

# An ab Initio Study of Observed and Hypothetical Polymorphs of Glycine

James A. Chisholm,<sup>\*,†</sup> Sam Motherwell,<sup>†</sup> Paul R. Tulip,<sup>‡</sup> Simon Parsons,<sup>§</sup> and Stewart J. Clark<sup>‡</sup>

Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, U.K., Pfizer Institute for Pharmaceutical Materials Science, Department of Materials Science & Metallurgy, University of Cambridge, Pembroke Street, Cambridge, CB2 3QZ, U.K., Department of Physics, University of Durham, South Road, Durham DH1 3LE, U.K., and Department of Chemistry, The University of Edinburgh, Kings Buildings, Mayfield Road, Edinburgh EH9 3JJ, U.K.

Received November 4, 2004

**ABSTRACT:** To investigate the merits of crystal structure prediction using ab initio computational techniques, we have used density functional (DFT) methods to investigate the relative stabilities of the four known crystalline phases of glycine and also a range of alternative putative crystal structures of the zwitterion. Energy differences are calculated using a range of exchange-correlation functionals, and it is found that the calculated relative stability of the phases is sensitive to the choice of functional. Energy differences are found to be on the order of a few tenths of a kilocalorie per mole with little separation in energy found between observed and putative structures. This result is similar to that typically obtained from force field calculations and confirms the difficulty of the task of predicting the structure of molecular crystals. Optimization of structures, including optimization of unit cells, highlights the limitation of DFT in describing the long-range dispersion interaction. Use of the local density approximation (LDA) is shown to over-bind crystals, whereas use of gradient-corrected functionals severely under-binds crystals. Calculated structural energy differences are presented, which show that, for the case of the LDA, the four observed glycine polymorphs receive a lower energy than all putative glycine structures considered.

## Introduction

The ability to predict the structure of molecular crystals computationally given only a molecular diagram is highly desirable. For example, for the pharmaceutical industry, the discovery of new polymorphs of a drug molecule is of great importance as these may display superior properties and can lead to new patents. If theoretical methods were able to predict possible polymorphs and their properties with a sufficient level of confidence then theory could be used to guide experiment to speed the drug development process.

It is also important to obtain a complete understanding of the thermodynamic stability of a drug product prior to release. The ability to predict the most thermodynamically stable form could in some cases provide a valuable warning that the developed form is not the most stable and that phase transformation, say during storage, is a possibility.

Much crystal structure prediction (CSP) work has focused on the use of empirical force fields to rank large lists of generated candidate structures in terms of minimized lattice energy or minimized free energy.<sup>1-3</sup> Recent blind tests of CSP have shown that such approaches have, so far, met with only limited success and it is still not possible to predict structures with confidence.<sup>4,5</sup>

A major difficulty in the prediction of molecular crystals is the large number of plausible candidate

structures that can be generated with very similar low lattice energy. A typical result occurs that several unobserved structures can be found with lower calculated lattice energy than the observed structure(s). For CSP to be successful it is important to gain an accurate picture of the relative stability of these putative structures. Unfortunately, evaluating energy differences to the required accuracy poses a great challenge for many theoretical methods.

An in depth assessment of the performance of lattice energy minimization techniques has been carried out by Day<sup>6</sup> using the Williams force field. Prediction runs were carried out for 50 small, rigid organic molecules and it was found that for around half of these the experimentally observed structure could be located in the lowest 5 predicted structures, or within 0.2 kcal/mol of the global minimum in lattice energy.

The aim of this work is to investigate the performance of density functional theory (DFT) for CSP work and whether such methods can be expected to provide a greater predictive capability compared to force field methods. Assessing whether energies obtained from DFT calculations represent an improvement or otherwise is difficult as benchmark values for actual structural energy differences of various candidate structures for a given material are not available. However we proceed by addressing the following questions.

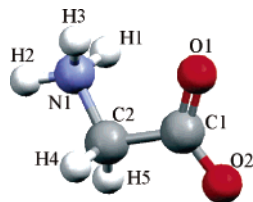
1. Do DFT energies predict the correct relative stability of the experimentally observed polymorphs of a molecular crystal?
2. Do DFT energies improve the energy ranking of the observed polymorphs in the energy ranked list compared to force field methods?
3. Do DFT energies agree with the common result found from force field calculations: that several putative

\* To whom correspondence should be addressed. Tel: +44 (0)1223 763921. Fax: +44 (0)1223 336063. E-mail: chisholm@ccdc.cam.ac.uk.

<sup>†</sup> Cambridge Crystallographic Data Centre and University of Cambridge.

<sup>‡</sup> University of Durham.

<sup>§</sup> The University of Edinburgh.



**Figure 1.** Atom numbering for the glycine zwitterion.

unobserved structures can be found with similarly low energy, separated by only a few tenths of a kcal/mol?

As an example for such a study, we have chosen to examine crystal structures of glycine. Glycine is known to crystallize in one of four forms in the space groups  $P21/n$ ,<sup>7</sup>  $P21$ ,<sup>8</sup>  $P32$ <sup>9</sup> and  $P21/n$ <sup>10</sup> which are labeled  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\delta$  respectively. The glycine molecule is pictured in Figure 1 where we define our notation for labeling the atoms. The  $\alpha$ ,  $\beta$  and  $\gamma$  phases are found at ambient pressure, with  $\alpha$  and  $\beta$  phases being metastable with respect to the  $\gamma$  phase.  $\delta$  glycine has recently been found to form under application of pressure,<sup>10</sup> the details of such pressure studies can be found elsewhere.<sup>11</sup> In the gaseous phase, glycine is in a nonionic form, while in all four of the crystal structures glycine is zwitterionic. Although zwitterionization causes an increase in energy with respect to the gas-phase molecule, it is thought that the zwitterionic crystals are stabilized by the increase in number of hydrogen bonds that can be formed in comparison to the number that would be formed in the nonionic case.

Our approach is to take the lowest energy putative glycine structures obtained from a crystal structure prediction performed using force field methods and re-optimize structures and recalculate relative energies using DFT. The force field prediction was performed by Day<sup>10</sup> who employed the Williams force field and an electrostatic model using multipoles as opposed to the commonly used point charge approximation. In that study the observed forms of glycine were located at positions 1( $\gamma$ ), 3( $\alpha$ ), 7( $\beta$ ) and 38( $\delta$ ) in the energy ranked list.

### Computational Details and Discussion

DFT calculations are performed using the plane-wave pseudopotential method as implemented in the Castep code.<sup>12</sup> We use Vanderbilt ultra-soft pseudopotentials<sup>13</sup> and a basis set cutoff energy of 450 eV. Brillouin zone integrations are performed on a symmetrized Monkhorst–Pack<sup>14</sup> k-point grid with a k-point spacing of 0.08 Å<sup>-1</sup>. Both k-point spacing and basis set cutoff are chosen to ensure total energies are converged to better than 0.1 meV/cell. The electronic structure is calculated by minimizing the total energy of the system using a preconditioned conjugate gradients algorithm. We note the use of plane waves avoids the basis set superposition error which is present for atom centered basis sets.

For all structures studied, both experimental and putative, we optimize structures fully allowing relaxation of unit cells and all atomic positions using a Broyden–Fletcher–Goldfarb–Shanno optimization scheme. Energy differences between structures can be extremely small, of the order of a few tenths of a kilocalorie per mole, and small changes in atomic

positions can give rise to significant changes in energy. Therefore it is important use high convergence criterion to ensure structural energy differences have converged. Structures are optimized until total energies have converged to within  $2 \times 10^{-5}$  eV/atom (0.005 kcal/mol of glycine molecules). In addition, forces are converged to a tolerance of  $2.5 \times 10^{-2}$  eV/Å, atomic displacements are converged to 0.01 Å and stresses are converged to 0.1 GPa.

The ability to fully optimize atomic positions, allowing relaxation of bond lengths and angles etc, represents one advantage over the use of force field methods where the treatment of flexible molecules can be difficult due to the partitioning of energy between inter and intramolecular components. However, the task of optimizing several structures in a prediction run is computationally intensive. Several geometry optimization steps may be required and for each step a self-consistent-field cycle is performed to recalculate the electronic structure. The cost of calculating the electronic structure typically scales as  $N^2 \log N$  where  $N$  is the number of atoms. A typical optimization for one glycine structure, containing 40 atoms in the unit cell, took 2 days on a 1.8 GHz Pentium 4 processor. A similar optimization using force field methods, in which the molecule conformation is held fixed and an electrostatic model using multipoles is used, takes less than 10 s.

An important question concerns the suitability of DFT methods for describing weakly bound molecular systems. DFT can be expected to provide a good description of the electrostatic, induction and the exchange-repulsion contributions to the total energy compared to empirical force field methods. However, a known limitation of DFT is its ability to faithfully model the long-range dispersion interaction. In the Kohn–Sham formalism<sup>15,16</sup> the exact form of the exchange-correlation functional is not known and approximate functionals must be used. Standard functionals include those that implement the local density approximation (LDA), which is exact for the case of a uniform electron gas, and the generalized gradient approximation (GGA), where account is taken of the charge density gradient at a point. Such local and semi-local functionals do not reliably describe long-range vdW energies. Methods such as 2nd order Moller–Plesset or coupled-cluster methods are often required to provide a higher level treatment of correlation and a more accurate description of dispersion. Such methods are computationally very intensive and it is of great interest to determine how well DFT, which includes some account of correlation, can perform for the task of CSP.

Several studies have sought to assess the performance of various exchange-correlation functionals for the description of weakly bound systems. Studies on rare gas diatomics have shown that use of the LDA leads to an overestimate in the strength of vdW forces while gradient corrected functionals can underestimate these forces.<sup>17–20</sup> This behavior holds also for the case of molecular crystalline systems. For example, in a recent study on polyethylene<sup>21</sup> the functional of Becke–Perdew (BP) was found to lead to extremely weak binding between molecules. The functional by Perdew, Becke and Ernzerhof (PBE) gave slightly better binding but still overestimated cell volumes by 22%. On the other

**Table 1. Unit Cell Parameters Calculated for  $\alpha$ ,  $\beta$ ,  $\gamma$ , and  $\delta$  Glycine Using Exchange-Correlation Functionals PW91, PBE, and LDA**

	$a$ (Å)	$b$ (Å)	$c$ (Å)	$\beta$ angle (deg)	vol (Å <sup>3</sup> )	no. of molecules	vol per molecule (Å <sup>3</sup> )
$\alpha$ -Glycine							
PW91	5.198	12.649	5.436	109.58	336.7	4	84.2
PBE	5.154	12.424	5.437	109.90	327.3	4	81.8
LDA	4.961	11.340	5.357	112.08	279.3	4	69.8
expt <sup>a</sup>	5.083	11.820	5.460	111.93	304.2	4	76.0
$\beta$ -Glycine							
PW91	5.110	6.448	5.389	111.34	165.3	2	82.6
PBE	5.114	6.465	5.392	111.41	166.0	2	83.0
LDA	4.927	5.823	5.326	113.73	139.9	2	70.0
expt <sup>b</sup>	5.094	6.286	5.383	113.21	158.44	2	79.2
$\gamma$ -Glycine							
PW91	7.213	7.213	5.483	90	247.0	3	82.3
PBE	7.213	7.213	5.484	90	247.1	3	82.4
LDA	6.759	6.759	5.394	90	213.4	3	71.1
expt <sup>c</sup>	6.975	6.975	5.437	90	230.6	3	76.9
$\delta$ -Glycine							
PW91	9.699	6.693	5.407	101.00	344.6	4	86.1
PBE	9.695	6.753	5.415	101.12	347.9	4	87.0
LDA	9.077	5.831	5.302	98.60	277.5	4	69.4
expt <sup>d</sup>	9.689	6.723	5.346	96.19	346.2	4	86.6

<sup>a</sup> Reference 25. <sup>b</sup> Reference 26. <sup>c</sup> Reference 27. <sup>d</sup> Reference 10.

hand, use of the LDA was shown to over-bind molecules. In recent work Tsuzuki and Luthi calculated interaction energies for vdW and hydrogen-bonded systems to assess the performance of the Perdew–Wang 91 functional (PW91).<sup>22</sup> This gradient corrected functional was shown to underestimate the strength of binding between molecules for all systems studied, although the performance of PW91 was shown to improve for the case of the hydrogen bonded systems.

It is clear then that DFT must be used with caution and that results obtained for weakly bound crystalline systems can be sensitive to the choice of exchange-correlation functional. The question for the present study is to what extent this limitation of DFT, or rather the limitation of the functionals, affects the reliability of calculated structural energy differences and thus the results of a crystal structure prediction. To help investigate this question we perform calculations using three functionals: LDA, PW91 and PBE. It is worth noting that the glycine molecule in the crystal is zwitterionic and the electrostatic interactions can be expected to be an important component of the lattice energy. DFT may be better suited for the prediction of a molecule like glycine than say for a hydrocarbon molecule.

DFT methods have been employed in previous work to study polymorphic forms of molecular crystals. Freeman et al employed the DMOL program and the functional of Becke–Lee–Yang–Parr to study the observed polymorphs of glycine.<sup>23</sup> In that study unit cells were constrained to experimental values.  $\alpha$  glycine was found to be the most stable form. Energy differences of +2.3 and +2.0 kcal were found for  $\beta$  and  $\gamma$  glycine, respectively. Rovira and Novoa studied the relative energies of several low energy polymorphic forms of acetic acid using both the GROMOS force field and a first-principles plane-wave method.<sup>24</sup> In that study the BP functional was employed and the problem of dispersion was addressed by the addition of a simple correction factor that scaled as  $r^{-6}$ . It is encouraging to see that

**Table 2. Relative Energies of Polymorphs of Glycine in kcal/mol Calculated with PW91, PBE, and LDA Exchange-Correlation Functionals<sup>a</sup>**

polymorph	PW91	PBE	LDA
$\alpha$	0 (0)	0	+0.30
$\beta$	+0.18 (+0.23)	+0.16	+0.62
$\gamma$	+0.21 (+0.21)	+0.16	0
$\delta$	+0.32	+0.25	+0.74

<sup>a</sup> Energies are expressed relative to the lowest energy structure, which is taken as a zero energy reference level. The values in parentheses are determined using norm-conserving pseudopotentials using a plane wave basis set cutoff energy of 800 eV.

**Table 3. Mulliken Charges Calculated for the Observed Phases of Glycine and a Nonzwitterionic Glycine Molecule in the Gas Phase<sup>a</sup>**

	$\alpha$	$\beta$	$\gamma$	$\delta$	mol
C1	0.64	0.65	0.66	0.66	0.68
C2	-0.50	-0.49	-0.50	-0.50	-0.56
H1	0.44	0.45	0.44	0.44	0.53
H2	0.45	0.44	0.44	0.46	0.47
H3	0.45	0.47	0.45	0.44	0.46
H4	0.35	0.34	0.34	0.35	0.36
H5	0.32	0.33	0.34	0.33	0.35
N	-0.83	-0.84	-0.82	-0.83	-1.00
O1	-0.66	-0.67	-0.67	-0.66	-0.59
O2	-0.67	-0.68	-0.67	-0.68	-0.69

<sup>a</sup> Atoms are labelled according to the scheme shown in Figure 1. Charges are expressed in units of the electronic charge,  $|e|$ .

the low-pressure experimental acetic acid structure is identified as the lowest energy structure in the DFT energy ranked list. Also, the relative ordering of the polymorphs was found to be insensitive to the inclusion of the dispersion correction term and to the use of alternative exchange-correlation functionals.

Another difficulty in the use of DFT for CSP work concerns the evaluation of the free energy. Due to the computational effort it is extremely time-consuming to evaluate the vibrational spectra and obtain the zero point vibrational energy (ZPVE) and vibrational entropy. Naturally such contributions may play a significant role in determining relative stabilities at a given temperature. However, in this work we present zero-temperature results without the ZPVE as an initial investigation as to the feasibility of DFT methods for CSP work.

## Results and Discussion

Lattice parameters calculated at ambient pressure for the four known phases of glycine are presented in Table 1. Also listed are experimental values determined at 120 K (neutron diffraction) for  $\alpha$ ,<sup>25</sup> 298 K (X-ray) for  $\beta$ <sup>26</sup> and 83 K (neutron diffraction) for  $\gamma$ .<sup>27</sup> Calculated volumes are found to be sensitive to the choice of exchange-correlation functional. Volumes calculated with LDA are lower than the finite temperature experimental values by 8%, 12%, 7% and 19% for  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\delta$  glycine, respectively. This is consistent with an overestimation of the weak intermolecular interactions by the LDA. We compare this result with recent findings by Montanari et al where separations between polyethylene chains calculated with LDA were found to be  $\sim$ 9% less than the experimental value.<sup>21</sup> Cell parameters calculated using PW91 are found to be similar to PBE values. In contrast to the LDA, PW91 and PBE predict cell volumes that are larger than experimental values

**Table 4. Optimized Bond Lengths and Torsion Angles for  $\alpha$ ,  $\beta$ , and  $\gamma$  Glycine Calculated Using PW91, PBE, and LDA<sup>a</sup>**

bonds	$\alpha$ -glycine				$\beta$ -glycine				$\gamma$ -glycine			
	PW91	PBE	LDA	expt	PW91	PBE	LDA	expt	PW91	PBE	LDA	expt
C4–O1	1.26	1.26	1.26	1.25	1.26	1.26	1.27	1.25	1.27	1.27	1.27	1.25
C4–O2	1.26	1.26	1.27	1.25	1.26	1.26	1.26	1.25	1.26	1.26	1.26	1.24
C5–N3	1.47	1.47	1.46	1.48	1.47	1.47	1.45	1.48	1.47	1.46	1.45	1.42
C5–C4	1.51	1.51	1.50	1.52	1.51	1.52	1.50	1.53	1.51	1.51	1.50	1.53
H6–N3	1.07	1.07	1.09	1.05	1.07	1.08	1.09	0.90	1.07	1.07	1.08	1.01
H7–N3	1.04	1.05	1.08	1.04	1.06	1.06	1.08	0.88	1.06	1.04	1.08	1.04
H8–N3	1.05	1.04	1.06	1.03	1.04	1.04	1.05	0.90	1.04	1.06	1.06	1.04
H9–C5	1.09	1.09	1.11	1.09	1.09	1.10	1.11	0.95	1.09	1.09	1.11	1.09
H10–C5	1.09	1.09	1.10	1.09	1.09	1.09	1.10	0.97	1.09	1.09	1.11	1.09

torsion	$\alpha$ -glycine				$\beta$ -glycine				$\gamma$ -glycine			
	PW91	PBE	LDA	expt	PW91	PBE	LDA	expt	PW91	PBE	LDA	expt
N3–C5–C4–O1	32.3	30.1	24.3	18.6	29.1	28.0	24.2	24.8	17.2	17.2	15.3	12.8

<sup>a</sup> Bond lengths are in Å, and torsion angles are in deg.

by between 5 and 10%. This result points to an underestimation of weak intermolecular interactions by both PW91 and PBE. Thus all functionals are found to provide poor predictions of glycine crystal densities.

In contrast, the structure of the glycine molecule is well reproduced by all functionals. Molecules obtained from optimized structures can be superimposed with molecules obtained from experiment to produce rms deviation values that are less than 0.1 Å. Calculated bond lengths are typically within 1% of finite temperature experimental values as shown in Table 4. Glycine is a flexible molecule and adopts a different conformation in each polymorph. This difference can be seen in the variation of the N3–C5–C4–O1 torsion angle which is highest in  $\beta$ -glycine (24.8°) and smallest in  $\gamma$ -glycine (12.8°). The present calculations are found to reproduce this change in torsion angle across the polymorphs (Table 4) although GGA values appear somewhat too large.

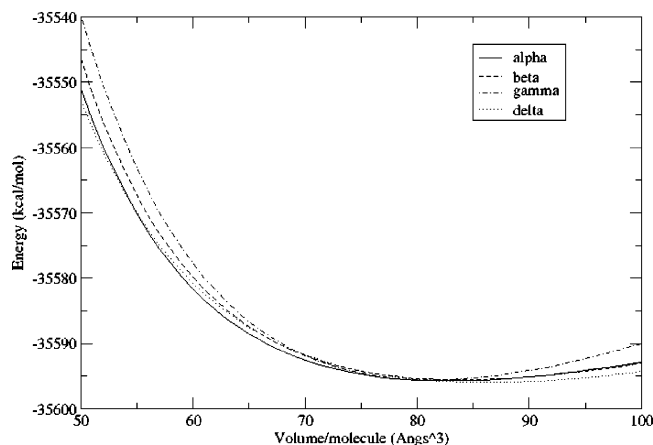
**Relative Stability of Observed Glycine Polymorphs.** The relative stability of glycine polymorphs has been addressed in recent thermo-chemical studies by Perlovich<sup>26</sup> and Boldyreva.<sup>28</sup> Results based on differences in solution enthalpies show the relative thermodynamic stability to be in the order  $\gamma$ – $\alpha$ – $\beta$  (most to least stable). Structural energy differences calculated using LDA, PW91 and PBE are presented in Table 2. We note two main points. First, calculated energy differences are very small, less than 1 kcal/mol, and are significantly smaller than previous theoretical estimations.<sup>23,29</sup> The small energy differences confirm the difficulty of predicting relative stabilities. The second point is that, similar to the prediction of crystal densities, calculated energy differences are also sensitive to the choice of exchange-correlation functional. Results obtained using PW91 and PBE are found to be similar. Both PW91 and PBE functionals predict  $\alpha$ -glycine to have the lowest energy with  $\gamma$  and  $\beta$ -glycine having energies around 0.2 kcal/mol higher than  $\alpha$ -glycine. Energies for  $\delta$ -glycine, the high pressure form, are predicted to be 0.32 and 0.25 kcal/mol higher than the  $\alpha$  glycine using PW91 and PBE, respectively.

It is noted that further calculations were performed to test the effect of the choice of pseudopotential. Energy differences obtained using norm-conserving pseudopotentials using a plane wave cutoff at 800 eV and the PW91 functional are shown in parentheses in Table 2. Energy differences are within 0.05 kcal/mol of values obtained using ultrasoft pseudopotentials.

Structural energy differences obtained using LDA differ from the PW91 and PBE results. LDA values predict the relative stability to be in the order:  $\gamma$ ,  $\alpha$ ,  $\beta$  then  $\delta$  (most to least stable). Thus LDA results do reproduce the relative stability of  $\alpha$ ,  $\beta$  and  $\gamma$  glycine found from experiment. This apparent success may be due to cancellation of errors when determining energy differences as the LDA is known to provide a poor description of hydrogen-bonded systems. Also the LDA predicts a greater separation in energy between polymorphs: 0.3 kcal/mol ( $\gamma$  to  $\alpha$ ) and 0.6 kcal/mol ( $\gamma$  to  $\beta$ ) compared to PW91 and PBE results.

The effects of pressure on these systems can also be important, noting that pressure is used to form the  $\delta$ -phase of glycine. Calculations were performed for the  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\delta$  phases of glycine, using the PW91 functional, for a range of external pressures applied to the computational cell. The resulting changes in volumes are shown in Figure 2. Curves using other GGA functionals are found to be very similar. The minimum of these curves show the relative ambient (zero) pressure energies, while the (negative) volume derivative shows the volume response to pressure. It can be seen that at ambient pressure the four phases are extremely close in energy as can also be seen from Table 2. As the pressure is increased, moving to smaller volumes, it can be seen that the  $\delta$  phase becomes the favored structure. The calculated pressure of transition is 2.4 GPa whereas the experimental value is only roughly known at 0.8 GPa.<sup>30</sup>

**Mulliken Charges.** Results of a Mulliken population analysis performed for a gas-phase glycine molecule and for  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\delta$  glycine are presented in Table 3. As the Mulliken analysis is formulated in terms of atomic orbitals,<sup>31</sup> we first project the electronic wave function from a plane wave to an atomic orbitals basis set using the scheme of Sanchez–Portal.<sup>32</sup> Charges are found to vary by only a small amount between the glycine polymorphs (around 0.01|e|). A more noticeable variation in the charge distribution can be seen between the crystalline molecules, which are zwitterionic, and the gas-phase molecule. For example, the charge on the nitrogen is found to be  $-1.00$  in the gas-phase molecule and  $-0.83$  in the crystal. It should be noted that the particular gas-phase conformation in Table 3 is that obtained by optimizing a glycine zwitterion. During optimization a proton transfers from the nitrogen to the nearest oxygen atom to form a more stable nonionic



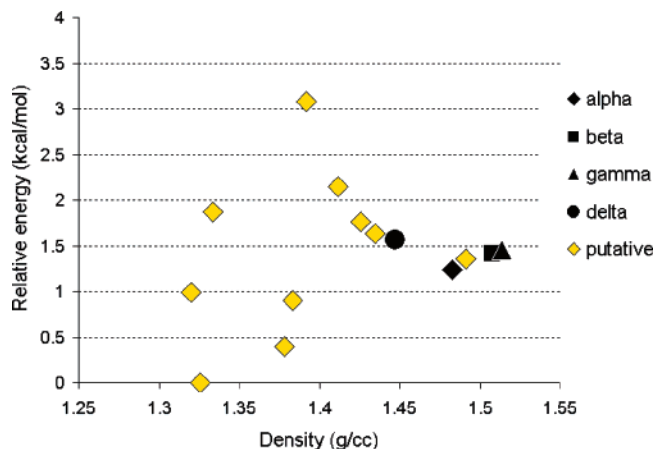
**Figure 2.** Energy as a function of volume for the four known phases of glycine calculated with the PW91 functional.

molecule (no search was made for the most stable conformation).

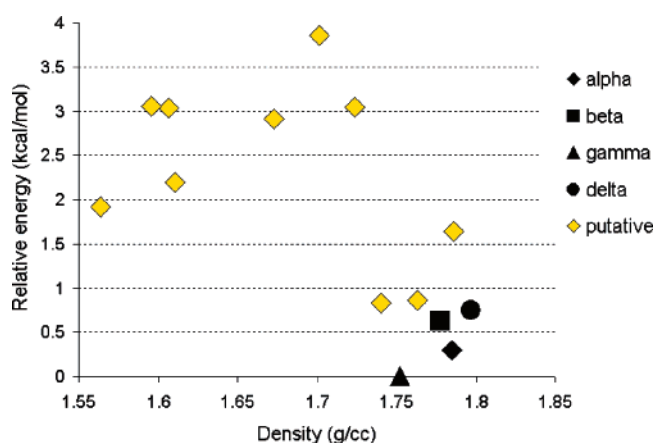
Such variation in charge distribution from the free molecule to the crystal highlights the potential problems and uncertainties associated with the use of force field charges which are typically derived to best reproduce the electrostatic potential of a particular isolated conformer. Electronic structure calculations provide a natural way to account for any significant variation in charge distribution that may result due to changes in molecular conformation or due to the influence of the crystal field, for example the formation of intermolecular hydrogen bonds.

**Relative Stability of Putative Glycine Polymorphs.** As mentioned, putative structures are obtained from a CSP run performed using the Williams force field.<sup>33</sup> Due to computational effort required we limit the DFT study to the four observed polymorphs plus the 10 lowest energy putative structures. We are interested to see if the ranking of the observed forms using DFT improves compared to the force field ranking and whether the energy of putative and observed structures are found within a similarly small energy range. Calculations on putative structures are performed with PW91 and LDA but not with PBE due to the computational effort required.

Structural energy differences calculated using PW91 are shown in Figure 3. Observed and putative structures are found to have very similar energies. All structures are contained in an narrow energy bracket of around 3 kcal/mol which is a little smaller than the 4.8 kcal/mol energy bracket obtained using force fields. In addition, several putative structures are found to have calculated energies lower than the observed forms. The energy ranked positions of the observed glycine polymorphs obtained using PW91 is as follows:  $\alpha$ -glycine is at rank 5,  $\beta$ -glycine rank 7,  $\gamma$ -glycine rank 8 and  $\delta$ -glycine is at rank 9. Thus a prediction, using PW91, does not improve the rank of the observed polymorphs compared to the force field results. For the CSP blind tests, participants are allowed to submit three structures for each blind test molecule.<sup>4</sup> If such a rule were followed here and the lowest three energy structures submitted, the PW91 would not lead to a successful prediction. Figure 3 also shows the densities calculated for each structure. It is found that the spread in



**Figure 3.** Structural energy differences of observed and putative glycine structures calculated using the PW91 functional.



**Figure 4.** Structural energy differences of observed and putative glycine structures calculated using the LDA.

densities predicted by PW91 is large (1.32–1.51 g/cm<sup>3</sup>) and that the observed glycine polymorphs are among the highest predicted densities.

Figure 4 shows structural energy differences calculated using the LDA. A different relative stability is found from PW91 results. The observed glycine polymorphs are now located in the following locations:  $\gamma$ -glycine is at rank 1,  $\alpha$ -glycine rank 2,  $\beta$ -glycine rank 3 and  $\delta$ -glycine is at rank 4 in the energy ranked list. This represents an improvement in the energy ranking compared to force field results and, using the basic approach of picking the lowest energy structures, the LDA prediction would be considered successful, although of course the prediction of crystal densities is poor. The apparent success of the LDA energy ranking may be due to the cancellation of errors when taking energy differences.

The difference in the strength of binding between LDA and PW91 can be seen in the predicted densities shown in Figure 3 and Figure 4. As discussed previously, the LDA predicts a much stronger binding between molecules compared to PW91. A consequence of this is that those structures that have a greater molecular packing efficiency are able to gain more energy compared to structures with a less efficient packing. Thus the energy ranking calculated using the LDA, shown in Figure 4, reflects the molecular packing

efficiency to a greater extent than PW91 energies with higher density structures having lower energies. Efficiency of packing has long been considered an important factor in determining the form of molecular crystals and this may also explain the apparent success of the LDA prediction.

### Conclusions

Density functional theory calculations have been performed for the four observed polymorphs of glycine and a range of putative low energy glycine structures to investigate the applicability of DFT for crystal structure prediction work. Structural data has been calculated for the four known forms of glycine allowing full relaxation of unit cells and atomic positions. Consistent with many previous DFT studies, the strength of intermolecular interaction and predicted crystal densities are sensitive to the choice of exchange-correlation functional. PW91 and PBE functionals are found to overestimate cell volumes whereas the LDA underestimates cell volumes or over-bind the molecules. Likewise, the relative stabilities of the observed glycine polymorphs are also sensitive to the choice of functional. PW91 and PBE give similar values while LDA values differ significantly and predict the relative stability of glycine polymorphs to be in the order:  $\gamma$ ,  $\alpha$ ,  $\beta$ ,  $\delta$  with  $\gamma$ -glycine as the most stable. This ordering of total energies at 0 K using the LDA is in agreement with the relative thermodynamic stability found from recent thermo-chemical studies.

A study of several putative glycine structures has shown that, similar to the findings obtained using force field calculations, DFT predicts small structural energy differences between predicted structures that are of the order of a few tenths of a kcal/mol, thus confirming the difficulty of the task of CSP. The choice of exchange-correlation functional also affects the relative energies of putative structures. The ranking of structures using PW91 leads to an unsuccessful prediction with four putative structures having a lower energy than that of the lowest ranked PW91 observed polymorph. In contrast, results obtained using the LDA predict the four known polymorphs of glycine to have lower energies than all putative structures considered. Thus the LDA results, although producing a poor prediction of crystal densities, does improve the ranking of observed structures in the energy ranked list compared to force field results.

Results obtained from force field predictions are known to vary considerably depending on the choice of force field, the electrostatic model adopted and the model used to evaluate intramolecular energies. This study demonstrates that prediction results obtained using DFT methods also vary depending on the approximations made. However, the 'successful' prediction obtained in this work indicates that, despite limitations present in the method, there may be some promise for using DFT for CSP work in the future, especially if a more accurate description of dispersion and crystal densities can be achieved. Further work will seek to extend this study to further molecular systems.

**Acknowledgment.** J. Chisholm gratefully acknowledges Dr. Graeme Day for providing hypothetical glycine structures and the Pfizer Institute for Pharmaceutical Materials Science for providing funding.

### Note Added after ASAP Publication

An earlier version of this paper published ASAP on May 5, 2005, contained an incorrect value for pressure of transition on line 357. This value is corrected and ref 10 is updated in this new version posted May 20, 2005.

**Supporting Information Available:** Structural data in cif format for optimized glycine structures. Also, calculated structural energy differences are available in a plain text file. This material is available free of charge via the Internet at <http://pubs.acs.org>.

### References

- (1) Mooij, W. T. M.; van Eijck, B. P.; Price, S. L.; Verwer, P.; Kroon, J. *J. Comput. Chem.* **1998**, *19* (4), 459.
- (2) Gavezzotti, A. *Acta Crystallogr.* **1996**, *B52*, 201.
- (3) Beyer, T.; Day, G. M.; Price, S. L. *J. Am. Chem. Soc.* **2001**, *123*, 5086.
- (4) Lommerse, J. P. M.; Motherwell, W. D. S.; Ammon, H. L.; Dunitz, J. D.; Gavezzotti, A.; Hofmann, D. W. M.; Leusen, F. J. J.; Mooij, W. T. M.; Price, S. L.; Schweizer, B.; Schmidt, M. U.; van Eijck, B. P.; Verwer, P.; Williams, D. E. *Acta Crystallogr.* **2000**, *B56*, 697.
- (5) Motherwell, W. D. S.; Ammon, H. L.; Dunitz, J. D.; Dzyabchenko, A.; Erk, P.; Gavezzotti, A.; Hofmann, D. W. M.; Leusen, F. J. J.; Lommerse, J. P. M.; Mooij, W. T. M.; Price, S. L.; Scheraga, H.; Schweizer, B.; Schmidt, M. U.; van Eijck, B. P.; Verwer, P.; Williams, D. E. *Acta Crystallogr.* **2002**, *B58*, 647.
- (6) Day, G. M.; Chisholm, J.; Shan, N.; Motherwell, W. D. S.; Jones, W. *J. Cryst. Growth Des.* **2004**, *4*, 1327.
- (7) Marsh, R. E. *Acta Crystallogr.* **1958**, *11*, 654.
- (8) Iitaka, Y. *Acta Crystallogr.* **1960**, *13*, 35.
- (9) Iitaka, Y. *Acta Crystallogr.* **1961**, *14*, 1.
- (10) Dawson, A.; Allan, D. R.; Belmonte, S. A.; Clark, S. J.; David, W. I. F.; McGregor, P. A.; Parsons, S.; Pulham, C. R.; Sawyer, L. *Cryst. Growth Des.*, published online May 18, 2005, <http://dx.doi.org/10.1021/cg049716m>.
- (11) Allan, D. R.; Clark, S. J.; Vos, W.; Brugmans, M.; Ackland, G. *J. Phys. Rev. B* **1998**, *58*, No. R11809.
- (12) Segall, M. D.; Lindan, P. J. D.; Probert, M. J.; Pickard, C. J.; Hasnip, P. J.; Clark, S. J.; Payne, M. C. *J. Phys.: Condens. Matter* **2002**, *14*, 2717.
- (13) Vanderbilt, D. *Phys. Rev. B* **1990**, *41*, 7892.
- (14) Monkhorst, H. J.; Pack, J. D. *Phys. Rev. B* **1972**, *13*, 5188.
- (15) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *B136*, 864.
- (16) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *A140*, 1133.
- (17) Perez-Jorda, J. M.; San-Fabian, E.; Perez-Jimenez, A. *J. Chem. Phys.* **1999**, *110* (4), 1916.
- (18) Zhang, Y.; Pan, W.; Yang, W. *J. Chem. Phys.* **1997**, *107* (19), 7921.
- (19) Kristyán, S.; Pulay, P. *Chem. Phys. Lett.* **1994**, *229*, 175.
- (20) Pérez-Jordá, J. M.; Becke, A. D. *Chem. Phys. Lett.* **1995**, *233*, 134.
- (21) Montanari, B.; Ballone, P.; Jones, R. O. *J. Chem. Phys.* **1998**, *108* (16), 6947.
- (22) Tsuzuki, S.; Luthi, H. P. *J. Chem. Phys.* **2001**, *114* (9), 3949.
- (23) Freeman, C. M.; Andzelm, J. W.; Ewig, C. S.; Hill, J. R.; Delly, B. *Chem. Commun.* **1998**, 2455.
- (24) Rovira, C.; Novoa, J. J. *J. Phys. Chem. B* **2001**, *105*, 1710.
- (25) Legros, J. P.; Kvik, A. *Acta Crystallogr.* **1980**, *B36*, 3052.
- (26) Perlovich, G. L.; Hansen, L. K.; Brauer-Brandl, A. *J. Therm. Anal. Calorim.* **2001**, *66*, 699.
- (27) Kvik, A. *Acta Crystallogr.* **1980**, *B36*, 115.
- (28) Boldyreva, E. V.; Drebushchak, V. A.; Drebushchak, T. N.; Paukov, I. E.; Kovalevskaya, Y. A.; Shutova, E. S. *J. Therm. Anal. Calorim.* **2003**, *73*, 409.
- (29) Raabe, G. *Z. Naturforsch.* **2000**, *A55*, 609.
- (30) Parsons, S. Private communication, 2003, to be published.
- (31) Mulliken, R. S. *J. Chem. Phys.* **1955**, *23*, 1833.
- (32) Sanchez-Portal, D.; Artacho, E.; Soler, J. M. *Solid State Commun.* **1995**, *95*, 685.
- (33) Day, G. M.; Motherwell, W. D. S.; Jones, W. *Cryst. Growth Des.*, in press.