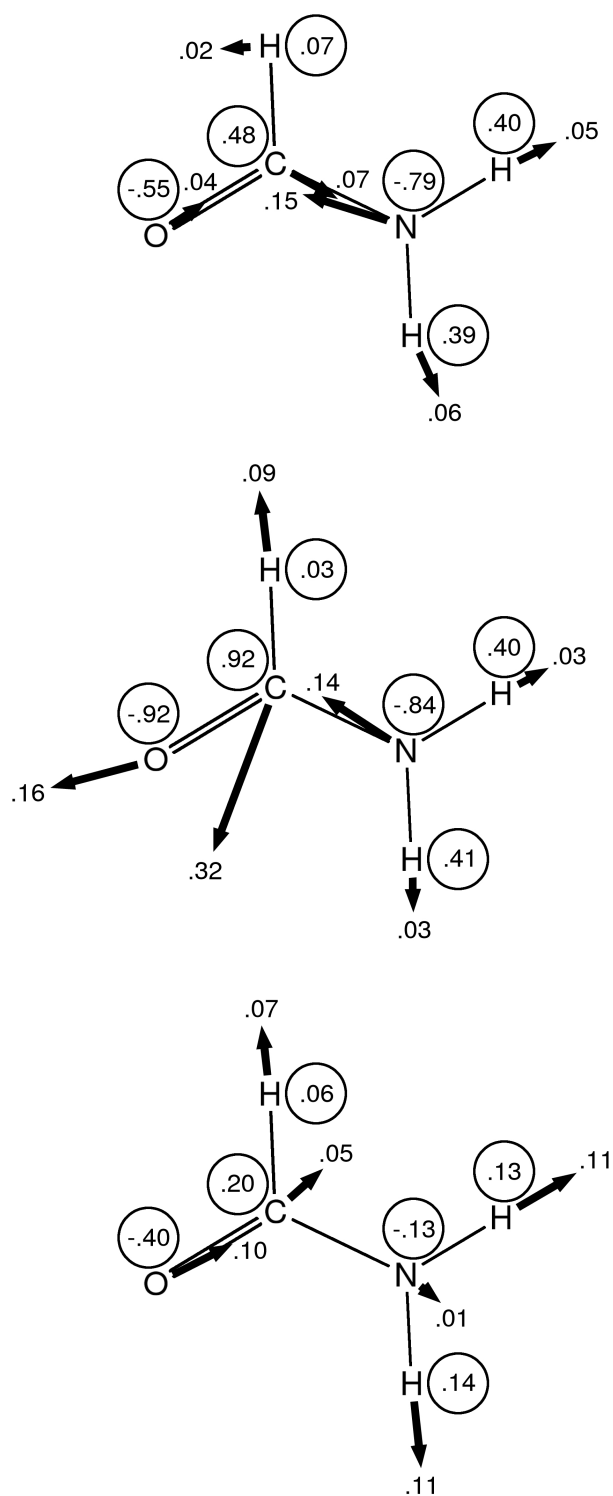


Figure 3: A comparison of atomic monopoles and dipoles of formamide calculated using three distinct methods: Force-Related, DMA, and Hirshfeld.

Table XI: Comparison of Predicted Crystal Structures ^a of Oxamide Obtained Using Atomic Monopoles, Dipoles, and Quadrupoles from the Force-Related, DMA, and Hirshfeld Methods to the Experimental Structure.

		Unit Cell Vectors		
Experimental ⁽²⁷⁾		Calculated		
		Force-Related	DMA	Hirshfeld
a	3.63	3.56(-.07)	3.64(.01)	3.64(.01)
b	5.19	5.16(-.03)	5.13(-.06)	5.14(-.05)
c	5.66	5.60(-.06)	5.51(-.15)	5.61(-.05)
α	83.70	82.83(-.87)	86.16(2.46)	82.54(-1.16)
β	114.10	113.91(-.19)	113.28(-.82)	113.82(-.28)
γ	115.10	115.39(.29)	115.41(.31)	115.77(.67)

^aSee footnotes of Table I.

Table XII: Comparison of Predicted Crystal Structures ^a of Urea Obtained Using Atomic Monopoles, Dipoles, and Quadrupoles from the Force-Related, DMA, and Hirshfeld Methods to the Experimental Structure.

		Unit Cell Vectors		
Experimental ⁽²⁷⁾		Calculated		
		Force-Related	DMA	Hirshfeld
a	5.66	6.72(1.06)	6.81(1.15)	5.47(-.19)
b	5.66	5.10(-.56)	5.01(-.65)	5.45(-.21)
c	4.71	3.67(-1.04)	3.68(-1.03)	4.78(.07)
α	90.00	90.01(.01)	90.09(.09)	89.99(-.01)
β	90.00	89.97(-.03)	89.96(-.04)	90.02(.02)
γ	90.00	89.62(-.38)	89.81(-.19)	90.00(.00)

^aSee footnotes of Table I.

The same data is presented for oxamide and urea in Tables XI and XII. For oxamide, the largest movements away from the experimental structure are a few degrees in the unit cell angles. The three predicted structures differ from one another by only about .1 Å in unit cell lengths, and by a few degrees in unit cell angles. For urea, the Force-Related and DMA models differ by only about .1 Å in unit cell lengths, and the Hirshfeld model differs from the other two models by about 1 Å in unit cell lengths. As a test for the possible existence of multiple minima on the energy surface of the Hirshfeld model, a second minimization was carried out, this time starting from the minimum-energy structure that was obtained using the Force-Related model. Again, the final energy-minimized structure was not changed by this alternative choice of the starting point.

For amides, *ab initio* calculations at the HF/6-31G* level tend to overestimate molecular dipoles by about 12%.⁽³⁷⁾ To determine the sensitivity of the predicted crystal structures to the expected overestimate of molecular dipoles, a second set of crystal minimizations was carried out, this time using the previously calculated atomic multipoles scaled uniformly by a factor of $\frac{1}{1.12}$. For all of the predicted crystal structures, changes induced by uniform scaling of atomic multipoles are minimal, the major effect being a small increase in most of the predicted unit cell lengths.

3.5 Comparison of Predicted Energies to Experiment.

A comparison of the predicted sublimation energies of formamide, oxamide, and urea obtained using atomic monopoles, dipoles, and quadrupoles to the experimental sublimation energies is given in

Table XIII: Comparison of Predicted Sublimation Energies ^a of Formamide, Oxamide, and Urea Obtained Using Atomic Monopoles, Dipoles, and Quadrupoles from the Force-Related, DMA, and Hirshfeld Methods to the Experimental Sublimation Energies.

	Energy (kcal/mol)		
	Force-Related	DMA	Hirshfeld
Formamide			
E _{nb}	-3.34	-2.28	-3.39
E _{elec}	-18.31	-20.53	-17.64
total	-21.65	-22.81	-21.03
expt ⁽²⁹⁾	-17.3		
Oxamide			
E _{nb}	-7.02	-4.80	-7.41
E _{elec}	-28.11	-32.04	-25.41
total	-35.13	-36.84	-32.82
expt ⁽²⁹⁾	-27.7		
Urea			
E _{nb}	-3.25	-3.30	-2.97
E _{elec}	-22.44	-23.57	-22.50
total	-25.69	-26.87	-25.47
expt ⁽²⁹⁾	-23.2		

^aAll unit cells within 16Å of the central unit cell are included in the calculation.

Table XIII. Predicted sublimation energies are too large by 10 to 25%.

4 DISCUSSION

For intermolecular electrostatic energy calculated at the experimental geometry of the crystal (Tables IV to VI), addition of atomic dipoles and quadrupoles causes maximum adjustments of 6.7, 13.5, and 15.2 kcal/mol for the Force-Related, DMA, and Hirshfeld models, respectively. Focusing temporarily on the Force-Related data, the contribution to calculated intermolecular electrostatic energy from multipole expansion terms containing atomic moments of dipole and higher is, for formamide and oxamide, roughly 2.5 kcal/mol per amide group in the direction of stronger binding. Although the corresponding energy contribution is less for the urea crystal, the magnitudes of individual terms remain significant. Based on theoretical considerations, it is clear that the collective contribution of these anisotropic terms depends on packing arrangement in a way that is potentially quite sensitive, in comparison to the coulomb term, to the relative positions and orientations of contacting groups. In contrast, the polarization contribution to the total intermolecular energy, also poorly represented by essentially all commonly used potential energy functions, is expected to be similarly large in magnitude (~20% as large as intermolecular electrostatic energy) yet relatively independent of packing arrangement and, therefore, less decisive in influencing the structure of close contacts. In this sense, the anisotropic electrostatic terms are expected to add to the remainder of the energy surface a higher frequency (more rapidly varying) component over the space of possible packing geometries. For an amide group in condensed phase, the data suggests that the height of this higher frequency component should be on the order of a few kilocalories per mole. As a specific example, consider two well-separated packing geometries for the urea crystal (Table XII),

the experimental geometry and the minimum energy geometry obtained using Force-Related multipoles through quadrupole. Anisotropic electrostatic terms engage the remainder of the energy surface, altering the energy difference between the two packing geometries by 3.35 kcal/mol in favor of the energy optimized structure.

For calculated unit cell lengths and angles (Tables I to III); addition of atomic dipoles and quadrupoles causes maximum adjustments of 1.2, 1.5, and 2.2 Å for the Force-Related, DMA, and Hirshfeld models, respectively; and of 1.2, 7.0, and 9.2°. For each of the three methods, the largest adjustment to unit cell lengths occurs for the urea crystal. It cannot be ruled out that the large structural shifts that were observed for this crystal in response to addition of atomic dipoles and quadrupoles could be just extreme sensitivity due to poor representation of the remainder of the total energy surface. In particular, the true repulsion + dispersion energy surface of the crystal may not be well represented here. Still, taken together, the magnitude of the anisotropic electrostatic contribution, its sensitivity to packing geometry and impact on calculated structures, suggest, inescapably, that atomic dipoles and quadrupoles may be a necessary component for accurate prediction of structure and energetics in amide crystals, and by association in biomolecular systems.

One could argue that, in the case of amides, much of intermolecular electrostatic energy originates from hydrogen bonding contacts, that anisotropic electrostatic terms act to increase electrostatic potential at donor and acceptor sites, that all low energy packing geometries must have these contacts, and that energetic calculations might be improved by choosing a set of partial atomic point charges that better reproduces electrostatic potential in these key contacting regions at the expense of worsening electrostatic potential in less important regions. In other words, the atomic partial charges are no longer a best overall electrostatic potential fit but rather a best fit over small regions of high electrostatic potential where hydrogen bonding contacts occur. Relevant to this discussion, a more detailed analysis of the anisotropic electrostatic contribution shows that, while the contribution of anisotropic electrostatic terms to hydrogen bonding interactions is in the direction of stronger binding (seen qualitatively for formamide in Figure 3) by an average of roughly .6 kcal/mol per hydrogen bond, the contribution to non-hydrogen bonding interactions is equally strong, and for each molecule, less than half of the added energy (Tables IV to VI) is attributable to a strengthening of hydrogen bonds.

At the quadrupole level, the extents of convergence of the three models were shown to be similar. As a consequence, changes to predicted quantities with the addition of atomic dipoles and quadrupoles are expected to decrease as the extent of convergence at the monopole level increases. At the monopole level, the Force-Related model is expected to be more fully converged than the DMA or Hirshfeld models. Consistent with these expectations, movements of predicted quantities with the addition of atomic dipoles and quadrupoles tend to be smaller for the Force-Related model than for the DMA or Hirshfeld models. The additional information that, at the monopole level, the Force-Related model has been shown to be competitive in the sense that its accuracy approaches the highest possible within the limitations of a partial atomic point charge model⁽³¹⁾ leads to an increased confidence that our conclusions concerning the significance of atomic dipoles and quadrupoles will not fail to hold for some as yet untested method.

Because of the possibility of errors in a given representation of the repulsion or polarization components of the total energy, the substitution of a more accurate representation of the electrostatic component is not guaranteed to cause predicted quantities to move closer to the experimental target. Consistent with this possibility, it can be seen from Tables I to III and X to XII that, in some cases, the addition of atomic dipoles and quadrupoles does cause the predicted crystal structure to move away from the experimental structure. A more direct path toward determining the importance of atomic dipoles and quadrupoles for accurate prediction of structure and energetics could be cleared by removing the uncertainty in the relationship between the accuracy of an electrostatic model and the distance from the experimental target. This next level of study will require more accurate representations of repulsion and possibly polarization.

For all of the methods, the intermolecular electrostatic energy is not fully converged at the quadrupole level. This lack of complete convergence has its origin in the small internuclear distances that occur in hydrogen bonds and other contacts between molecules. In addition to convergence difficulties, the goal

of obtaining an accurate model of the electrostatic energies of close contacts is further complicated by penetration. However, as an alternative to explicit modeling of the penetration energy,^(36,38) the nature of the distance dependence may enable representation of this term in the form of a modification to the repulsion parameters.

5 Acknowledgment

We thank C. S. Ewig for helpful discussion.

References

- [1] A. T. Hagler, E. Huler, and S. Lifson, *J. Am. Chem. Soc.*, **96**, 5319–5327 (1974).
- [2] A. T. Hagler, and S. Lifson, *J. Am. Chem. Soc.*, **96**, 5327–5335 (1974).
- [3] F. A. Momany, L. M. Carruthers, R. F. McGuire, and H. A. Scheraga, *J. Phys. Chem.*, **78**, 1595–1620 (1974).
- [4] S. Lifson, A. T. Hagler, and P. Dauber, *J. Am. Chem. Soc.*, **101**, 5111–5121 (1979).
- [5] S. Lifson, A. T. Hagler, and P. Dauber, *J. Am. Chem. Soc.*, **101**, 5122–5130 (1979).
- [6] K. T. No, O. Y. Kwon, S. Y. Kim, K. H. Cho, C. N. Yoon, Y. K. Kang, K. D. Gibson, M. S. Jhon, and H. A. Scheraga, *J. Phys. Chem.*, **99**, 13019–13027 (1995).
- [7] M. D. Beachy, D. Chasman, R. B. Murphy, T. A. Halgren, and R. A. Friesner, *J. Am. Chem. Soc.*, **119**, 5908–5920 (1997).
- [8] F. L. Hirshfeld, *Acta Cryst.*, **B27**, 769–781 (1971).
- [9] A. T. Hagler, and A. Lapicciarella, *Biopolymers*, **15**, 1167–1200 (1976).
- [10] Z. Berkovitch-Yellin, and L. Leiserowitz, *J. Am. Chem. Soc.*, **102**, 7677–7690 (1980).
- [11] A. J. Stone, and S. L. Price, *J. Phys. Chem.*, **92**, 3325–3335 (1988).
- [12] S. L. Price, *Molecular Simulation*, **1**, 135–156 (1988).
- [13] P. Coppens, *J. Phys. Chem.*, **93**, 7979–7984 (1989).
- [14] S. Yashonath, S. L. Price, and I. R. McDonald, *Molec. Phys.*, **64**, 361–376 (1988).
- [15] D. J. Willock, M. Leslie, S. L. Price, and C. R. A. Catlow, *Mol. Cryst. Liq. Cryst.*, **235**, 217–224 (1993).
- [16] J. P. Ritchie, A. S. Copenhaver, and E. M. Kober, *Proceedings of the 10th International Symposium on Detonation*, Office of Naval Research, (1993).
- [17] D. J. Willock, S. L. Price, M. Leslie, and C. R. A. Catlow, *J. Comput. Chem.*, **16**, 628–647 (1995).
- [18] D. S. Coombes, S. L. Price, D. S. Willock, and M. Leslie, *J. Phys. Chem.*, **100**, 7352–7360 (1996).
- [19] A. J. Stone, *Chem. Phys. Lett.*, **83**, 233–239 (1981).
- [20] F. L. Hirshfeld, *Theoret. Chim. Acta*, **44**, 129–138 (1977).
- [21] J. Ladell, and B. Post, *Acta Cryst.*, **7**, 559–564 (1954).
- [22] T. Ottersen, *Acta Chem. Scand.*, **A29**, 939–944 (1975).

- [23] E. D. Stevens, *Acta Cryst.*, **B34**, 544–551 (1978).
- [24] B. H. Torrie, C. O'Donovan, and B. M. Powell, *Molec. Phys.*, **82**, 643–649 (1994).
- [25] E. M. Ayerst, and J. R. C. Duke, *Acta Cryst.*, **7**, 588–590 (1954).
- [26] G. de With, and S. Harkema, *Acta Cryst.*, **B33**, 2367–2372 (1977)
- [27] S. Swaminathan, B. M. Craven, and R. K. McMullan, *Acta Cryst.*, **B40**, 300–306 (1984).
- [28] S. Swaminathan, B. M. Craven, M. A. Spackman, and R. F. Stewart, *Acta Cryst.*, **B40**, 398–404 (1984).
- [29] H. G. M. de Wit, J. C. van Miltenburg, and C. G. de Kruif, *J. Chem. Thermodynamics*, **15**, 651–663 (1983).
- [30] U. Dinur, and A. T. Hagler, *J. Chem. Phys.*, **91**, 2949–2958 (1989).
- [31] U. Dinur, and A. T. Hagler, *J. Chem. Phys.*, **91**, 2959–2970 (1989).
- [32] U. Dinur, *J. Comput. Chem.*, **12**, 91–105 (1991).
- [33] U. Dinur, *J. Comput. Chem.*, **12**, 469–486 (1991).
- [34] *Gaussian 90*, M. J. Frisch, M. Head-Gordon, G. W. Trucks, J. B. Foresman, H. B. Schlegel, K. Raghavachari, M. A. Robb, J. S. Binkley, C. Gonzalez, D. J. Defrees, D. J. Fox, R. A. Whiteside, R. Seeger, C. F. Melius, J. Baker, R. L. Martin, L. R. Kahn, J. J. P. Stewart, S. Topiol, and J. A. Pople, Gaussian, Inc., Pittsburg, PA, 1990.
- [35] C. G. Gray, *Can. J. Phys.*, **54**, 505–512 (1976).
- [36] R. J. Wheatley, and J. B. O. Mitchell, *J. Comput. Chem.*, **15**, 1187–1198 (1994).
- [37] S. L. Price, J. S. Andrews, C. W. Murray, and R. D. Amos, *J. Am. Chem. Soc.*, **114**, 8268–8276 (1992).
- [38] J. P. M. Lommerse, S. L. Price, and R. Taylor, *J. Comput. Chem.*, **18**, 757–774 (1997).