# First-principles intermolecular binding energies in organic molecular crystals

W.F. Perger [a,*], Ravindra Pandey [a], Miguel A. Blanco [b], Jijun Zhao [c]

[a] *Department of Physics, Michigan Tech University, 1400 Townsend Drive, Houghton, MI 49931, USA*
[b] *Departamento de Química Física y Analítica, Universidad de Oviedo, 33006 Oviedo, Spain*
[c] *Institute for Shock Physics, Washington State University, USA*

## Abstract

The intermolecular binding (lattice) energies are calculated for the molecular crystals cyclotrimethylene trinitramine, pentaerythritol, and pentaerythritol tetranitrate using the CRYSTAL 98/03 and GAUSSIAN 98 programs, the DMOL program and the CASTEP program and compared with experiment. Calculating the theoretical intermolecular binding energy as a tool for testing the intrinsic quality of a calculation is explained. Optimization of the Gaussian basis sets for the CRYSTAL 98/03 programs versus adding more terms to the Gaussian basis set is examined. Increasing the size of the basis set consistently reduces basis-set superposition error whereas basis set optimization improved the quality of the calculation relatively little.
© 2004 Elsevier B.V. All rights reserved.

## 1. Introduction

The accurate calculation of organic molecular crystal (OMC) properties represents a formidable challenge of contemporary interest. Some OMCs are also high-explosive materials for which a detailed first-principles knowledge of their properties is useful for safe handling, storage, and yield. Three levels of interaction in OMCs have been identified [1]: electronic, intramolecular and intermolecular. The relatively strong intramolecular binding is responsible for the fact that vibrational frequencies for the isolated molecule (calculated, for example, with the GAUSSIAN 98 program) compare reasonably well with those in experiments using crystals, for example, PETN [2] (although, of course, the lattice modes cannot be calculated this way). In this case, the agreement between theory and experiment for these high-frequency vibrational modes supports the idea that the properties of the molecule are more-or-less unaf-

fected by the presence of the lattice. However, the binding of the molecules in the lattice, being due to the relatively weak van der Waals interaction, must be accurately described if other properties, such as the lattice vibrational modes (acoustic and optical), elastic constants, or the lattice energy itself, are to compare well with experiment. To address this need for an accurate intermolecular potential, one can consider the molecular crystal at a perturbed isolated molecule level where the existence of adjacent molecules is simulated by introducing an external Coulomb potential produced by the charge distribution of these adjacent molecules into the molecular Hamiltonian [3]. This approach was used by Krijin et al. [4,5] using a local-density approximation (LDA) for crystalline α-oxalic acid dihydrate.

If the molecular interaction energy is to be determined, then terms in addition to the electrostatic one must be included, such as with the two-body additivity approximation [3], which was used by Avoird et al. [6,7]. Neither of these approaches is fully periodic, in contrast to the all-electron method used in the CRYSTAL [8] program, and the work of Dovesi et al. [9] on urea represents one of the first efforts using this first-principles Hartree–Fock (HF) linear combination of atomic

---

* Corresponding author. Fax: +90-648-72949.
*E-mail addresses:* wfp@mtu.edu (W.F. Perger), pandey@mtu.edu (R. Pandey).
*URL:* http://www.ee.mtu.edu/faculty/wfp.html (W.F. Perger).

orbitals (HF-LCAO) approach. As noted in that work, although the CRYSTAL program is, in principle, superior to the afore-mentioned approaches, several practical issues arise, such as basis set completeness and numerical accuracy, which must be carefully considered prior to reaching a conclusion as to the quality of a specific calculated property. Some guidance in this regard has been given in a recent work by Spackman and Mitchell [10] on Gaussian basis sets for molecular crystals and their influence on lattice energies revealed progressively smaller basis-set superposition error (BSSE) using the 6-21G**, 6-31G**, and double-zeta plus polarization (DZP) sets, in that order.

Because the weak intermolecular interaction is responsible for the mechanical and elastic properties, which are in turn of importance for a detailed understanding of the shock initiation-to-detonation transition, one of the goals of the current work is to calculate the lattice energy for several OMCs, using different theoretical methods, and compare with experiment. The materials PETN, $[C(CH_2ONO_2)_4]$, and RDX, $(CH_2NNO_2)_3]$, were chosen for the current study because they are high-explosives; PE, $C(CH_2OH)_4$, was chosen because, although not an energetic material, it is a tetragonal crystalline structure similar to PETN, but with 21 atoms per primitive unit cell, it is somewhat simpler computationally than PETN, which has 56 atoms per

unit cell. Structures for these materials are shown in Fig. 1.

The internal co-ordinates were taken from Choi [11] for RDX, Ladd [12] for PE, and Kitaigorodskii [13] for PETN. Different basis sets are used and contrasted, and issues of basis set optimization, as well as a new script using a parallel version of CRYSTAL 98 for this purpose, are described.

The weak intermolecular potential is difficult to model and therefore it is important to understand, in the context of the CRYSTAL 98/03 program, the relative merits of increasing the size of the basis set versus optimization of the exponents of a smaller set of Gaussian functions, which ultimately represent the atomic orbitals. Because the computational demands for OMCs are typically very high, it is therefore important to a priori know, if possible, which of these competing alternatives will provide the highest quality representation of the intermolecular potential and, ultimately, OMC properties. Consequently, one of the primary objectives of this work is to examine in detail this question, so as to provide guidance for future investigations in OMCs.

## 2. Methodology

The intermolecular binding (lattice) energy per molecule is defined here as:

$$-E_{lattice} = E_{crystal}/n - E_{molecule}, \qquad (1)$$

where $E_{crystal}$ is the crystal energy per unit cell, and $n$ is the number of molecules per unit cell. With this definition, including the minus sign, $E_{lattice} > 0$ for a stable OMC. Eq. (1) requires two, distinct, calculations for each value of the lattice energy. Notice the difference with respect to the conventional binding energy, in which the reference corresponds to isolated atoms. Thus, binding energies are in the order of hundreds of millihartrees ($mE_h$), whereas lattice energies are in the order of tens of $mE_h$. Given the typical values of atomic energies in OMC, ranging in hundreds of $mE_h$, computing $E_{lattice}$ requires more than five significant figures of precision in the individual calculations, and thus basis set completeness and numerical error must be considered in order to achieve a reasonable value. As such, the accurate determination of the lattice energy is sensitive to the overall quality of the basis set employed, as well as the level of correlation included. In addition to improving upon the basis sets currently used for OMCs, comparison with independent theoretical approaches offers further insight as to the overall quality of the calculation. Conformational changes are not included in the present study.

The first theoretical approach used in this work was the CRYSTAL program (both 98 [8] and 03 [14] versions), as it had been used with success for other studies
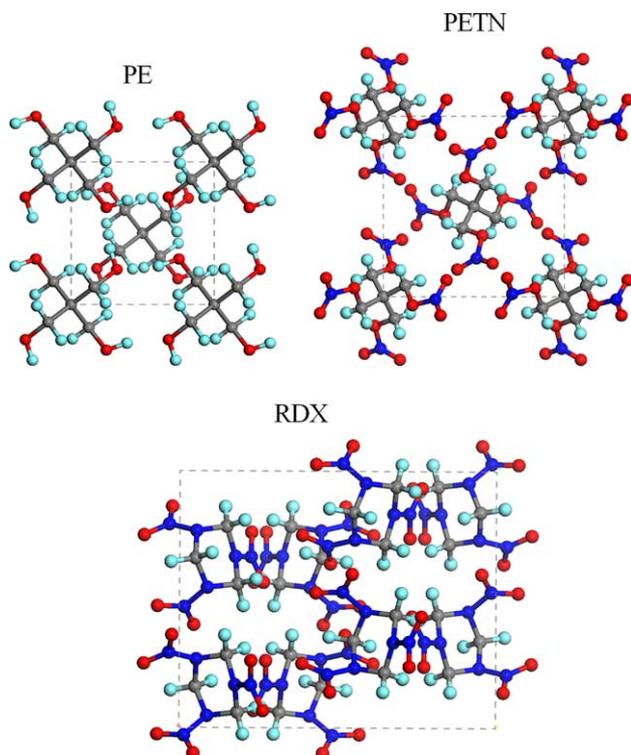


Fig. 1. The atomic structures of molecular crystals, PE, PETN, and RDX (all from top view of the three-dimensional lattice shown by the light grey dashed lines). Hydrogen atoms: light green balls; carbon atoms: grey balls; nitrogen atoms: blue balls; oxygen: red balls.

with molecular crystals [9,10,15,16]. In that approach, a given crystalline orbital, $\psi_i(\mathbf{k},\mathbf{r})$, is expressed as a linear combination of Bloch functions, $\phi_\mu(\mathbf{k},\mathbf{r})$, which are defined in terms of a set of fixed atomic orbital basis functions, $\zeta_\mu(\mathbf{r} - \mathbf{A}_\mu - \mathbf{g})$, centered on atom $\mathbf{A}_\mu$ within cell $\mathbf{g}$ [8]

$$\psi_i(\mathbf{k},\mathbf{r}) = \sum_\mu a_{\mu,j}(\mathbf{k})\phi_\mu(\mathbf{k},\mathbf{r}), \tag{2}$$

where

$$\phi_\mu(\mathbf{k},\mathbf{r}) = \sum_g \zeta_\mu(\mathbf{r} - \mathbf{A}_\mu - \mathbf{g})e^{i\mathbf{k}\cdot\mathbf{r}}. \tag{3}$$

The local functions, $\zeta_\mu(\mathbf{r} - \mathbf{A}_\mu - \mathbf{g})$, are in turn defined by a contraction, i.e., a fixed linear combination of $n_G$ Gaussian-type functions (GTFs)

$$\zeta_\mu(\mathbf{r} - \mathbf{A}_\mu - \mathbf{g}) = \sum_{j=1}^{n_G} d_j G(\alpha_j;\ \mathbf{r} - \mathbf{A}_\mu - \mathbf{g}). \tag{4}$$

Finite basis set LCAO schemes, while having the advantage of being able to accurately describe the core and valence atomic orbitals with few functions relative to a plane-wave basis set, have the disadvantage of certain well-known characteristics, such as BSSE [9]. Furthermore, computational demands have limited the use of basis sets to 6-21G in urea just a little more than a decade ago [9] and more recently to 6-21G for systems as large as RDX [16]. In the present work, we extend the quality of the basis sets used in two, distinct, ways: (1) expanding the GTFs to 6-311G* and DZP; (2) performing basis set optimization on the exponents of the most diffuse Gaussians representing the valence electrons. The coefficients and exponents of the GTFs were all taken from the Pacific Northwest National Laboratory's website [17]. Given the difficulty of optimizing all the exponents for all the Gaussians in a given basis set, the approach taken was to select the exponents of the most diffuse Gaussians describing the valence orbitals (including polarization functions, if they exist) for optimization, then an optimization script [18] was used to call a parallel version of CRYSTAL 98 using parallel-virtual machine (PVM). The script chooses a set of exponent values around the nominal value, then calls the CRYSTAL 98/PVM program, checks the result for convergence, then selects the value of the exponent which minimizes the total energy.

In order to compare with the CRYSTAL 98/03 results and the available experiments, two other density func-

tional theory (DFT) programs, DMOL [19] and CASTEP [20], have been used. The CRYSTAL calculations were obtained using the HF exchange potential and DFT exchange–correlation potentials. The DFT calculations were performed using LDA exchange with VWN correlation, PW–GGA for exchange and correlation, and the B3LYP exchange–correlation hybrid potential (all exchange–correlation potentials are defined in [8]). The DMOL calculations are done using a double numerical basis. In the CASTEP calculations, we use ultrasoft pseudo-potentials with a large cut-off energy of 500 eV for the plane-wave basis. In both cases, the LDA and GGA–PW91 exchange–correlation functionals have been used.

## 3. CRYSTAL and GAUSSIAN basis sets

The choice of the basis set is a necessary but not sufficient condition for obtaining accurate first-principles calculations in crystalline systems [8]. For OMCs, although the intramolecular bonding may be described well with basis sets borrowed from years of development by quantum chemists (e.g., the standard Pople sets [21]), the intermolecular interaction must also be accurately described and it is not obvious that the basis sets developed for molecular calculations will, without adaptation, be suitable for many problems of interest for crystalline systems. Although in principle to obtain greater accuracy one could simply increase the number of GTFs in Eq. (4) to improve the representation of the crystal orbital, in practice this approach is intractable for many OMCs, owing to the rapid increase in the number of integrals necessary as a function of the number of GTFs. Furthermore, it is not obvious that simply increasing the size of the basis set by adding more Gaussians whose parameters were not determined for a crystalline calculation, will necessarily be more accurate than a smaller basis set which has been optimized for the crystalline system. To examine in detail the effects of the basis set on the lattice energy, the OMCs RDX, PE, and PETN, with characteristics shown in Table 1, were studied. The first basis set chosen was the common 6-21G, which used s(6)sp(2)sp(1) for carbon, nitrogen, and oxygen, and s(2)s(1) for hydrogen. The second set was the 6-311G*, with s(6)sp(3)sp(1)sp(1)d(1) for carbon, nitrogen, and oxygen, and s(3)s(1)s(1) for hydrogen. A DZP functions

Table 1
Relevant information for the molecular crystals studied

| Material | $\frac{\text{Atoms}}{\text{molecule}}$ | $\frac{\text{Molecules}}{\text{unit cell}}$ | Point group | Space group | Structure |
|---|---|---|---|---|---|
| RDX | 21 | 8 | $C_s$ | Pbca | Orthorhombic |
| PE | 21 | 2 | $S_4$ | I$\bar{4}$ | Tetragonal |
| PETN | 29 | 2 | $S_4$ | P$\bar{4}2_1$c | Tetragonal |

set was also chosen owing to its recent success in calculating lattice energies for other molecular crystals [10]. In addition to the different basis sets, different exchange–correlation potentials and functionals were also chosen: HF, density-functional theory with B3LYP, PW–GGA, and LDA for exchange and VWN for correlation (LDA–VWN) [8]. Of the crystals studied, RDX represents the greatest computational challenge, with 1176 Gaussian basis functions required which results in an integrals file size of roughly 60 GBytes for the 6-311G* basis set.

The optimization of the exponents of just the most diffuse Gaussians can be a computationally difficult task because it requires the simultaneous adjustment of several parameters in order to achieve a minimum of the total energy; each adjustment of a given parameter is another complete run of the CRYSTAL program, which often means dozens or hundreds of runs for obtaining one set of optimized parameters. The procedure for optimization of the selected Gaussian exponents proceeds in the following manner. A given exponent is arbitrarily selected and a range of values around the nominal one is chosen, requiring one calculation of the CRYSTAL program per value in this range. A determination is then made as to where in that range the minimum energy occurs; the value of that exponent which minimizes the total crystalline energy is then saved and the process repeated for the next exponent. Once every exponent is optimized using this procedure, another complete pass is made, searching for the value of each exponent in turn which optimizes the total energy. This process is slow and not guaranteed to not get caught in local minima of the total energy surface. However, with the use of the parallel version of the CRYSTAL 98 program and PVM on a 15-node Linux cluster of 32-bit, 2.2 GHz machines, the total execution time is now reasonable. For example, the optimization of the DZP basis set for PE required 2.5 days, using 5 nodes. As for the issue of unambiguously identifying the global minimum of the total energy, the conjugate gradient approach employed by the LOPTCG script [8] was used as a cross-check and it gave very similar results for all exponents in urea.

## 4. Results

Table 2 shows a comparison of calculations of the lattice energy per mole for the three OMCs selected, using the CRYSTAL 98/03, DMOL, and CASTEP programs, and compared with experiment. The CRYSTAL 03 program (as opposed to CRYSTAL 98) was used for all DFT calculations (see the discussion below Table 3). In addition to the different basis sets employed, different exchange–correlation potentials were also chosen. All calculations obtained using the CRYSTAL 98/03 programs were corrected with the counterpoise method, as this was previously determined to be an important effect for molecular crystals [9,10]. The density-functional theory LDA approximation was found to reproduce the lattice energy per mole reasonably well, often slightly overestimating it. However, this agreement may be the consequence of a cancellation of errors, because it has recently been shown that LDA overestimates the

Table 2
Lattice energy per mole (in units of kJ/mol, 1 eV/molecule = 96.485 kJ/mol) for the chosen molecular crystals, with experimental values and calculations based on different programs and methods

| | 6-21G | | 6-311G* | | | DZP | | DMOL | | CASTEP | |
|------|------|------|------|------|------|------|------|------|------|------|------|
| | HF | B3 | HF | LDA | B3 | B3 | GGA | LDA | GGA | LDA | GGA |
| RDX | 65.0 | 43.1 | 40.4 | 115.5 | 36.3 | 38.3 | 57.8 | 143.8 | 63.30 | 135.6 | 61.47 |
| PE | 109.4 | 121.6 | 68.1 | 177.2 | 88.5 | 88.0 | 110.9 | 218.8 | 130.4 | 199.7 | 117.2 |
| PETN | 80.2 | 31.8 | 26.6 | 113.1 | 15.5 | 16.4 | 46.9 | 155.1 | 53.36 | 150.6 | 58.12 |

Experimental values: RDX, 130.1 [23]; PE, 163. [24]; PETN, 151.9 [25,26].

For the CRYSTAL calculations (first 7 columns), B3 indicates the B3LYP, LDA indicates the LDA–VWN, and GGA the PW–GGA exchange–correlation potentials. For the DMOL and CASTEP calculations, GGA refers to the GGA–PW91 exchange–correlation potential.

For CASTEP calculations, the ultrasoft pseudo-potential with 500 eV cut-off energy was used. All the structures for the crystal are taken from experiments.

Table 3
Lattice energy per mole (in units of kJ/mol, 1 eV/molecule = 96.485 kJ/mol) for PE calculated using the CRYSTAL 98/03 programs, including BSSE corrections and using optimization of the valence and polarization functions, where applicable

| | 621G | 621G-opt | 631G** | 631G**-opt | 6311G* | DZP | DZP-opt |
|-------|------------|------------|-----------|------------|-----------|-----------|-----------|
| HF | 109.4(−43) | 123.9(−53) | 67.7(−47) | 71.5(−30) | 68.1(−50) | 64.7(−18) | 72.4(−18) |
| B3LYP | 121.6(−56) | 123.0(−64) | 86.4(−41) | 86.7(−35) | 88.5(−37) | 88.0(−28) | 92.9(−22) |

The experimental value is 163 kJ/mol [24].

strength of the intermolecular potential in elastic constants [22]. It is significant to note from Table 2 that LDA results obtained from the different programs with different basis sets exhibit rough agreement with each other for all three OMCs, implying that programs using completely different types of basis sets can produce comparable results assuming the same level of correlation is included. Likewise, the PW–GGA results using CRYSTAL and a DZP basis set showed reasonable agreement with the GGA–PW91 results from DMOL and CASTEP.

Generally speaking, the 6-21G basis set with B3LYP results from the CRYSTAL 03 program produced results in better agreement with experiment but this agreement is most likely fortuitous. Regarding the two Gaussian basis sets chosen, the results of Table 2 do not show an improvement in the agreement with experiment as the CRYSTAL basis set was increased from 6-21G to 6-311G* to DZP, despite the dramatic increase in computational complexity.

Because of this increase in computational resources required as a function of increasing basis set size, it is of practical importance for large molecular crystalline calculations to know whether or not a smaller, but optimized, basis set might perform as well as a larger, unoptimized one. To answer this question, for PE the 6-21G, 6-31G**, and DZP basis sets were optimized by minimization of the total energy by varying the exponents of the valence GTFs. Table 3 shows the relative improvement in the lattice energy per mole compared with experiment for PE, for both the HF and B3LYP correlation potentials. Table 3 also shows the lattice energy per mole for the DZP basis set, and with optimization of valence and polarization terms. As with the 6-21G set, the optimization of the DZP set also improved the agreement with experiment. Of particular interest in Table 3 are the magnitudes of the BSSE, given in parentheses. There is a consistent trend towards lower BSSE with increasing size of the basis set. Furthermore, the optimization of the basis set has comparatively little effect on the BSSE. The relatively large BSSE found with the 6-21G basis set, combined with the fact that the lattice energies which it produces are clearly separated from the grouping of values arising from the other basis sets, suggests that, at least for lattice energies, the 6-21G basis set, although widely used for calculations of this kind, is inadequate. Finally, the DFT calculations obtained from the CRYSTAL 98 program exhibited inexplicable trends in the BSSE with increasing basis-set size and the CRYSTAL 03 program was therefore used for all DFT calculations reported here. As is observed from the BSSE values in Table 3, the trend of decreasing BSSE with increased basis set size is more-or-less the same for both the HF and DFT results. It is also to be noted that, according to the CRYSTAL 03 manual [14, pp. 182], 'The DFT

part of the code has been completely rewritten.' Hartree–Fock results were identical from the CRYSTAL 98 and CRYSTAL 03 programs.

## 5. Conclusions and future work

The lattice energy represents a test of the ability of a first-principles theoretical approach to model the relatively weak intermolecular potential found in OMCs. Three crystal modeling programs were used to calculate the lattice energy of RDX, PE, and PETN, each employing a different type of basis set: the CRYSTAL 98/03 programs, which uses GTFs to represent the atomic orbitals, DMOL, which uses a double-numerical basis set, and CASTEP, which uses plane waves. For the use of the CRYSTAL 98/03 programs, the better agreement with experiment observed by using the 6-21G basis set, with or without optimization, is fortuitous and a basis set with the complexity of the DZP set is needed because of its systematically lower BSSE. Although the 6-21G basis set has been widely used for studies in molecular crystals, the results of this work show that for lattice energies at least a 6-31G** set should be used. The computational demands of using the DZP set appear warranted, in agreement with the findings of Spackman and Mitchell [10] on other molecular crystals. The remaining discrepancies between the lattice energies calculated here and experiment is likely due to the inadequacy of the exchange–correlation potential used. Optimization of the most diffuse Gaussians does lead to a better estimate of the intermolecular potential as exhibited by closer agreement to experimental lattice energies, but the improvement is typically on the order of several percent. When the same exchange–correlation potential was chosen, LDA, the CRYSTAL 03, DMOL, and CASTEP programs showed relatively good agreement for all three molecular crystals studied. However, the HF potential does not include correlation effects and LDA overestimates the van der Waals interaction. The PW–GGA exchange–correlation gave results consistently closer to experiment than did the B3LYP, although it would be premature to make a definitive conclusion based on this limited evidence. Additional investigation on the choice of appropriate exchange–correlation potentials for calculation of organic molecular crystalline properties is therefore a priority for future work. Note that the structures used in all calculations were taken from experiment; if fully optimized, the theoretical lattice energies should be closer to experiment. Finally, in the present study, the optimized basis set affected both the intra- and inter-molecular potentials, and the relative amounts were roughly an order of magnitude larger change for the inter-molecular potential. It is therefore of interest to separate the

effects of basis set optimization on these potentials, which suggests a series of optimizations performed on suitable molecular systems.

## Acknowledgements

## References

[1] E.A. Silinsh, V. Ĉápek, Organic Molecular Crystals: Interaction Localization and Transport Phenomena, AIP Press, New York, 1994.

[2] Y.A. Gruzdkov, Y.M. Gupta, J. Phys. Chem. A 105 (2001) 6197.

[3] A.I. Kitaigorodskii, Molecular Crystals and Molecules, Academic Press, New York, 1973.

[4] M.P.C.M. Krijin, D. Feil, J. Chem. Phys. 89 (1988) 4199.

[5] M.P.C.M. Krijin, H. Graafsma, D. Feil, Acta Crystallogr. Sect. B 44 (1988) 609.

[6] W.J. Briels, A.P.J. Jansen, A. van der Avoird, Adv. Quantum Chem. 18 (1988) 131.

[7] A. van der Avoird, P.E.S. Wormer, F. Mulder, R.M. Berns, Top. Curr. Chem. 93 (1980) 1.

[8] V.R. Saunders, R. Dovesi, C. Roetti, M. Causà, N.M. Harrison, R. Orlando, C.M. Zicovich-Wilson, CRYSTAL 98 User's Manual, University of Torino, Torino, 1998.

[9] R. Dovesi, M. Causà, R. Orlando, C. Roetti, V. Saunders, J. Chem. Phys. 92 (12) (1990) 7402.

[10] M.A. Spackman, A.S. Mitchell, Phys. Chem. Chem. Phys. 3 (2001) 1518.

[11] C.S. Choi, E. Prince, Acta Cryst. B28 (1972) 2857.

[12] M.F.C. Ladd, Acta Cryst. B35 (1979) 2375.

[13] A.I. Kitaigorodskii, Organic Chemical Crystallography, Consultants Bureau Enterprises, New York, 1961.

[14] V.R. Saunders et al., CRYSTAL 2003 User's Manual, University of Torino, Torino, 2003.

[15] A.B. Kunz, Phys. Rev. B 53 (15) (1996) 9733-I.

[16] W.F. Perger, Chem. Phys. Lett. 368 (3–4) (2003) 319.

[17] Available from <http://www.emsl.pnl.gov/forms/basisform.html/>, Environmental Molecular Sciences Laboratory, 2003.

[18] Available from <http://www.tcm.phy.cam.ac.uk/~mdt26/crystal.html> (2003).

[19] B. Delley, J. Chem. Phys. 92 (1990) 508.

[20] V. Milman, B. Winkler, J.A. White, C.J. Packard, M.C. Payne, E.V. Akhmatskaya, R.H. Nobes, Int. J. Quantum Chem. 77 (2000) 895.

[21] D.J. Hehre, L. Radom, P.V.R. Schleyer, J.A. Pople, Ab initio Molecular Orbital Theory, Wiley, New York, 1986.

[22] J. Zhao, J.M. Winey, Y.M. Gupta, W.F. Perger, in: Shock Compression of Condensed Matter-2003, 13th APS Topical Group Meeting on Shock Compression of Condensed Matter, American Physical Society, Portland, OR, 2003.

[23] J. Rosen, C. Dickinson, J. Chem. Eng. Data 14 (1969) 120.

[24] D. Semmingsen, Acta Chem. Scand. A42 (1988) 279.

[25] J.S. Chickos, in: J.F. Liebman, A. Greenberg (Eds.), Molecular Structure and Energetics, VCH Publishers, New York, 1987 (Chapter 2).

[26] G. Edwards, Trans. Faraday Soc. 49 (1953) 152.