Predicting Structure of Molecular Crystals from First Principles

Rafał Podeszwa,[1,2] Betsy M. Rice,[3] and Krzysztof Szalewicz[2]
[1]*Faculty of Chemistry, University of Warsaw, Pasteura 1, 02-093 Warsaw, Poland*
[2]*Department of Physics and Astronomy, University of Delaware, Newark, Delaware 19716, USA*
[3]*U.S. Army Research Laboratory, Aberdeen Proving Ground, Maryland 21005, USA*
(Received 2 April 2008; published 10 September 2008)

A recently developed method, symmetry-adapted perturbation theory based on the density-functional description of monomers [SAPT(DFT)], is shown to be sufficiently accurate and numerically efficient to facilitate predictions of the structure of molecular crystals from first principles. In one application, a SAPT(DFT) potential was used to generate and order polymorphs of the cyclotrimethylene trinitramine crystal, resulting in the lowest-energy structure in excellent agreement with the experimental crystal. In a different application, a SAPT(DFT)-based calculation reproduced the lattice energy of the benzene crystal to within a few percent.

DOI: 10.1103/PhysRevLett.101.115503 PACS numbers: 61.50.Ah, 31.15.xp, 34.20.Gj, 61.50.Lt

Molecular crystals are bound by intermolecular (van der Waals) forces, and knowledge of such force fields should be sufficient to predict crystal structures. In principle, accurate force fields can be obtained using electronic structure methods, but, for reasons discussed below, this has not yet been achievable in practice. Thus, theoretical investigations of crystal structures typically rely on empirical force fields that are parametrized using experimental information. Unfortunately, the predictive capability of such fields is limited, since a given field can describe well only the system used for its parametrization and thus is often not transferable even to polymorphs of this system. As a result, prediction of crystal structures has been considered an impossible task. In 1988, Maddox published [1] a provocative op-ed stating that "one of the continuing scandals" is that computational scientists are not able to predict crystal structures from molecular structures. This opinion was echoed first by Ball [2] and more recently by Desiraju [3], who wrote that the issue "eluded scientists for more than 50 years" and emphasized the low success rate of crystal structure predictions in the blind tests conducted by the Cambridge Crystallographic Data Center [4]. A subsequent third blind test [5] had a lower success rate than the previous tests. Very recently, a fourth, yet unpublished, test was performed, and the success rate has apparently improved [6]. One of the key issues in predicting crystal structures is the accuracy of the force fields. This accuracy is also critical for calculations of lattice energies at experimental crystal structures [7]. The force fields can be computed *ab initio* using wave-function (WF) based methods, but until recently the accuracy achievable for molecules containing more than a few atoms was far from quantitative and was insufficient for determination of crystal structures. One might have hoped that the problem could be resolved by the development of density-functional theory (DFT), which can be applied to systems containing hundreds of atoms. Unfortunately, conventional DFT methods fail badly in describing intermolecular interactions for which dispersion is the dominant component; such systems include molecular organic crystals. Recently, a method has been proposed [8–11] which combines symmetry-adapted perturbation theory (SAPT) [12] of intermolecular interactions with the Kohn-Sham DFT representation of monomers. SAPT is a perturbational approach starting from isolated Hartree-Fock or Kohn-Sham monomers and imposing the correct permutational symmetry on perturbed wave functions. In the latter case, some time-consuming terms describing intramonomer electron correlation can be omitted, resulting in a very low-cost method which gives predictions as accurate as those of high-level WF-based methods [13–15]. SAPT(DFT) has been so far applied mainly to dimers. The present work examines its ability to predict structures and lattice energies of extended objects: molecular crystals.

The effectiveness of first-principles approaches depends critically on the degree of flexibility of monomers, since the dimensionality of the potential energy surfaces increases dramatically when internal degrees of freedom are included. For a sufficiently rigid molecule, the gas-phase structure—which can be determined very accurately from *ab initio* calculations—will be close to the structure within the corresponding crystal; thus, a rigid-monomer potential employing the *ab initio*-optimized structure can be used. The cyclotrimethylene trinitramine (RDX) molecule—chosen for the present study since its crystal has been the subject of many literature investigations [16] and since extensive SAPT(DFT) calculations have been performed [17] for its dimer—is fairly rigid, but it does possess two nearly isoenergetic low-energy conformers [18], one of them very close to the structure observed in the RDX crystal [19]. A fully first-principle route would require the development of two rigid-monomer potentials, one for each conformer, which would be too costly at the present time. Therefore, at this one point we departed from this route and used the experimentally determined RDX monomer geometry [19].

We first generated a number of RDX polymorphs using the molecular packing and lattice minimization methods with MOLPAK/WMIN computer codes [20] and a simplified version of the SAPT(DFT) RDX dimer potential [17]. We had to develop such a version since WMIN codes cannot accommodate the published form [17]. This was done by refitting the atom-atom potentials with a formula containing Coulomb, exponential, and inverse sixth power terms. The simplification had only a small impact on crystal predictions. We checked this by performing one isothermal-isostress molecular dynamics (NsT MD) simulation with DL_POLY-2 software [21] on the low-energy polymorph with both potentials: The cell vectors of the time-averaged geometry differed by 0.2%–0.7% in the two cases. The molecular packing step systematically sampled 51 of the most-commonly observed coordination geometries of molecular crystals, producing 6859 candidate structures for each geometry. Lattice energy minimizations were then performed for the 500 most dense structures. We selected the 14 lowest-energy structures for the NsT MD simulations at ambient conditions, performed using the original potential [17]. The rigid-molecule approximation was assumed, and no symmetry constraints were imposed. One of the simulations produced an amorphous structure and will be neglected in the analysis below. In addition to the usual averaging of thermodynamic properties and crystalline parameters, time-averaged configurations were obtained by recording all of the atomic positions in the supercell along the trajectories. In order to determine whether space group symmetry was preserved during the simulations for each polymorph, a hypothetical ideal RDX crystal was generated by applying the appropriate crystalline space group operations to the averaged atomic positions for one of the symmetry-equivalent molecules in the unit cell. The maximum deviation of center-of-mass fractional coordinates from the ideal positions was 0.0032 fractional units. The maximum deviation of the Euler angles that define the orientation of the molecular principal axes of inertia was 3.01°. Clearly, the space group symmetries are well maintained throughout the simulations.

Table I lists our results for the four lowest-energy polymorphs. The lattice energy of our minimum-energy structure I is below the lattice energies of all of the other structures by at least 4.7 kJ/mol, thus eliminating ambiguity in stability rankings. Additionally, the crystallographic parameters of structure I are essentially identical to those of the experimental one: The lattice parameters differ by 0.1%–0.4%. The energies, which differ by 0.7%, are also close. However, the experimental value corresponds to a higher temperature and probably has a significant uncertainty which limits its usefulness as a benchmark. Figure 1 shows the superposition of the predicted unit cell onto the experimental one. The agreement with experiment is excellent, even better than typically achieved in simulations with empirical potentials fitted to a given set of crystal data. To estimate the uncertainties of our force field, we evaluated the main neglected component, the nonadditive three-body contributions to the lattice energy for the experimental geometry of the crystal: the HF energy, the dispersion energy [23], and the leading asymptotic dispersion term, obtaining the values of 3.2, 4.2, and 1.3 kJ/mol, respectively. The consecutive contributions were obtained using 28, 7, and almost 8×10^6 symmetry-unique trimers. These contributions are relatively small compared to the two-body component, and, therefore, the lattice parameters should change very little. Moreover, the total effect is dominated by the dispersion energy, which is fairly isotropic and would be similar for all of the polymorphs. Therefore, the relative energies of the polymorphs are unlikely to change after including the many-body forces. The absolute value of the lattice energy changes by 6.7% for structure I, suggesting that the agreement of the pairwise additive energy with experiment was partly due to the fact that the basis set incompleteness error of the pair potential and the three-body effects have opposite sign.

Whereas the agreement with experiment achieved for the structure of the RDX crystal is clearly excellent, the

TABLE I. Lowest-energy polymorphs of the RDX crystal. a, b, c, and β—lattice parameters, ρ—density, and the theoretical energy is the time-averaged potential energy per molecule from MD simulations at 298 K, whereas the experimental counterpart is the sublimation enthalpy (with the sign reversed). Zero-point energy correction is not included in the theoretical energy.

	Exp.[a]	I[b]	II	III	IV
Space group	*Pbca*	*Pbca*	*Pbca*	$P2_1/c$	$P2_1/c$
a (Å)	13.182	13.23	13.38	8.13	7.24
b (Å)	11.574	11.60	11.68	10.70	11.68
c (Å)	10.709	10.72	10.76	9.66	10.72
β (deg)	90.0	90.0	90.0	96.41	71.48
ρ (g/cm^3)	1.806	1.791	1.755	1.765	1.717
Energy (kJ/mol)	−130.2	−129.3	−124.6	−118.8	−118.6

[a] Refs. [19,22]. The measured value of the sublimation enthalpy [22] was averaged over the range 329–371 K.

[b] The minor differences compared to the results of Ref. [17] are mainly due to a truncation of the number of significant digits in the monomer's geometry by the MOLPAK package.

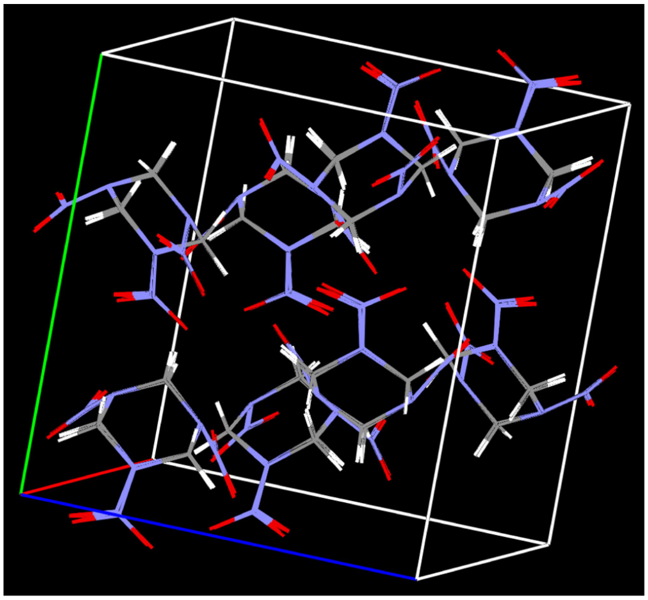

FIG. 1 (color online). The predicted RDX unit cell superimposed onto the experimental structure.

agreement for the lattice energy is less meaningful, since the experimental value may be a poor benchmark. To enable more precise conclusions about the accuracy of the SAPT(DFT) force fields, we have computed a quantity which is well known from experiments and which was regarded [7] as pathologically difficult to predict from first principles, the lattice energy of the benzene crystal at the experimental geometry [24]. We calculated SAPT(DFT) interaction energies (analogously to Ref. [15] but the potential fit has not been utilized) of the 15 nearest nonequivalent dimers, corresponding to 50 pairs by symmetry. The two-body interactions beyond the resulting cutoff of 11 Å were computed from the asymptotic expansion of Ref. [15] and were well converged by taking into account 8737 pairs.

As shown in Table II, our two-body lattice energy of −56.30 kJ/mol agrees already quite well with the experimental sublimation enthalpy at 298 K equal to −44.6 kJ/mol [29], about an order of magnitude better agreement than in the case of the recent predictions from Ref. [7]. However, several other contributions and corrections should be considered for a more reliable comparison. First, we were aware from our recent work [23] that for benzene the nonadditive effects can be substantial. We have therefore computed the three-body contributions to the lattice energy using the nonadditive SAPT(DFT) dispersion energy and the supermolecular second-order Møller-Plesset (MP2) energy, a method referred to as the MP2 + SDFT approach [23]. We considered all 96 nonequivalent trimers for which the average intermolecular distance is less than 9 Å (corresponding to 188 trimers by symmetry). The nonadditive dispersion contribution from larger trimers was computed from the asymptotic triple-dipole formula. Also the MP2 component has a long-range induction contribution which was computed from an appropriate asymptotic expression. As shown in Table II, the total three-body contribution is a fairly significant fraction (12.2%) of the two-body energy.

TABLE II. Contributions to lattice energy (in kJ/mol) of the benzene crystal.

	This work	Literature
2-body short-range	−54.82	
2-body long-range	−1.48	
Total 2-body	−56.30	(−111.3)–(−34.4)[a]
3-body MP2	0.45	
3-body dispersion short-range	6.19	
3-body dispersion long-range	0.35	
3-body induction long-range	−0.12	
Total 3-body	6.87	
2 + 3-body	−49.43	−51.06[b]
Lattice extrapolation to 0 K	−0.9[c]	
2 + 3-body for 0 K lattice	−50.33	−52.25[d]
Zero-point energy	2.78	2.80[e]
Lattice energy at 0 K	−47.55	−50.5[f]

[a] Ref. [7], various *ab initio* and semiempirical methods.

[b] From the potential of Ref. [25].

[c] From Ref. [26].

[d] Ref. [27], lowest-energy structure obtained by optimization of an empirical potential.

[e] Ref. [28].

[f] Obtained from the median experimental value of $H_{\mathrm{sub}}^{298} =$ 44.6 kJ/mol [29] and the calculated temperature correction (see text) equal to 5.90 kJ/mol.

The sum of the two- and three-body contributions can be compared to lattice energies predicted by benzene empirical potentials. Note that this comparison is appropriate, since such empirical pair potentials include effectively many-body contributions. Our value is 3.2% smaller in magnitude than we got with the pair potential of Ref. [25]. Furthermore, a molecular-mechanics optimization of the benzene crystal structure using another empirical potential has been published by van Eijck *et al.* [27]. To compare with this work, we should correct for the fact that the lattice geometry used by us was measured at 138 K [24], whereas the calculations of Ref. [27] correspond to 0 K. This (very small) correction for the lattice contraction has been taken from the literature [26]. As seen in Table II, our corrected lattice energy agrees to within 3.7% with the empirical result of Ref. [27].

To compare with experiment, we needed the zero-point vibrational energy which was obtained by calculating numerically the energy second derivatives of the central molecule within a cluster of 35 benzene molecules described by the potential of Ref. [15]. The resulting value equals 2.78 kJ/mol, in very good agreement with Ref. [28]. Monomer relaxation upon sublimation should also be taken into account; however, we were unable to do so since the crystal geometry at 0 K is unknown. Use of the 138 K geometry would result in some double counting with the correction for the lattice-geometry extrapolation to 0 K. Nevertheless, we believe that we have included all of the effects larger than about 1 kJ/mol, and the major uncertainties of our lattice energy at 0 K are probably due to the inaccuracies of the two- and three-body components.

The lattice energy cannot be directly measured; however, it is equal to the sublimation enthalpy at 0 K. The available sublimation enthalpy [29] measured at 298 K can be corrected for temperature effects using a formula developed in Ref. [30] (for the range $T = 4$–270 K but it can be extrapolated to 298 K). By integrating this formula over the range 4–298 K, we obtained a value of 15.81 kJ/mol from which one has to subtract $4RT$—the enthalpy of the ideal gas at 298 K plus the gas expansion work. This gives a correction of 5.90 kJ/mol, which, when added to the median value of the experimental sublimation enthalpies [29] at 298 K, results in the lattice energy at 0 K equal to -50.5 kJ/mol. The difference of 5.8% between this result and our final lattice energy of -47.55 kJ/mol is comparable to the analogous differences for the lattice energies computed from the empirical potentials and to the spread of experimental values [29]. This difference is also commensurate with internal assessments of uncertainties of theoretical components: The accuracy of the two-body potential was estimated [15] to be about 1 kJ/mol, and it was found in Ref. [23] that the MP2 + SDFT approach tends to overestimate the three-body energies by about 20%, which would result in an about 1 kJ/mol too large contribution to the lattice energy of benzene. Out of the contributions neglected by us, higher than three-body interactions and the effects of monomer geometry relaxation upon sublimation may also be of the order of 1 kJ/mol.

In conclusion, we have demonstrated that the SAPT (DFT) method is capable of producing force fields for interactions of relatively large organic molecules that enable reliable predictions of crystal structures of these compounds. In a subsequent paper, we will show that such predictions can also be made for polymorphic forms of the RDX crystal and for additional crystal properties. The predictions can be done entirely from first principles, eliminating reliance on empirical force fields and enabling treatment of compounds for which experimental data are unavailable. We expect our method to find broad applications in crystal design, in particular, to screening novel materials and drug candidates, screening molecules for cocrystallization, and identification of low-energy polymorphs of pharmaceutical compounds.

This research was supported by an ARO DEPSCOR grant, by ONR, by NSF Grant No. CHE-0555979, by the Polish Science Foundation grant Homing, and by a DoD HPCMP Challenge Project.

Note added in proof.—Two related studies of crystal modeling [31,32] and a calculation of lattice energy of the benzene crystal based on the two-body coupled-cluster approach [33] have just been published.

[1] J. Maddox, Nature (London) **335**, 201 (1988).

[2] P. Ball, Nature (London) **381**, 648 (1996).

[3] G. R. Desiraju, Nature Mater. **1**, 77 (2002).

[4] W. D. S. Motherwell *et al.*, Acta Crystallogr. Sect. B **58**, 647 (2002).

[5] G. M. Day *et al.*, Acta Crystallogr. Sect. B **61**, 511 (2005).

[6] K. Sanderson, Nature (London) **450**, 771 (2007).

[7] W. B. Schweizer and J. D. Dunitz, J. Chem. Theory Comput. **2**, 288 (2006).

[8] A. J. Misquitta and K. Szalewicz, Chem. Phys. Lett. **357**, 301 (2002).

[9] A. J. Misquitta, B. Jeziorski, and K. Szalewicz, Phys. Rev. Lett. **91**, 033201 (2003).

[10] A. J. Misquitta, R. Podeszwa, B. Jeziorski, and K. Szalewicz, J. Chem. Phys. **123**, 214103 (2005).

[11] A. Hesselmann and G. Jansen, Chem. Phys. Lett. **367**, 778 (2003).

[12] K. Szalewicz, K. Patkowski, and B. Jeziorski, Structure and Bonding (Berlin) **116**, 43 (2005).

[13] A. Hesselmann, G. Jansen, and M. Schütz, J. Chem. Phys. **122**, 014103 (2005).

[14] R. Podeszwa, R. Bukowski, and K. Szalewicz, J. Chem. Theory Comput. **2**, 400 (2006).

[15] R. Podeszwa, R. Bukowski, and K. Szalewicz, J. Phys. Chem. A **110**, 10 345 (2006).

[16] P. M. Agrawal, B. M. Rice, L. Q. Zheng, and D. L. Thompson, J. Phys. Chem. B **110**, 26 185 (2006).

[17] R. Podeszwa, R. Bukowski, B. M. Rice, and K. Szalewicz, Phys. Chem. Chem. Phys. **9**, 5561 (2007).

[18] B. M. Rice and C. F. Chabalowski, J. Phys. Chem. A **101**, 8720 (1997).

[19] C. S. Choi and E. Prince, Acta Crystallogr. Sect. B **28**, 2857 (1972).

[20] J. R. Holden, Z. Du, and H. L. Ammon, J. Comput. Chem. **14**, 422 (1993).

[21] W. Smith and T. R. Forester, J. Mol. Graphics **14**, 136 (1996).

[22] J. M. Rosen and C. Dickinson, J. Chem. Eng. Data **14**, 120 (1969).

[23] R. Podeszwa and K. Szalewicz, J. Chem. Phys. **126**, 194101 (2007).

[24] G. E. Bacon, N. A. Curry, and S. A. Wilson, Proc. R. Soc. A **279**, 98 (1964).

[25] W. Jorgensen and D. Severance, J. Am. Chem. Soc. **112**, 4768 (1990).

[26] A. I. M. Rae and R. Mason, Proc. R. Soc. A **304**, 487 (1968).

[27] B. P. van Eijck, A. L. Spek, W. M. Mooij, and J. Kroon, Acta Crystallogr. Sect. B **54**, 291 (1998).

[28] M. Nakamura and T. Miyazawa, J. Chem. Phys. **51**, 3146 (1969).

[29] J. S. Chickos and J. W. E. Acree, J. Phys. Chem. Ref. Data **31**, 537 (2002).

[30] R. C. Lord, Jr., J. E. Ahlberg, and D. H. Andrews, J. Chem. Phys. **5**, 649 (1937).

[31] A. J. Misquitta, G. W. A. Welch, A. J. Stone, and S. L. Price, Chem. Phys. Lett. **456**, 105 (2008).

[32] M. A. Neumann, F. F. J. Leusen, and J. Kendrick, Angew. Chem., Int. Ed. **47**, 2427 (2008).

[33] A. L. Ringer and C. D. Sherrill, Chem. Eur. J. **14**, 2542 (2008).