



**HAL**  
open science

# First principles determination of structural, electronic and lattice dynamical properties of a model dipeptide molecular crystal

Paul Robert Tulip, Simon Bates

► **To cite this version:**

Paul Robert Tulip, Simon Bates. First principles determination of structural, electronic and lattice dynamical properties of a model dipeptide molecular crystal. *Molecular Physics*, 2009, 107 (20), pp.2201-2212. 10.1080/00268970903224955 . hal-00521569

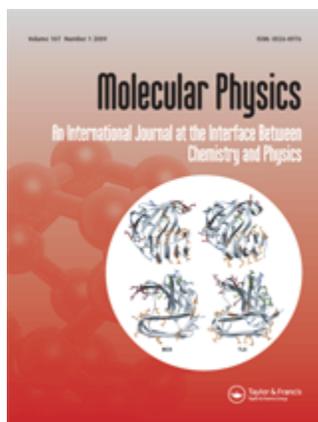
**HAL Id: hal-00521569**

**<https://hal.science/hal-00521569>**

Submitted on 28 Sep 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**First principles determination of structural, electronic and lattice dynamical properties of a model dipeptide molecular crystal**

Journal:	<i>Molecular Physics</i>
Manuscript ID:	TMPH-2009-0187.R1
Manuscript Type:	Full Paper
Date Submitted by the Author:	14-Jul-2009
Complete List of Authors:	Tulip, Paul; University of Edinburgh, School of Physics and Astronomy Bates, Simon; University of Edinburgh, School of Physics and Astronomy
Keywords:	ab initio, density functional theory, molecular crystals, lattice dynamics
<p>Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.</p> <p>tulip.zip</p>	



## RESEARCH ARTICLE

## First principles determination of structural, electronic and lattice dynamical properties of a model dipeptide molecular crystal.

P.R. Tulip<sup>a\*</sup> and S.P. Bates<sup>a</sup><sup>a</sup>School of Physics and Astronomy and SUPA, University of Edinburgh, Edinburgh, UK

(June 2009)

We present the results of *ab initio* calculations of the structural, electronic and lattice dynamical properties of the solid-state crystal of the glycyl-L-alanine dipeptide. Intramolecular bond lengths are found to be in good agreement with experimental values; lattice constants are in reasonable agreement, although it is found that discrepancies do exist. A hierarchy of hydrogen bond strengths is found, with those between (oppositely-charged) amine and carboxy functional groups being strongest. The crystal is found to be an indirect-bandgap material, with indirect bandgaps  $\approx 4.95$  eV, compared to a direct bandgap of 5.00 eV. Analysis of the electronic structure reveals that the electronic states in the near vicinity of the energy gap arise from carboxylate and amide oxygen atoms. The arrangement of both molecules and hydrogen bonds in the unit cell is found to manifest itself in increased bandwidth along specific reciprocal space directions, reflecting coupling brought about by hydrogen bonds. Determination of the zone-centre lattice dynamical behaviour permits the IR absorption spectrum to be explained. Intermolecular hydrogen bonds are found to couple intramolecular motions in adjacent molecules, revealing the importance of an accurate treatment of intermolecular interactions, even for high-frequency vibronic modes.

**Keywords:** *ab initio*; density functional theory; molecular crystals; lattice dynamics

## 1. Introduction

Recent years have witnessed sustained and intense research efforts into proteins and peptides, using a combination of experimental [1–3], theoretical [4–6] and computational methods [7–10]. There are several reasons for this interest: firstly, peptides, as the building blocks of proteins are of fundamental biological importance - successfully understanding the biological physics of these molecules will provide a major advance in our biological understanding. Secondly, misfolding and aggregation of peptides and proteins has been implicated as being of critical importance in understanding the pathologies of diseases such as Alzheimer's and Creutzfeldt-Jakob's [11, 12]. Promising potential treatment strategies for these diseases involve the development of peptidic inhibitors that may prevent, or reverse, aggregation [13–15].

There are, in addition, technological imperatives for this interest: hybrid bio-nanoparticle materials offer the promise of novel technological applications. For example, conjugation of peptides with metallic nano-clusters results in modification of optical, electronic and conformational properties [16–20]; peptide adsorption on metallic and mineral surfaces, meanwhile, is of relevance to the development of

---

\*Corresponding author. Email: ptulip@ph.ed.ac.uk

1 biosensors and bioarrays [21–24], whilst peptide-inorganic interfaces offer novel  
2 means of materials synthesis and nanoscale fabrication [25–28].

3 The common theme underpinning these disparate research avenues is the desire  
4 to develop, at the molecular level, a detailed and accurate description of the inter-  
5 molecular interactions responsible for these phenomena and to thereby elucidate  
6 the relationship between peptide structure and function. Computational modelling  
7 is an invaluable tool that may be used to effect this goal; valuable insights have  
8 been gained using a variety of modelling techniques, as examination of the relevant  
9 literature reveals [29–34].

10 One of the most powerful elements of the computational physicist's armoury is  
11 provided by quantum mechanical electronic structure calculations. The power of  
12 these methods is intimately related to the realisation that the key to understand-  
13 ing the properties of any condensed matter system lies within the electronic struc-  
14 ture: once the ground-state electronic structure has been accurately determined,  
15 this allows, in principle, a complete understanding of its physical properties. The  
16 Hohenberg-Kohn-Sham density functional theory (DFT) [35, 36] represents the  
17 state-of-the-art for such calculations. A long history of successful applications to a  
18 wide range of systems [37–41] attests to the power and flexibility of the method;  
19 and indeed, it is sufficiently powerful that it may be used as a predictive tool in its  
20 own right, rather than merely serving as a tool to confirm experimental results.

21 In this paper we make use of DFT calculations to investigate and elucidate the  
22 structural, electronic, dielectric and lattice dynamical properties of the solid state  
23 phase of the dipeptide glycyl-l-alanine. This dipeptide has recently been the sub-  
24 ject of computational and experimental work: the gas-phase structure and vibra-  
25 tional properties have been determined [42], whilst some attempt has been made  
26 to extend this work to incorporate the condensed phase through a dielectric con-  
27 tinuum treatment of solvent effects [42, 43]. The hydration structure of the peptide  
28 in water has been studied using both classical and *ab initio* molecular dynamics  
29 methods [44, 45]. More recently, the solution structure of the peptide in water at  
30 a 1:20 solute-solvent concentration has been investigated using neutron diffraction  
31 methods [46], whilst a recent study by the current authors [47] has addressed the  
32 nature of peptide aggregation in this system and elucidated the solvent structure  
33 using classical molecular dynamics simulations. The latter work has provided the  
34 stimulus for the current study: molecular crystals provide a well-defined system  
35 that may be analysed at a high level of theoretical treatment; in particular, the  
36 electronic structure itself can be investigated in some depth, allowing the nature  
37 of the intermolecular interactions to be elucidated. To the best of our knowledge,  
38 this is the first such systematic investigation of this dipeptide's solid state prop-  
39 erties; previous investigations have concerned specific topics such as NMR shielding  
40 tensors [48, 49]. **Indeed, there are relatively few works analysing the elec-  
41 tronic structure of biomolecular crystals in depth; however, meaningful  
42 exploitation of such systems demands that the electronic structure be  
43 understood in detail, and the relationship to the crystal and molecular  
44 structure be understood. The objective of this work is to provide such  
45 a characterisation for this particular peptide.**

46 The paper is structured as follows: in the next section we provide an overview  
47 of the theoretical and computational methods employed in this study, whilst in  
48 section 3 we discuss the structural properties. In section 4, we address the electronic  
49 structure. In section 5 we focus our attention upon the lattice dynamical and  
50 dielectric properties; finally, we conclude in section 6.

## 2. Computational Approach

We carry out our calculations within the plane wave pseudopotential (PWP) formalism of the Hohenberg-Kohn formulation of density functional theory (DFT) [35, 36]. This is a powerful technique for the accurate computation of ground state properties. Despite the existence of the well-known *bandgap problem* [50], *i.e.* that DFT commonly estimates incorrect bandgaps, the qualitative topology of band-structures obtained using DFT is generally considered to be correct, and we are therefore justified in using this concept to understand the electronic structure of the systems under consideration. The calculations are performed using the CASTEP [51] code, using norm-conserving pseudopotentials of the Kleinman-Bylander [52] form. The valence electron wavefunctions are expanded in a plane-wave basis set, with a kinetic energy cut-off of 1000 eV. Reciprocal space integrations have been carried out using a  $3 \times 3 \times 3$  Monkhorst-Pack grid [53]; this yields eight symmetry-reduced k-points, and converges total energies to better than 1 meV/atom, a degree of accuracy comparable to that with regard to kinetic energy cut-off. Exchange and correlation are treated using the well-known generalised gradient approximation (GGA) due to Perdew and Wang [54]. This is known to be capable of accurately treating hydrogen-bonded molecular systems [55, 56], **although it should be acknowledged that such semi-local functionals are intrinsically incapable of treating van der Waals interactions.**

Our calculations fall into several distinct stages: firstly we determine the geometrical structures adopted at equilibrium by minimising the forces acting on nuclei using the well-known Hellmann-Feynman theorem. This involves optimisation of all atomic coordinates, in addition to allowing the unit cell to relax. We take experimentally determined structures [57] as our initial atomic positions. **This approach amounts to identifying the minimum in the lattice energy. A large body of research into crystal structure prediction [58–60] indicates that such minimisation of the intermolecular potential energy represents a successful method for optimising crystal packing, and that density functional theory may be used to quantify such interactions, either directly, or through a DFT-derived forcefield.** Once we have optimised these structures, we then solve the single-particle Kohn-Sham equations at a series of reciprocal space points to determine the eigenvalue spectrum.

Having determined the structural and electronic properties, we determine the lattice dynamical and dielectric properties of the crystal using the variational formalism of density functional perturbation theory (DFPT) of Gonze and Lee [61], as implemented within the CASTEP code [51, 62]. DFPT is a powerful method capable of determining the lattice dynamical properties of the crystal via a harmonic treatment; as such, it is incapable of treating anharmonicities. However, it has been successfully applied to such organic molecular systems [56, 62–64], which justifies our confidence in the accuracy of the assignments presented here.

In comparison to frozen-phonon approaches, the DFPT approach to lattice dynamics possesses the advantage that the zone-centre excitations may be readily accessed; further, that because it is formulated at the linear response level, it is possible to compute the system response to a combined perturbation (e.g. a lattice perturbation and an applied field). This allows the macroscopic electric fields induced by long-wave atomic displacements to be naturally incorporated within the calculation. These fields lead to the so-called “LO-TO splitting” [65] of zone-centre modes. Successful treatment of this phenomenon is essential for an accurate description of the normal modes.

In this work, we concern ourselves solely with the determination of the vibra-

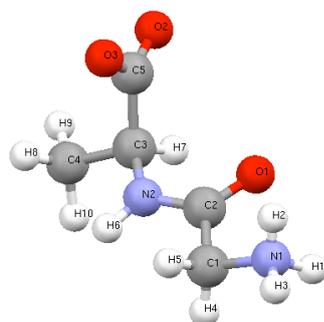


Figure 1. Structure of the glycyl-L-alanine molecule and the labelling scheme used in this work.

tional excitation spectrum at the Brillouin zone centre, as opposed to determining a full dispersion curve. The zone-centre phonons are of physical importance in that it is these modes that couple to electromagnetic radiation. Further, it should be noted that, whilst typically, DFPT is capable of describing vibrational frequencies to within  $50 \text{ cm}^{-1}$ , molecular crystals may prove to be demanding tests of the theory. In particular, the soft potential surfaces of such systems ensure that extremely stringent (and computationally demanding) convergence criteria must be employed to determine low-lying intermolecular modes accurately [66]. For this reason we focus our attention instead upon the intramolecular modes, which are less demanding to describe adequately.

It is convenient at this stage to consider the molecular structure and labelling scheme employed in this study. In Figure 1 the zwitterionic structure of the glycyl-L-alanine molecule is illustrated, along with the numbering convention used throughout.

### 3. Structural Properties

The initial crystal structure for the geometry optimisation was taken from the X-ray diffraction data of Wang and Paul [57] at room temperature. These data suggest that the crystal forms an orthorhombic structure, with space group  $P_{212121}$ , with four molecules per unit cell. In Figure 2 we present this structure. It can be seen that, if the structure is viewed along the  $a$ -axis, then the molecules adopt a herring-bone-type-layering structure, in which layers “stack” along the  $c$  axis. This maximises inter-molecular hydrogen bonding contacts. Our geometry optimisation calculations relax all internal degrees of freedom in addition to the unit cell parameters whilst maintaining a fixed system symmetry. In Table 1 the optimised unit cell parameters are listed. It can be seen that these are in good agreement with those obtained experimentally for both XC functionals employed (of the order of 2%). In the  $b$  and  $c$  directions, the optimised lattice constants are longer than the experimental values, which is consistent with the known tendency of the GGA to underestimate binding [41, 50]. In contrast, the  $a$ -axis is underestimated relative to experiment, suggesting instead a tendency to over-bind in this case. The experimental cell volume is  $695.79 \text{ \AA}^3$ . The PW91 and PBE functionals predict larger cell volumes at  $716.60 \text{ \AA}^3$  and  $719.88 \text{ \AA}^3$ , respectively. It should be noted that these calculations result in residual forces of less than  $1 \text{ meV/\AA}$  per atom, indicating that the optimised atomic locations are very close to the potential minimum. In addition, in contrast to the experiments, which obtain a structure at

Table 1. Experimental and theoretical lattice parameters as used in this study. Experimental values are obtained from ref. [57]. All values quoted are in Angstrom.

a			b			c		
PW91	PBE	Experimental	PW91	PBE	Experimental	PW91	PBE	Experimental
9.565	9.550	9.693	9.683	9.755	9.524	7.737	7.727	7.537

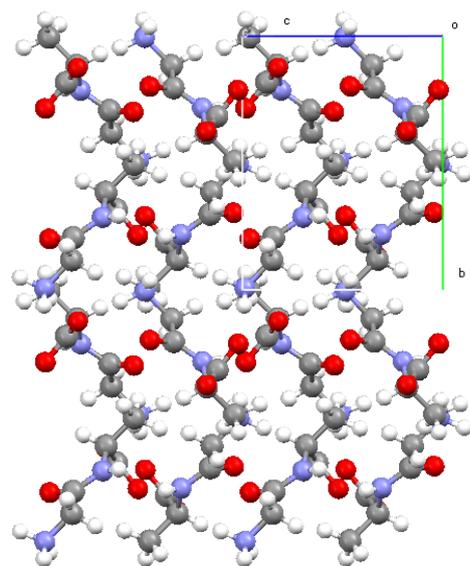
Table 2. Experimental and theoretical bond lengths. Experimental values are from reference [57]; again, all values are quoted in Angstrom.

	PW91	PBE	PBE [49]	B3LYP/6-31+G(d) [67]	Experimental
N1-C1	1.475	1.475		1.476	1.475
N1-H1	1.048	1.048			1.000
N1-H2	1.050	1.051			0.861
N1-H3	1.051	1.052			0.950
C1-C2	1.520	1.521		1.509	1.527
C1-H4	1.089	1.093			0.984
C1-H5	1.089	1.088			1.005
C2-O1	1.206	1.207	1.237	1.230	1.223
C2-N2	1.343	1.344		1.315	1.338
N2-C3	1.449	1.449		1.444	1.455
N2-H6	1.033	1.034		0.850	0.939
C3-C4	1.513	1.513		1.525	1.524
C3-H7	1.089	1.090			0.976
C4-H8	1.091	1.091			1.030
C4-H9	1.087	1.089			1.067
C4-H10	1.088	1.089			1.119
C3-C5	1.538	1.539		1.547	1.538
C5-O2	1.242	1.242	1.262	1.240	1.249
C5-O3	1.235	1.235	1.268	1.255	1.256

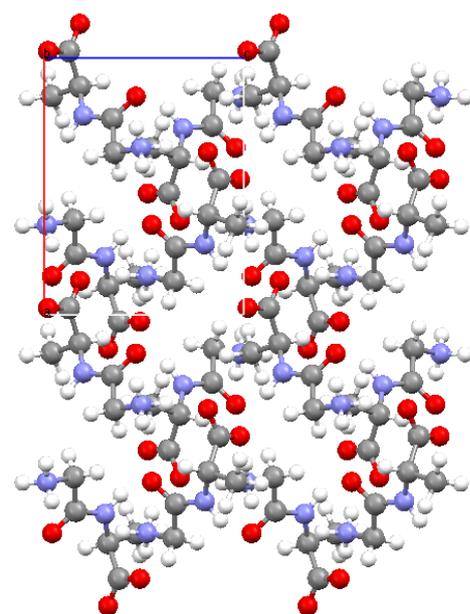
room temperature, the DFT calculations are zero-temperature calculations; it is conceivable, therefore, that temperature effects may also explain at least some of the small discrepancy observed. **Whilst such agreement is gratifying, it is likely that it is, to some extent, fortuitous. Semi-local XC functionals such as the GGA are incapable of treating van der Waals interactions; in molecular crystals, this may lead to inaccurate lattice constants [55]. It is likely that in this particular system, the zwitterionic nature of the glycyl-L-alanine molecule ensures that there is a significant electrostatic contribution to the intermolecular hydrogen bonds, and hence to the lattice energy, and thus the neglect of van der Waals interactions is not significant.**

In Tables 2 and 3, the quality of the intra-molecular structures obtained are compared with the experimental values, through comparison of bond lengths and dihedral angles. The values in Table 2 indicate extremely good agreement with the experimental results for the intramolecular bond lengths. Note, however, that there is markedly less agreement for bond lengths involving hydrogen atoms. It is likely that this reflects the well-known difficulties of accurately locating hydrogen atoms via X-ray diffraction, as opposed to arising from a systematic problem in the theoretical treatment. The quality of the calculations compares favourably with other theoretical values found in the literature and quoted in Table 2. **Optimisation of the crystal structure using the PBE functional [68] reveals that both functionals predict essentially identical bond lengths. This indicates, therefore, that the discrepancies that occur between our results and those of Zheng *et al.* [49] arise owing to numerical factors (choice of pseudopotential, kinetic energy cut-off, k-point sampling, or some combination thereof).**

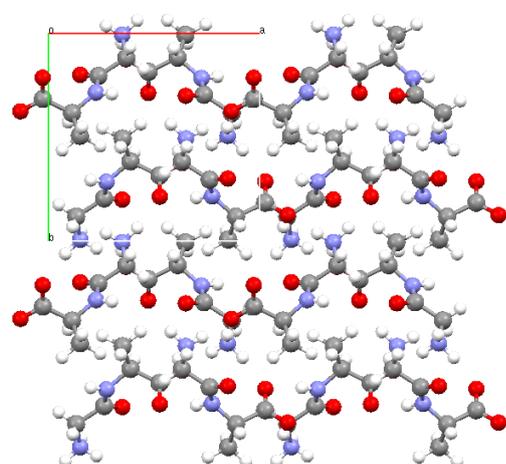
The calculated values of Kolev *et al.* [67] are obtained using DFT in conjunction with the B3LYP XC functional and the 6-31+G(d) basis set. It is difficult to tell



(a)



(b)



(c)

Figure 2. Crystal structure of the glycyl-L-alanine crystal. In each case a  $2 \times 2 \times 2$  supercell is presented in order to illustrate the crystal packing. In (a), the crystal is viewed along the  $a$  axis; in (b), along the  $b$  axis. In (c), it is viewed along the  $c$  axis.

Table 3. Experimental and theoretical dihedral angles. Experimental values are taken from reference [57]. All angles quoted are in degrees.

	PW91 (this study)	PBE (this study)	B3LYP/6-31+G(d) [67]	Experimental
$\omega_{C^\alpha-C'-N-C^\alpha}$	-172.37	-172.57	173.0	-173.94
$\phi_{C'-N-C^\alpha-C'}$	-81.63	-83.33		-78.36
$\psi_{N-C^\alpha-C'-N}$	-159.54	-160.49		-163.54

Table 4. Selected intermolecular contacts - comparison between experiment (reference [57]) and theoretical values obtained in this work. All lengths are quoted in Angstrom

	PW91 (this study)	PBE (this study)	Experimental
O2-H1	1.789	1.799	1.776
O2-H2	1.724	1.723	1.923
O3-H3	1.729	1.729	1.789
O3-H6	1.881	1.891	1.969

whether the discrepancies arise owing to differences in the functional used, or are basis-set artifacts; in addition, an ambiguity lies in these authors' work in that it is not clear as to whether their calculations refer to a gas-phase molecule, a periodic solid, or a cluster approximation to the crystal.

Having assessed the quality of the intra-molecular interactions through the bond lengths, in Table 3 we explore them further, through the determination of selected dihedrals. The calculated dihedrals are those commonly used to characterise the conformational preferences of peptides, namely  $\phi$ ,  $\psi$  and  $\omega$ . In common with the X-ray data, the calculations indicate that the molecules are in the *trans* state. **Agreement between experiment and theory for the dihedrals is poorer than that found for the bond lengths; similar results are found in gas-phase calculations.**

In Table 4, we examine the lengths of selected intermolecular contacts. It is noticeable that there is a variety of hydrogen bond lengths present, suggesting a hierarchy of hydrogen bond strengths; further, that each carboxylate oxygen atom can be seen to participate in two hydrogen bonds, i.e. it acts as a bifurcated acceptor. One of these bonds is an inter-layer bond, whilst the other is intra-layer in character. The herring-bone-like packing of the molecules is such that hydrogen bonding involving the O1 atom does not occur. This is because the molecules pack so as to achieve the maximum number of interactions between the oppositely-charged amine and carboxy groups, i.e. it is interactions between these groups that drive the formation of the crystal structure. It is of interest to note that a recent examination of the structure of the aqueous solution using classical molecular dynamics [47] obtained a similar result, in which molecular aggregation was dominated by charge-charge interactions. **Note that in some cases, disagreement between hydrogen bond lengths is marked. Some of this may be attributed to inaccuracies in the hydrogen positions as inferred from X-ray crystallography, whilst another contribution may be provided by the neglect of van der Waals interactions, as discussed earlier.**

Summarising, therefore, it seems that intramolecular bonding interactions are well-described by the XC functional employed in this study, with discrepancies in the crystal structure arising in the main from differences in the treatment of intermolecular hydrogen bonds. Having discussed the structural features of the crystal in some depth, we now focus upon the electronic structure, and attempt to elucidate the relationship between the two.

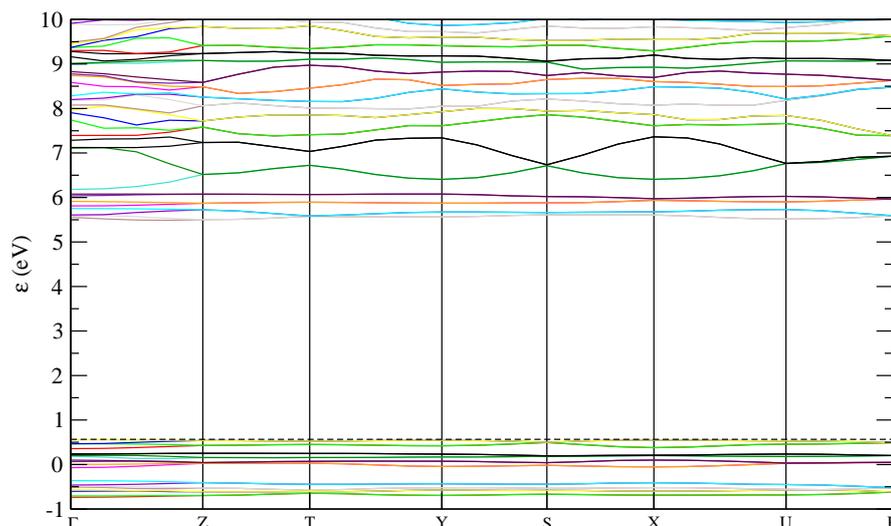


Figure 3. Electronic bandstructure of the glycyl-L-alanine crystal. This is a “zoomed-in” view, in which only the bands in the vicinity of the bandgap are displayed.

#### 4. Electronic Structure

In Figure 3 we present the electronic bandstructure. Note that, owing to the presence of four molecules per unit cell, each molecular orbital splits into four bands. Examination of the bandstructure reveals that the crystal has three indirect bandgaps, which are found to arise from transitions between a conduction band minimum at Z and valence band maxima at the  $\Gamma$ -point, Y and X. Although not corresponding to the true excitation energies, the Kohn-Sham eigenenergies may be taken as a zeroth order approximation to them [70]. Following this approach yields estimates for the indirect bandgaps of 4.95 eV, 4.96 eV and 4.95 eV, respectively. In comparison, the direct bandgap corresponding to  $\Gamma - \Gamma$  transitions is 5.00 eV. As DFT calculations often obtain underestimates of the bandgap [71], this suggests that the glycyl-L-alanine crystal possesses an insulating character. These values are in close agreement with existing theoretical estimates of the bandgap for both crystalline alanine [55] and glycine [71]. In the absence of experimental data with which to compare directly, we note that experimental optical absorption measurements indicate a bandgap of 5.11 eV for crystalline  $\alpha$ -glycine; it is perhaps possible to anticipate that the experimental value for glycyl-L-alanine is similar.

The valence bands displayed are extremely narrow, displaying little dispersion across the Brillouin zone, indicating that electronic orbitals are primarily localised on each molecule in the unit cell. Such localisation is consistent with hydrogen bonding being the dominant mechanism responsible for the crystal structure; similar findings have been reported concerning the electronic structure of amino acid molecular crystals [55, 71]. Inspection of the eigenvalues reveals that for the highest occupied valence bands the Davydov splitting (i.e.  $E_{HOMO} - E_{HOMO-3}$ ) arising

1 at the  $\Gamma$ -point is 0.20 eV; this is a relatively large splitting, and indeed is actually  
2 larger than the bandwidth arising from dispersion.  
3

4 Interestingly, whilst the lowest conduction bands are similarly flat, the bands  
5 lying between 6 and 7 eV show markedly more dispersion; this is characterised  
6 by a bandwidth of 0.70 eV. In particular, the dispersion is seen to be greatest  
7 along specific directions in reciprocal space, namely along the lines  $\Gamma \rightarrow Z$ ,  $Y \rightarrow S$ ,  
8  $S \rightarrow X$  and  $X \rightarrow U$ . This can be understood by reference to the crystal unit cell as  
9 follows: the first direction mentioned corresponds to dispersion along the  $c^*$ -axis;  
10 in real space, this is perpendicular to the  $ab$  plane. This suggests that dispersion  
11 in this direction arises owing to the coupling of electronic states on neighbouring  
12 molecules, which is likely to be due to the O2-H1 hydrogen bonds, which are roughly  
13 directed perpendicular to this plane. Similar explanations may be offered for the  
14 other directions: the second direction corresponds to dispersion along the  $b^*$ -axis,  
15 i.e. perpendicular to the  $ac$  plane; this is found to be due to O3-H3 hydrogen  
16 bonds. The third direction in which marked dispersion occurs is along the  $a^*$ -axis;  
17 this suggests that O2-H2 and O3-H6 hydrogen bonds are responsible for this. The  
18 magnitude of the Davydov splitting is also greater; at the  $\Gamma$ -point, this is 1.10 eV,  
19 compared to a splitting of approximately 0.2 eV for the lowest lying conduction  
20 bands at the zone centre. Finally, marked dispersion occurs along the  $c^*$ -axis, which  
21 as discussed earlier, corresponds to coupling occurring via O2-H1 hydrogen bonds.  
22

23 Determination of the density of states, along with its decomposition in terms of  
24 angular momentum states, permits further insights into the electronic structure.  
25 To that end, in Figure 4, this quantity is shown. It is evident that the bands in  
26 the vicinity of the bandgap (both valence and conduction) are primarily of  $p$ -type  
27 character, whilst conduction bands between about 6 and 7 eV are primarily of  
28  $s$ -type character. The small amount of hybridisation that this indicates for the  
29 highest occupied states is consistent with the inference that the electronic orbitals  
30 are localised, as discussed earlier in the context of the bandstructure plot.  
31

32 It is desirable to decompose these densities of states further, such that atomic  
33 contributions to the bandstructure may be obtained. In Figure 5, such local densi-  
34 ties of states are presented for each atom type present in the crystal. It is apparent  
35 that the dominant contribution to the highest occupied valence states is provided  
36 by oxygen  $p$ -states. Further decomposition of this quantity (not shown) reveals that  
37 the carboxylate and carbonyl oxygens provide a comparable contribution to these  
38 states. This is interesting to note, especially if one recalls that the latter do not  
39 participate in hydrogen bonding, and it could therefore perhaps be anticipated that  
40 they do not contribute substantially to intermolecular interactions. It is also found  
41 that the carbonyl oxygen contributes marginally more to the LUMO states than the  
42 carboxylate oxygens. Nitrogen and carbon atoms each provide a contribution that  
43 is between approximately one sixth and one fifth that of the oxygen atoms to the  
44 HOMO states, with the nitrogen providing slightly more states. Decomposing the  
45 nitrogen local density of states into both amine and amide nitrogen contributions  
46 (not shown), it is found that the amine nitrogen makes a negligible contribution  
47 to the bandstructure in the vicinity of the bandgap, consistent with recent work  
48 concerning the electronic structure of the glycine molecular crystal [71]. In contrast  
49 to the amine nitrogen atom, the amide nitrogen atom is found to provide a more  
50 substantial contribution to the electronic states near the bandgap (not shown). In  
51 common with the oxygen atoms, it provides considerably more (by a factor of 2)  
52 to the valence band states than the conduction band states. Whilst the HOMO  
53 states are primarily O  $2p$  in character, the LUMO states are revealed to be an  
54 admixture of  $p$  electrons from the oxygen, carbon and nitrogen atoms, with the  
55 dominant contribution arising from the carbons.  
56  
57  
58  
59  
60

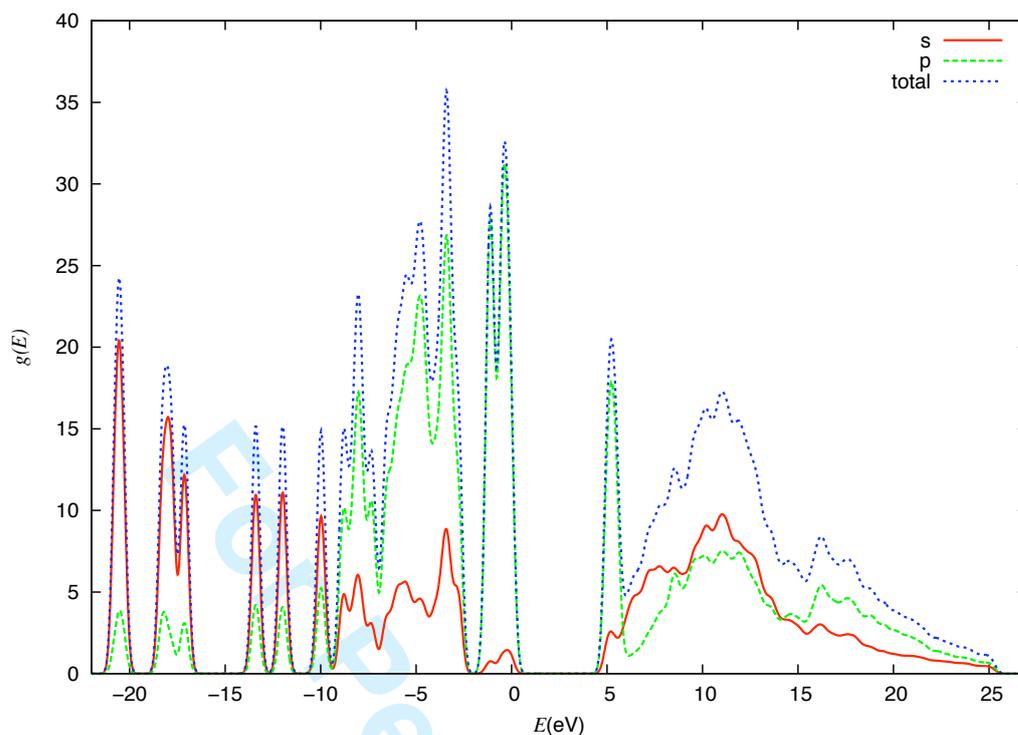


Figure 4. Density of states (DOS) for the glycyl-L-alanine crystal. Presented are both the total density of states and a decomposition in terms of angular momentum character. As in all of the DOS plots presented in this work, a Gaussian smearing of 0.2 eV has been applied.

Taken together with the partial density of states, these plots indicate that the major contributions to the band of dispersive states between 6 and 7 eV arise from carbon *p* and hydrogen *s*-electrons. Figure 4 demonstrates that both types contribute appreciably, suggesting a marked degree of *sp*-hybridisation. This suggests that these states possess a more covalent character than the lowest lying states, which is consistent with the greater dispersion observed in the bandstructure.

## 5. Lattice Dynamics

Prior to discussing the zone-centre lattice dynamics, it is worthwhile to briefly elucidate the solid-state vibrational selection rules. The glycyl-L-alanine crystal possesses 237 normal modes (i.e.  $3N-3$ , where 3 rigid translations must yield zero frequency by the acoustic sum rule); we note at this point that the degree to which the acoustic sum rule is satisfied provides some measure of the numerical precision and quality of the calculation. Our calculations satisfy this rule to within  $20 \text{ cm}^{-1}$ , and indeed no imaginary frequencies are obtained, indicating that our calculated structure is mechanically stable and located close to the potential minimum. The factor group of the glycyl-L-alanine crystal is  $D_2$ ; a factor group analysis therefore allows one to determine the transformation properties of the phonons. These are found to transform as

$$\Gamma = 60a + 59b_1 + 59b_2 + 59b_3 \quad (1)$$

where the *a* irrep is the only symmetry species that is not infra-red-active, and all symmetry species are Raman-active. Of these 237 modes, we shall concentrate

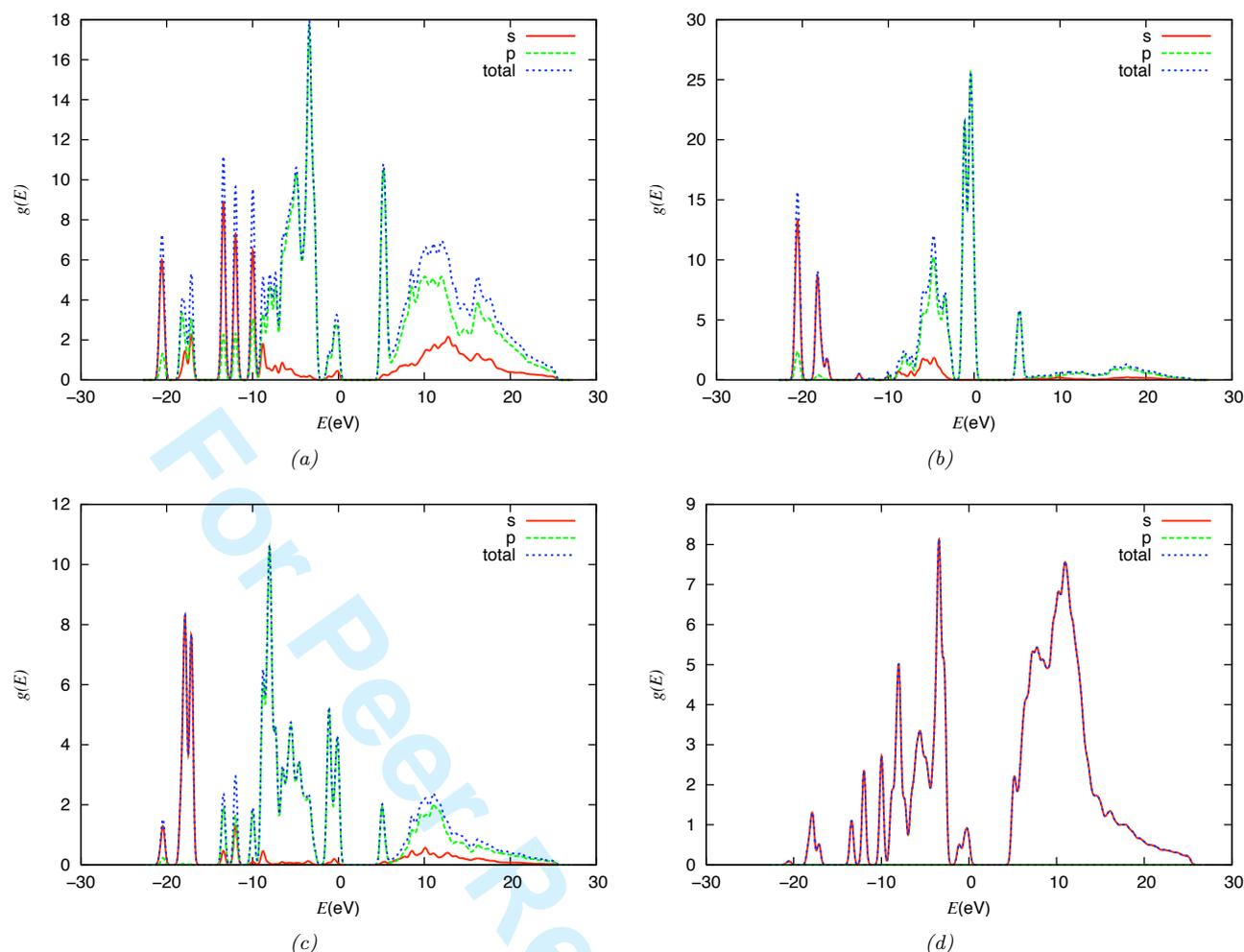


Figure 5. Local densities of states: in (a) for carbon atoms; in (b) for oxygen atoms; in (c) for nitrogen atoms; (d) hydrogen atoms.

our attention upon the vibronic modes. To motivate the discussion, in Figure 6, we present the calculated IR spectrum for the crystalline system. This is determined as described in [72]. Visualisation of the phonon eigenvectors corresponding to given mode frequencies allows the oscillations responsible for the IR features to be understood.

If we perform such an analysis and examine the IR spectrum, we initially address the major peaks. The large peak at around  $3160\text{ cm}^{-1}$  corresponds to stretches of the amide NH bond. It is not entirely decoupled from other atomic motions, however; on the contrary, oscillations are also observed in the NH<sub>3</sub> group. Furthermore, the NH<sub>3</sub> oscillations are different in the different “layers” of the herring-bone structure; layers of alternating oscillators are formed, in which, in one layer, the NH<sub>3</sub> group undergoes symmetric stretches, then in the next, undergoes asymmetric stretches, which appear almost like a “rotation” of the functional group. The presence of a peak at this frequency is found in the solid-state IR data presented by Kolev *et al.* [67]. It is interesting to note, however, that their calculations assign this oscillation a frequency of  $3216\text{ cm}^{-1}$ ; it is not clear, though, whether their data refer to a crystal, or an isolated molecule. If it is the latter, then this would seem to be consistent with a red-shifting of the vibrational frequency upon crystallisation. Given that the amide group participates in a hydrogen bond with the carboxy group in the solid state, such a red-shift can be anticipated; however, compared

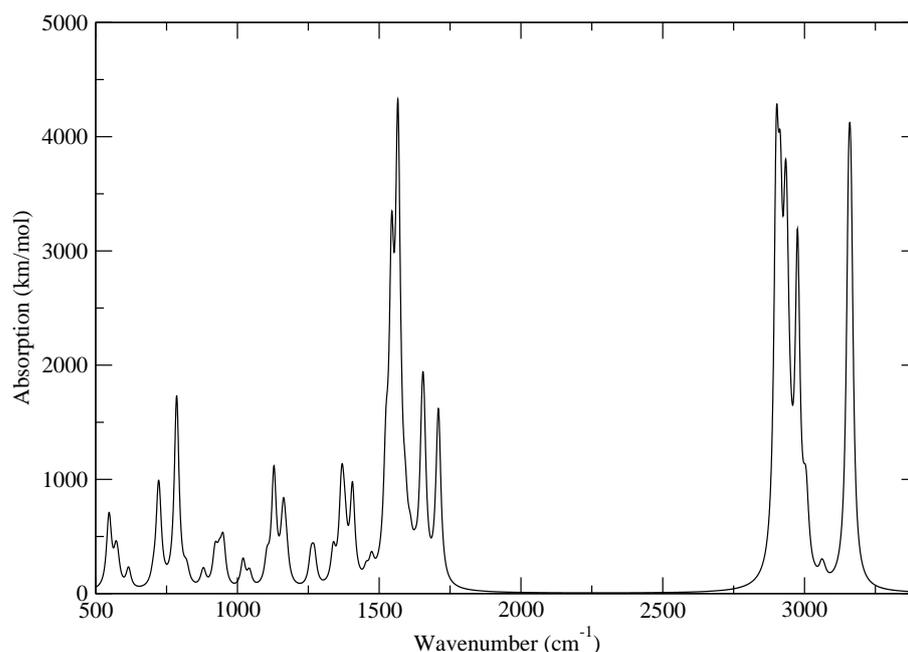


Figure 6. IR spectrum of the glycyl-L-alanine crystal. A Gaussian smearing with of  $20 \text{ cm}^{-1}$  has been applied.

to the magnitude of red-shifts inferred from calculations on amino acids [56], the red-shift appears to be smaller than expected.

The second set of peaks of interest form a triplet at  $\sim 2900 \text{ cm}^{-1}$ ,  $2936 \text{ cm}^{-1}$  and  $2977 \text{ cm}^{-1}$ . The lowest frequency peak absorbs more strongly than the other two; this mode consists of asymmetric stretches in the terminal amine group. This peak forms a doublet with that at  $2936 \text{ cm}^{-1}$ . This latter feature is revealed to arise from symmetric stretching of the bonds of the  $\text{NH}_3$  group in one layer, alternating with layers in which the bonds undergo a symmetric bending motion. It can perhaps be anticipated that such a motion will not result in as large a transition moment, explaining why this mode absorbs less strongly. The mode at  $2997 \text{ cm}^{-1}$  is somewhat similar, except that the symmetric stretches have instead been replaced by asymmetric stretches. We note that such modes do not excite any other molecular motions.

The other major absorption features are a doublet at  $\sim 1547 \text{ cm}^{-1}$  and  $1566 \text{ cm}^{-1}$ . The lower frequency peak corresponds to a more complex collective motion, with the main motions arising in the carboxyl and amine groups, and lesser motions observed in the amide  $\text{NH}$  group. The carboxyl group motions are as follows: again, a “layer” structure is found, in which the carboxy groups in one layer undergo asymmetric stretches, with those in the next layer undergoing a rocking motion that is aligned along the molecular axis. The amine groups also undergo motion which is different in each layer: in one layer the  $\text{NH}$  bonds undergo a bending motion, in which all three bend in phase; in the next layer, the bonds undergo what appears to be a combination of bending and bonding, again, in phase. In the former case, the amide  $\text{NH}$  bond rocks along the molecular axis; in the latter, in contradistinction, the amide bond rocks perpendicular to the molecular axis. These differences in

Table 5. Optical ( $\epsilon^\infty$ ) and static ( $\epsilon_0$ ) contributions to the dielectric permittivity.

$\epsilon^\infty$			$\epsilon_0$		
2.37	0.00	0.00	3.59	0.00	0.00
0.00	2.31	0.00	0.00	4.64	0.00
0.00	0.00	2.32	0.00	0.00	3.71

the motion between layers may be understood in terms of the hydrogen bonding network present in the crystal; for example, in cases where the amide NH oscillates parallel to the molecular axis, then this may correlated with the carboxy group rocking; conversely, NH rocking perpendicular to the molecular axis correlates with asymmetric stretches of the carboxy group. That is, the carboxy and amide group oscillations are coupled through the intermolecular hydrogen bonds present in the crystal. This demonstrates that, even for high frequency intramolecular modes, intermolecular interactions play an important part in determining the nature of the oscillations. Thus, the effect of intermolecular interactions is *not* simply to cause a shift in the frequency of intramolecular vibrations; rather, they may also couple intramolecular vibrations on neighbouring molecules. We note briefly that the oscillations leading to the higher frequency peak are the same; however, the amplitude is increased, which explains the increased size of the absorption peak at this frequency.

The smaller peak at  $\sim 1660 \text{ cm}^{-1}$  arises from pronounced bending of the NH bonds within the amide group that are almost entirely decoupled from any other atomic motions. At  $\sim 1714 \text{ cm}^{-1}$ , we find the amide I bands, arising from stretching of the amide CO bonds. Given that this involves backbone atoms, in contrast to the above excitations, we find a collective motion in which other atoms are also induced to move; for example, some rocking of the amide NH bonds is found, in which alternating layers execute this motion either parallel to, or perpendicular to, the molecular axis. As above, these differences are correlated with differences in the carboxy group oscillations. Bending of the NH bonds in the amine group are also found.

Finally, in Table 5, we present both the optical permittivity,  $\epsilon^\infty$ , and the static dielectric permittivity,  $\epsilon_0$ . It is noticeable that the static (i.e. ionic) contribution to the permittivity is greater than the optical (i.e. electronic) contribution. This result can be anticipated for the following reason: both the IR absorption and the static permittivity are proportional to the oscillator strength; therefore, significant IR absorption, as observed above, implies a large non-zero oscillator strength. This will then manifest itself in the static contribution to the permittivity.

## 6. Conclusions

In this work we have examined, in detail, the structural, electronic and lattice dynamical properties of the glycyl-L-alanine molecular crystal using density functional theory. Our calculations reveal that the crystal is bound together by a hydrogen bond network, with molecules being arranged within the unit cell to maximise the number of hydrogen bonds possible between the oppositely-charged amine and carboxy functional groups. Indeed, the calculations reveal that a hierarchy of hydrogen bond strengths exists within the crystal, with those between amide and carboxy oxygens being weaker than those between oppositely-charge functional groups. Recent studies of organic molecular crystals [55] have suggested the importance of adequately treating van der Waals interactions, with a failure to do so manifesting itself in overestimates of lattice constants. In contrast, in this work

1 the theoretically-determined lattice constants are found to be within 2% of the  
2 experimental values, suggesting that hydrogen bonds are dominant in stabilising  
3 the crystal, and that the neglect of van der Waals interactions implicit in the GGA  
4 XC functionals employed in this study is not significant. The slight overestimation  
5 of the  $b$  and  $c$  axes relative to experiment is consistent with the known behaviour  
6 of GGA functionals (i.e. they tend to underbind); in contrast, the underestimate  
7 of the  $a$  axis suggests that, in this case, the GGA may over-bind.

8  
9 Bandstructure calculations indicate that the glycyl-l-alanine crystal is a wide-  
10 indirect-gap insulator, with indirect bandgaps of 4.95 eV, as compared to a di-  
11 rect bandgap of 5.00 eV. The details of the bandstructure, namely, the directional  
12 dependencies of the bandwidths are found to be directly correlated with the hydro-  
13 gen bond orientations in the unit cell, suggesting that this, and by extension, the  
14 molecular arrangement, plays an important role in determining the bandstructure,  
15 and that it is coupling of electronic states via these hydrogen bonds that results in  
16 band dispersion. Dispersion is found to be particularly prevalent in the conduction  
17 band states; the valence band states show little dispersion. This is consistent with  
18 localised molecular orbitals and primary intermolecular interactions via hydrogen  
19 bonds; it is further supported by the density of states plots, which indicate that  
20 these states primarily arise from oxygen  $2p$ -states, and indicate a lack of hybridi-  
21 sation. In contrast, the dispersive conduction band states are found to possess a  
22 more pronounced hybridisation between hydrogen  $s$ - and carbon  $p$ -states.

23  
24 The lattice dynamical and IR absorption calculations reveal that major absorp-  
25 tion features arise owing to excitation of specific functional groups and/or bonds  
26 whereas, in contrast, the weak absorption features rather arise from motions of  
27 the molecular backbone, and hence couple several functional groups resulting in  
28 motions that are more difficult to characterise. However, even in the intramolec-  
29 ular modes, the presence of intermolecular hydrogen bonds is found to couple  
30 intramolecular motions in neighbouring molecules, demonstrating that an accurate  
31 treatment of intermolecular forces is necessary to obtain accurate intramolecular  
32 modes.

33  
34 It is hoped that these calculations will permit greater insight into the electronic  
35 and lattice dynamical properties of peptidic molecular crystals, and that these may  
36 yield insights into the structural motifs observed in liquid-state peptides. More par-  
37 ticularly, molecular crystal calculations allow, through crystal polymorphism, the  
38 effects of varying inter- and intramolecular interactions and molecular conforma-  
39 tion upon electronic properties to be understood. Understanding this relationship  
40 potentially opens up avenues by which electronic and optical properties may be  
41 modulated by mechanical (conformational) changes in the molecular structure,  
42 which is of interest to technological exploitation of such molecules. We aim, there-  
43 fore, to use this work as a foundation to investigate and elucidate this relationship.

#### 44 45 46 47 48 **Acknowledgements**

49  
50 The authors would like to acknowledge the UK HECToR supercomputing facility,  
51 which was used to perform these calculations.

#### 52 53 54 **References**

- 55  
56 [1]H. Imamura and M. Kato, *Proteins - Structure Function and Bioinformatics* **75** (4), 914 (2009).  
57 [2]L.L. Lee, H. Ha, Y.T Chang *et al.*, *Protein Science* **18** (2), 277 (2009).  
58 [3]B. Fierz and T. Kiefhaber, *J. Am. Chem. Soc.* **12** (3), 672 (2007).  
59 [4]N.S. Hinrichs and V.S. Pande, *J. Chem. Phys.* **126** (24), 244101 (2007).  
60

- 1 [5] F. Noe and S. Fischer, *Curr. Op. Struct. Biol.* **18** (2), 154 (2008).
- 2 [6] Y. Suzuki, J.K. Noel and J.N. Onuchic, *J. Chem. Phys.* **128** (2), 025101 (2008).
- 3 [7] P.R. Tulip, R.Z. Troitzsch, E. Cerasoli *et al.*, in preparation (2009).
- 4 [8] S. Kannan and M. Zacharias, *J. Struct. Biol.* **166** (3), 288 (2009).
- 5 [9] M. Khalili and D.J. Wales, *J. Chem. Theor. Comput.* **5** (5), 1380 (2009).
- 6 [10] S. Hofinger, B. Almeida, and UHE Hansmann, *Proteins - Structure Function and Bioinformatics* **68**
- 7 (3), 662 (2007).
- 8 [11] P. Campiglia, C. Esposito, M. Scrima, *et al.*, *Chem. Biol. Drug Des.* **69** (2), 111 (2007).
- 9 [12] C. Soto, R.J. Kascsak, G.P. Saborio, *et al.*, *The Lancet* **355** (9199), 192 (2000).
- 10 [13] D.J. Gordon, R. Tappe and S.C. Meredith, *J. Peptide Res.* **60** (1), 37 (2002).
- 11 [14] L. Bennett, R. Williams, H. Ecroyd *et al.*, *Aust. J. Dairy Tech.* **60** (1), 117 (2009).
- 12 [15] C. Nerelius, J. Johansson and A. Sandregen, *Front. Bioscience* **14**, 1716 (2009).
- 13 [16] S.Y. Chung, S. Lee, C. Liu *et al.*, *J. Phys. Chem. B* **113** (1), 292 (2009).
- 14 [17] T. Tabarin, A. Kulesza, R. Antoine *et al.*, *Phys. Rev. Lett.* **101** (21), 213001 (2008).
- 15 [18] R. Mitric, J. Petersen, A. Kulesza *et al.*, *J. Chem. Phys.* **127** (13), 134301 (2007).
- 16 [19] R. Mitric, J. Petersen, A. Kulesza *et al.*, *Chem. Phys.* **343** (2-3), 372 (2008).
- 17 [20] A. Kulesza, R. Mitric and V. Bonacic-Koutecky, *J. Phys. Chem. A* **113** (16), 3783 (2009).
- 18 [21] E.M. Krauland, B.R. Peelle, K.D. Wittrup *et al.*, *Biotech. Bioeng.* **97** (5), 1009 (2007).
- 19 [22] J.E. Zull, J. Reedmundell, Y.W. Lee *et al.*, *J. Indust. Microbiol.* **13** (3), 137 (1994).
- 20 [23] M. Mijajlovic and M.J. Biggs, *J. Phys. Chem. C* **111** (43), 15839 (2007).
- 21 [24] K.E. Sapsford, Y.S. Shubin, J.B. Delehanty *et al.*, *J. Appl. Microbiol.* **96**, 47 (2004).
- 22 [25] M. Sarikaya, C. Tamerler, A.K.Y. Jen *et al.*, *Nat. Mater.* **2**, 577 (2003).
- 23 [26] E. Katz and I. Willner, *Angew. Chem. Int. Ed.* **43**, 6042 (2004).
- 24 [27] C.M. Niemeyer, *Angew. Chem. Int. Ed.* **40**, 4128 (2001).
- 25 [28] S.V. Patwardhan, G. Patwardhan and C.C. Perry, *J. Mat. Chem* **17**, 2875 (2007).
- 26 [29] E.E. Oren, C. Tamerler and M. Sarikaya, *Nano Lett.* **5** (3), 415 (2005).
- 27 [30] L.M. Ghiringelli, B. Hess, N.F.A. van der Vegt *et al.*, *J. Am. Chem. Soc.* **130**, 13460 (2008).
- 28 [31] S. Monti, V. Carravetta, C. Battocchio *et al.*, *Langmuir* **24** (7), 3205 (2008).
- 29 [32] R.Z. Troitzsch, H. Vass, W.J. Hossack *et al.*, *J. Phys. Chem. B* **112** (14), 4290 (2007).
- 30 [33] S. Köppen, O. Bronkalla and W. Langel, *J. Phys. Chem. C* **112**, 13600 (2008).
- 31 [34] A.A. Mungikar and D. Forciniti, *Biomacromolecules* **5** (6), 2147 (2004).
- 32 [35] P. Hohenberg and W. Kohn, *Phys. Rev.* **136** (3B), B864 (1964).
- 33 [36] W. Kohn and L.J. Sham, *Phys. Rev.* **140** (4A), 1133 (1965).
- 34 [37] M.C. Payne, M.P. Teter, D.C. Allan *et al.*, *Rev. Mod. Phys.* **64** (4), 1045 (1992).
- 35 [38] R.Z. Troitzsch, P.R. Tulip, J. Crain *et al.*, *Biophys. J.* **95** (11), 5014 (2008).
- 36 [39] C.J. Pickard and R.J. Needs, *Phys. Rev. Lett.* **102** (12), 125702 (2009).
- 37 [40] V. Thakor, J.B. Staunton, J. Poulter *et al.*, *Phys. Rev. B* **68** (13), 134412 (2003).
- 38 [41] R.M. Martin, *Electronic Structure: Basic Theory and Practical Methods* (Cambridge University Press,
- 39 2004).
- 40 [42] G. Nandini and D.N. Sathyanarayana, *J. Mol. Struct. Theochem.* **638**, 79 (2003).
- 41 [43] R. Jacob and G. Fischer, *J. Mol. Struct.* **613** (1-3), 175 (2002).
- 42 [44] T. Liang and T.R. Walsh, *Mol. Sim.* **33** (4-5), 337 (2007).
- 43 [45] H.W. Hugosson, A. Laio, P. Maurer *et al.*, *J. Comput. Chem.* **27** (5), 672 (2006).
- 44 [46] S.E. McLain, A.K. Soper, I. Diadone *et al.*, *Ang. Chem. Int. Ed.* **47**, 9059 (2008).
- 45 [47] P.R. Tulip and S.P. Bates, *J. Chem. Phys.* **131**, 015103 (2009).
- 46 [48] Y. Wei, D-K. Lee and A. Ramamoorthy, *J. Am. Chem. Soc.* **123**, 6118 (2001).
- 47 [49] A. Zheng, S-B. Liu and F. Deng, *J. Comput. Chem.* **30**, 222 (2009).
- 48 [50] R. O. Jones and O. Gunnarsson, *Rev. Mod. Phys.* **61** (3), 689 (1989).
- 49 [51] S.J. Clark, M.D. Segall, C.J. Pickard *et al.*, *Zeit. Krist.* **220** (5-6), 567 (2005).
- 50 [52] L. Kleinman, D. M. Bylander, *Phys. Rev. Letters* **48**, 1425 (1982).
- 51 [53] H.J. Monkhorst and J.D. Pack, *Phys. Rev. B* **13** (2), 5188 (1976).
- 52 [54] J.P. Perdew and Y. Wang, *Phys. Rev. B* **46**, 12947 (1992).
- 53 [55] P.R. Tulip and S.J. Clark, *Phys. Rev. B* **71** (19), 195117 (2005).
- 54 [56] P.R. Tulip and S.J. Clark, *Phys. Rev. B* **74** (6), 064301 (2006).
- 55 [57] A.H.J. Wang and I.C. Paul, *Cryst. Struct. Comm.* **8** (2), 269 (1979).
- 56 [58] S.L. Price, *Cryst. Eng. Comm.* **6**, 344 (2004).
- 57 [59] C. Ouvrard and S.L. Price, *Cryst. Growth Des.* **4**, 1119 (2004).
- 58 [60] M.A. Neumann, F.J.J. Lensen and J. Kendrick, *Angew. Chem. Int. Ed.* **47**, 2427 (2008).
- 59 [61] X. Gonze and C. Lee, *Phys. Rev. B* **55** (16), 10355 (1997).
- 60 [62] K. Refson, P.R. Tulip and S.J. Clark, *Phys. Rev. B* **73** (15), 155114 (2006).
- [63] P.R. Tulip and S.J. Clark, *J. Chem. Phys.* **121** (11), 5201 (2004).
- [64] B. Winkler, J.D. Gale, A.K. Refson *et al.*, *Phys. Chem. Minerals* **35** (1), 25 (2008).
- [65] G.K. Horton and A.A. Maradudin, *Dynamical Properties of Solids Vol 1* (North-Holland Publishing
- Company, Amsterdam 1974).
- [66] P.U. Jepsen and S.J. Clark, *Chem. Phys. Lett.* **442** (4-6), 145 (2007).
- [67] Ts. Kolev, B.B. Koleva, S.Y. Zareva *et al.*, *Inorg. Chim. Acta* **359**, 4367 (2006).
- [68] J.P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [69] R.S. Mulliken, *J. Chem. Phys.*, **23** (10), 1833 (1955).
- [70] A. Seidl, A. Görling, P. Vogl *et al.*, *Phys. Rev. B* **53** (7), 3764 (1996).
- [71] M.Z.S. Flores, V.N. Freire, R.P. dos Santos *et al.*, *Phys. Rev. B* **77** (11), 115104 (2008).
- [72] P.R. Tulip, Ph.D. thesis, University of Durham, 2004. <<http://portellen.phycmt.dur.ac.uk/sjc/>>