Structural and electronic properties of L-amino acids

P. R. Tulip and S. J. Clark
Department of Physics, University of Durham, Science Laboratories, South Road, Durham, DH1 3LE, United Kingdom
(Received 20 August 2004; revised manuscript received 22 November 2004; published 26 May 2005)

The structural and electronic properties of four L-amino acids alanine, leucine, isoleucine, and valine have been investigated using density functional theory (DFT) and the generalized gradient approximation. Within the crystals, it is found that the constituent molecules adopt *zwitterionic* configurations, in agreement with experimental work. Lattice constants are found to be in good agreement with experimentally determined values, although certain discrepancies do exist due to the description of van der Waals interactions. We find that these materials possess wide DFT band gaps in the region of 5 eV, with electrons highly localized to the constituent molecules. It is found that the main mechanisms behind crystal formation are dipolar interactions and hydrogen bonding of a primarily electrostatic character, in agreement with current biochemical understanding of these systems. The electronic structure suggests that the amine and carboxy functional groups are dominant in determining band structure.

DOI: 10.1103/PhysRevB.71.195117 PACS number(s): 61.66.Hq, 71.20.Rv, 71.15.Mb

I. INTRODUCTION

Determination of the band structure of a crystalline system is fundamental to understanding the electronic structure of a crystal, and hence its physical properties. While the electronic structures of conventional semiconducting crystals have been exhaustively studied using *ab initio* methods, only a small number of papers covering this aspect of molecular crystals exists,[1–5] which may be explained with reference to the computational complexity required to study such systems. However, advances in computer hardware are meaning that increasingly complex systems are now falling within the remit of *ab initio* methods.

Recent years have seen increasing interest in the properties of molecules of biological and biochemical interest, such as the amino acids,[6–10] although the majority of this work has concerned the application of *ab initio* methods to problems of conformational analysis for isolated molecules in the gaseous phase. In the solid state they often exist as molecular crystals, stabilized largely by electrostatic interactions. These have attracted surprisingly little interest from a theoretical point of view, save for some work concerning the shielding tensors of carbon-13.[11,12] It is therefore of interest to investigate the geometric and electronic structure of these systems for a number of reasons. Such calculations allow an understanding of experimentally established phenomena in terms of the underlying quantum mechanics. This provides an insight into the nature of the interactions between amino acids, principally the hydrogen bond, allowing a further understanding of the biological functionality of proteins, composed as they are, of amino acids. A firm understanding of the structural and electronic properties of these systems, and the relationship between the two, is an essential starting point for further investigation into the physics of these systems. Lastly, molecular crystals, forming a state intermediate between molecules and the conventional solid state, offer a rich vein of physical phenomena to be investigated, and so provide a rigorous testing ground of the applicability of *ab initio* methodologies.

In this paper we investigate the structural and electronic properties of the solid state phases of four amino acids alanine, valine, leucine, and isoleucine, and attempt to elucidate the connections between the structural and electronic properties. The paper is structured as follows. The next section provides an overview of the theoretical approach used in this work, while Sec. III describes the structural properties. Section IV covers electronic properties, while in Sec. V we draw conclusions on our results.

II. COMPUTATIONAL APPROACH

We carry out the calculations within the plane-wave pseudopotential formalism of density functional theory (DFT).[13,14] This is a powerful technique for the accurate computation of ground-state properties, here performed using the CASTEP code.[15] The electron-ion interactions are described using nonlocal norm-conserving pseudopotentials of the Kleinman-Bylander form.[16] The valence electron wave functions are expanded in a plane-wave basis set to a kinetic energy cutoff of 1000 eV which converges total-energy properties to better than 1 meV/atom. Reciprocal space integrations have been carried out over varying numbers of symmetry-reduced k points, where the energy convergence is the same tolerance as the plane-wave cutoff. Exchange and correlation are treated using the well-known generalized gradient approximation (GGA) due to Perdew and Wang.[17] This is known to perform significantly better for hydrogen-bonded molecular systems than the local density approximation (LDA).[18,19]

Our calculations fall into two distinct stages. First we determine the geometrical structures adopted at equilibrium by minimizing the forces acting on nuclei using the well-known Hellmann-Feynman theorem. This involves optimization of all internal degrees of freedom using a Broyden-Fletcher-Goldfarb-Shanno algorithm. We take experimentally determined structures[20–23] to determine our initial atomic positions. These initial structures are zwitterionic, that is, the amino and carboxy functional groups ionize via the donation of a proton from the carboxy group to the amino group. Once we have optimized these structures, we then solve the single-

particle Kohn-Sham equations at a series of reciprocal space points of high symmetry to determine the eigenvalue spectrum.

III. STRUCTURAL PROPERTIES

The determination of the structure of molecular crystal systems is a formidable problem, as has been discussed by Gdanitz:[24] although one may optimize the intramolecular structure relatively easily, the intermolecular structure represents a more challenging task due to the nature of the intermolecular forces and the large number of degrees of freedom typically present. Although, in principle, *ab initio* calculations are able to obtain the crystal structure from the constituent molecular structure the task is theoretically demanding as molecular crystals are typically bound by van der Waals forces; current exchange-correlation functionals such as the LDA or the GGAs do not describe such nonlocal interactions accurately. In this work, the molecules that we consider are polar in character, and thus, as discussed later, we can expect that dipole-dipole interactions and hydrogen bonds will dominate. We use the experimentally determined structures[20–23] to provide initial atomic positions and crystal symmetries, after which we may then minimize the atomic forces to determine an equilibrium geometry consistent with the system symmetry. One concern associated with the use of experimentally determined structures to provide initial geometries is that this could bias the result of the calculation. In this situation, one would expect the lattice parameters and atomic positions to vary only a small amount over the course of the calculation from the initial positions, as the calculation attempts to identify the minimum of the potential surface involved. However, this is not the case in these calculations; rather, examination of the lattice parameters and atomic coordinates during the course of the calculations reveals that these quantities may, at intermediate stages of the calculation, vary between 4% and 10% from the final equilibrium values. This suggests that the effects of biasing are not pronounced. Of course, one could investigate this more fully by taking an initial geometry far from that obtained from experiment; in this case, however, one would be hindered by the intense computational requirements of such an optimization, and moreover, by the possibility of local minima caused by the conformational freedom of the constituent molecules. We thus feel justified in the procedure used in this work.

Experimentally, L-alanine crystallizes in the orthorhombic structure with space group $P2_12_12_1$, while the remainder of the systems under consideration crystallize in the monoclinic structure with space group $P2_1$. In all cases there are four molecules per unit cell.

In Fig. 1 we show the structure of L-valine. One can observe that the molecules have formed *zwitterions*; that is, the amino and carboxy functional groups ionize, forming oppositely charged functional groups, via the donation of a proton from the carboxy group to the amino group; the resulting amino group is now NH_3^+, while the carboxy group is COO^-. This is expected, as it is well known that these zwitterions are stabilized in the solid state by electrostatic, polarization, and hydrogen-bonding interactions with the crystalline

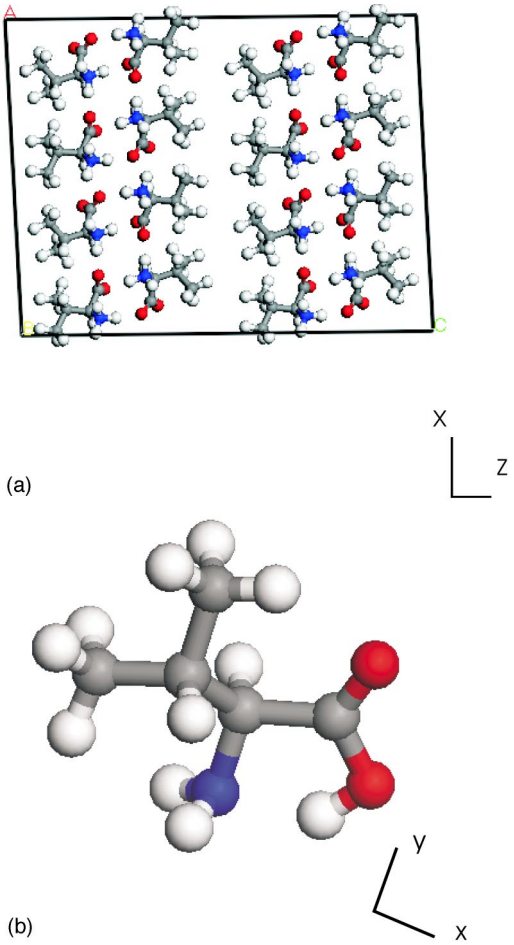

FIG. 1. The structure of L-valine: supercell used in this study in (a); free molecule in (b). Note the "layer" formation in (a). Carbon atoms are represented by gray atoms, hydrogen by white, and oxygen by dark gray (red). The nitrogen atoms are depicted as black (blue). This color convention is used throughout this work.

environment.[6] This is in contrast to the gaseous phase, as can be seen by examining Fig. 1(b) illustrating the structure of an isolated valine molecule. In the crystal one can see that the molecules have arranged themselves such that the positively ionized amine functional groups are in close proximity to the negatively ionized carboxy groups of neighboring molecules. This is readily understood via classical electrostatics: a bond between molecules may be formed via these charged functional groups. The molecules may also be bound by hydrogen bonding between the carboxy groups and the amine groups. Such bonding is well known to take place in these molecular systems.[25,26] It can be seen in Fig. 1 that the crystal structure allows oxygen atoms in the carboxy groups to participate in multiple hydrogen bonds; this is common in such structures.[25] However, such bonds are not generally of the same strength, and this is reflected in different bond lengths. For example, the O-H bonds between neighboring molecules in the *a* axis, where the H atom forms part of a donor group common to both bonds, and both oxygen atoms are in the same carboxy group, range from 1.71 to 1.82 Å.

This though does not address why it is energetically favorable for the molecules to form zwitterions: valine has a

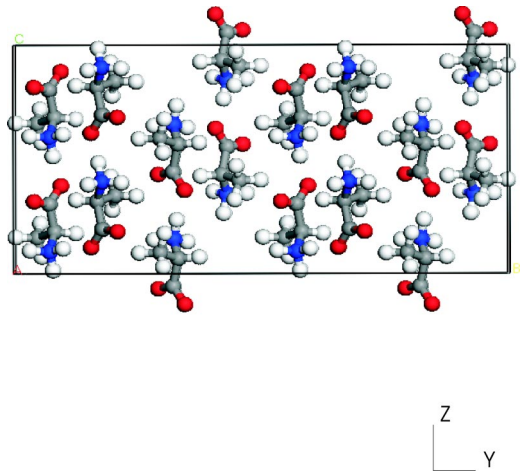

FIG. 2. Structure of alanine: alanine supercell. A percolating network of hydrogen bonds is formed in the structure.

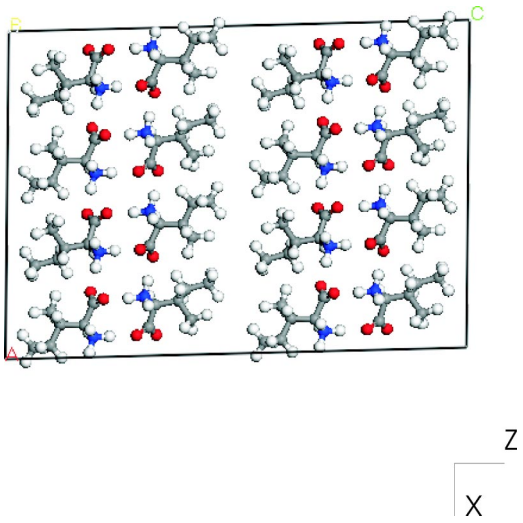

FIG. 3. Structure of isoleucine.

nonzero dipole moment; therefore the possibility of forming bonds with neighboring molecules must exist via dipole-dipole interactions. However, to form a zwitterion must involve some energetic penalty: calculations on the isolated molecules suggest that the energy cost of forming the zwitterionic species is of the order of 0.68 eV. Therefore, why is it energetically favorable to zwitterionize? This is a point that we return to later on, when we discuss the electronic structure of L-valine.

One consequence of the molecular arrangement, as shown in Fig. 1, is the formation of "layers" in the *c* axis, which one would expect to be weakly bonded to each other, in contrast to the dipolar interactions and hydrogen bonding of the intralayer arrangement.

We may contrast this structure with that of L-alanine, shown in Fig. 2. Again, the molecules form zwitterions but the molecular arrangement differs from that of L-valine. Examining this $2\times2\times2$ supercell, looking along the *a* axis reveals a pattern of interlocking "layers," each two molecules "thick" of oppositely aligned molecules. This arrangement allows the interaction energy between the oppositely ionized functional groups of neighboring molecules to be minimized, hence making this arrangement energetically favorable. It is worth noting that in contrast to L-valine, no weakly bonded layers appear in the structure. This is significant, as it may explain why, for example, the agreement between experiment and theory (shown in Table I) is so much better for L-alanine as opposed to that exhibited for L-valine. This can be understood as follows: charge localization means that it is unlikely that the interlayer bonding in L-valine is due to any form of charge overlap. Further, the "interfaces" of each layer are composed of CH_3 groups, so dipolar interactions are unlikely to be significant. Thus one could reasonably assume that nonlocal van der Waals–type interactions are responsible for the interlayer bonding. Semilocal exchange-correlation functionals such as the GGA are poor at describing such nonlocal interactions, and thus one could expect some inaccuracy to be present. Contrariwise, in L-alanine, no such layers exist; the intermolecular and interlayer interactions appear to be primarily dipolar and hydrogen-bond-like, which are accurately described by GGAs. One would therefore expect to find the agreement that is indeed found. This is also the most likely explanation for the fact that in some circumstances, the theoretical lattice constants have been found to be greater than the experimental lattice constants. In connection with this, though, it should be noted that GGAs are occasionally responsible for underbinding in solids, resulting in lattice constants that are slightly too long.[27]

Interestingly, for L-isoleucine, poor agreement between the experimentally and theoretically derived lattice parameters occurs in the *a* axis. Examining Fig. 3 it is evident that the bonding in this direction involves hydrogen bonds

TABLE I. Structural parameters of the molecular crystals L-alanine, L-isoleucine, L-leucine, and L-valine; experimental (cited above) versus theoretical (this study) results. All lattice parameters are quoted in angstroms.

	a		*b*		*c*	
	Theory	Experiment	Theory	Experiment	Theory	Experiment
L-alanine	6.18	6.03	12.29	12.34	5.83	5.78
L-leucine	14.72	14.67	5.29	5.32	9.60	9.61
L-isoleucine	10.20	9.68	5.31	5.30	14.48	13.96
L-valine	9.71	9.68	5.23	5.25	12.51	11.93

TABLE II. Theoretically determined atomic positions for L-alanine compared to the experimentally determined positions (Ref. 20).

	x		*y*		*z*	
	Experiment	Theory	Experiment	Theory	Experiment	Theory
H(1)	−0.301	−0.285	0.064	0.064	−0.806	−0.797
H(2)	−0.231	−0.214	0.179	0.198	−0.777	−0.780
H(3)	−0.401	−0.404	0.150	0.156	−0.957	−0.971
H(4)	−0.558	−0.559	0.239	0.248	−0.654	−0.646
H(5)	−0.778	−0.772	0.108	0.109	−0.856	−0.852
H(6)	−0.847	−0.833	0.103	0.103	−0.590	−0.558
H(7)	−0.685	−0.660	0.014	0.005	−0.694	−0.681
C(1)	−0.439	−0.428	0.142	0.141	−0.398	−0.390
C(2)	−0.523	−0.513	0.161	0.162	−0.644	−0.632
C(3)	−0.725	−0.706	0.091	0.091	−0.698	−0.684
N	−0.344	−0.341	0.138	0.143	−0.814	−0.805
O(1)	−0.271	−0.270	0.084	0.080	−0.372	−0.368
O(2)	−0.550	−0.528	0.185	0.188	−0.239	−0.230

between carboxy and amino groups. Thus the above argument is unsatisfactory in this particular case. Given the convergence tolerances used in this work, it is unlikely that this is due to numerical reasons. Examination of the structure reveals that the hydrogen bonds in the *a* axis consist entirely of bonds involving the single-bonded oxygen atom (i.e., the O^- atom in COO^-), in contrast to the situation in L-leucine, for example, where bonds of both types occur in this direction. This suggests strongly that this is the underlying reason for this discrepancy.

TABLE III. Theoretically determined atomic positions for L-leucine compared to experimentally determined positions (Ref. 21).

	x		*y*		*z*	
	Experiment	Theory	Experiment	Theory	Experiment	Theory
H(1)	0.625	0.623	0.528	0.503	0.959	0.956
H(2)	0.507	0.507	0.246	0.213	0.870	0.859
H(3)	0.579	0.578	0.061	0.085	0.984	0.990
H(4)	0.587	0.589	0.039	0.008	0.802	0.820
H(5)	0.747	0.744	0.162	0.136	0.810	0.813
H(6)	0.743	0.740	0.175	0.168	0.993	0.994
H(7)	0.800	0.800	0.609	0.584	0.816	0.806
H(8)	0.744	0.741	0.750	0.739	1.033	1.029
H(9)	0.863	0.859	0.780	0.788	1.026	1.025
H(10)	0.817	0.819	0.534	0.537	1.122	1.123
H(11)	0.941	0.956	0.517	0.480	0.898	0.904
H(12)	0.914	0.907	0.230	0.213	0.817	0.818
H(13)	0.930	0.914	0.244	0.232	1.000	1.002
C(1)	0.618	0.615	0.512	0.495	0.738	0.731
C(2)	0.636	0.636	0.389	0.370	0.880	0.874
C(3)	0.736	0.734	0.283	0.273	0.898	0.895
C(4)	0.810	0.809	0.485	0.470	0.905	0.901
C(5)	0.807	0.807	0.649	0.644	1.030	1.026
C(6)	0.903	0.902	0.345	0.341	0.908	0.906
N	0.575	0.574	0.164	0.153	0.887	0.886
O(1)	0.587	0.582	0.382	0.358	0.639	0.634
O(2)	0.636	0.635	0.743	0.723	0.731	0.719

TABLE IV. Theoretically determined atomic positions for L-isoleucine compared to experimentally determined positions (Ref. 22).

	x		*y*		*z*	
	Experiment	Theory	Experiment	Theory	Experiment	Theory
H(1)	0.303	0.288	−0.348	−0.336	0.396	0.383
H(2)	0.452	0.451	−0.286	−0.255	0.415	0.391
H(3)	0.359	0.349	−0.197	−0.182	0.480	0.477
H(4)	0.419	0.408	0.133	0.173	0.394	0.385
H(5)	0.438	0.425	−0.142	−0.113	0.254	0.247
H(6)	0.337	0.304	0.357	0.401	0.223	0.214
H(7)	0.493	0.468	0.287	0.339	0.252	0.248
H(8)	0.466	0.449	0.380	0.413	0.088	0.080
H(9)	0.343	0.331	0.177	0.175	0.070	0.065
H(10)	0.499	0.494	0.085	0.098	0.100	0.100
H(11)	0.198	0.189	−0.261	−0.263	0.239	0.243
H(12)	0.250	0.238	−0.186	−0.152	0.144	0.136
H(13)	0.163	0.140	−0.003	0.038	0.196	0.204
C(1)	0.206	0.198	0.131	0.157	0.384	0.385
C(2)	0.347	0.330	0.018	0.047	0.367	0.359
C(3)	0.363	0.343	−0.019	0.018	0.259	0.254
C(4)	0.408	0.383	0.231	0.264	0.216	0.209
C(5)	0.431	0.416	0.217	0.236	0.109	0.108
C(6)	0.230	0.221	−0.128	−0.096	0.204	0.207
N	0.366	0.354	−0.228	−0.199	0.420	0.405
O(1)	0.116	0.119	−0.017	0.019	0.413	0.424
O(2)	0.185	0.176	0.360	0.382	0.364	0.362

Examination of the atomic positions as determined experimentally and theoretically in Tables II–V lends credence to this view, as in all cases individual intramolecular degrees of freedom can be dealt with satisfactorily; evidence for this is the good agreement found for the atomic positions of an individual molecule of the crystal. Rather, this suggests that, in accordance with the above discussion, it is the intermolecular interactions that can be problematic.

L-isoleucine appears structurally to be similar to L-valine, as shown in Fig. 3. One may immediately note the interesting manner in which the functional groups on one molecule form hydrogen bonds with their oppositely charged counterparts on the neighboring molecule. Thus individual molecules within the unit cell are bound by two hydrogen bonds. Again, like L-valine, examining a supercell shows the same layer formation. One may postulate the following explanation for such formations occurring. L-alanine possesses a short side chain; therefore it is possible for the molecules to "interlock" to a greater extent than is possible if the functional groups are attached to a long side chain. It should also be noted that this interlocking manifests itself in the percolating network of hydrogen bonds that can be observed in Fig. 2. Thus it appears to be an issue of packing: the shorter the side chain, the closer the molecules can be packed, and therefore the greater the number of bonds that can be made between functional groups on neighboring molecules. Considering an example where one has a large side chain, then one can see that the packing arrangements are limited to the case where the functional groups can hydrogen bond with a molecule opposite, and perhaps form a bond to a molecule lying parallel to itself, resulting in "layer formation." This is a direct result of the orthorhombic nature of L-alanine compared to the monoclinic structures of the other three crystals under consideration. Perhaps unsurprisingly, L-leucine appears to be largely similar, and one may make the same observations for this structure as have been made for the previous ones.

Current methods of carrying out calculations on molecular crystals[28] frequently use free molecule structures and parameters, such as charge densities, in order to simulate the crystal, i.e., a free molecule is placed in the field due to the other molecules within the unit cell and then a calculation is carried out: the above raises serious issues about the validity of such an approach, for clearly it is inappropriate to assume that the molecule does not relax its internal degrees of freedom in the solid state. This is in agreement with the work of Puschnig and co-workers, who have discussed the importance of treating the full three-dimensional symmetry of the crystal in such calculations.[29,30]

IV. ELECTRONIC STRUCTURE

In Fig. 4 we show the band structure for L-valine. Despite the existence of the well-known band gap problem[19] (DFT

TABLE V. Theoretically determined atomic positions for L-valine compared to the experimentally determined positions (Ref. 23)

	x		*y*		*z*	
	Experiment	Theory	Experiment	Theory	Experiment	Theory
H(1)	0.678	0.690	0.439	0.461	0.385	0.387
H(2)	0.541	0.520	0.379	0.384	0.409	0.408
H(3)	0.643	0.649	0.302	0.281	0.486	0.492
H(4)	0.569	0.559	−0.021	−0.052	0.379	0.381
H(5)	0.523	0.517	0.255	0.268	0.226	0.231
H(6)	0.514	0.507	−0.078	−0.070	0.091	0.095
H(7)	0.612	0.622	−0.243	−0.259	0.167	0.174
H(8)	0.467	0.457	−0.181	−0.190	0.210	0.220
H(9)	0.690	0.683	0.290	0.290	0.081	0.090
H(10)	0.760	0.762	0.380	0.405	0.195	0.210
H(11)	0.792	0.800	0.128	0.096	0.152	0.164
C(1)	0.779	0.779	−0.031	−0.040	0.363	0.371
C(2)	0.634	0.636	0.083	0.078	0.347	0.351
C(3)	0.597	0.599	0.132	0.126	0.222	0.232
C(4)	0.543	0.543	−0.111	−0.113	0.167	0.177
C(5)	0.719	0.718	0.242	0.237	0.157	0.171
N	0.624	0.624	0.326	0.318	0.412	0.413
O(1)	0.873	0.871	0.112	0.105	0.400	0.409
O(2)	0.794	0.795	−0.262	−0.265	0.335	0.345

commonly underestimates band gaps), the qualitative topology of DFT-obtained band structures is generally considered to be correct, and we are therefore justified in using this concept to understand the electronic structure. Certain features are immediately apparent, namely, the wide band gap of 5.27 eV. It may also be noted that the occupied states display little dispersion, and therefore are likely to correspond closely to the molecular orbitals, i.e., to be localized to individual molecules within the unit cell, as one would expect for a molecular crystal. It implies that the intermolecular attraction, and hence the geometric structure, is not primarily dictated by electronic bonding, differing from that of the isolated molecules, but is rather due to either van der Waals forces, hydrogen bonding, or dipole-dipole interactions. The

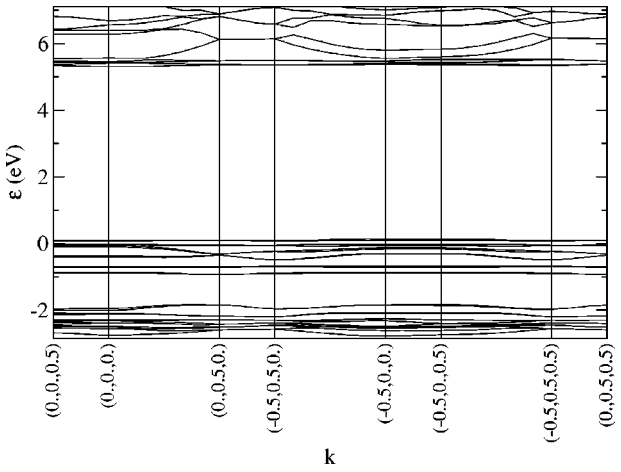

FIG. 4. The electronic band structure of L-valine.

nonzero dipole moment of valine, coupled with the polar nature of the functional groups involved, and their known tendency to form hydrogen bonds, strongly suggests that dipole-dipole and hydrogen-bonding interactions may be expected to dominate. This provides an explanation of the point alluded to in the previous section concerning the zwitterionization of L-valine in the solid state: namely, that the only way in which a stable crystal may form is by increasing the dipole-dipole interaction by zwitterionization: this is because of the localization of the electron charge distribution to individual molecules. Further, zwitterionization has implications for the nature of the hydrogen bond between the amino and carboxy groups. In principle, it would seem that a hydrogen bond could form between these groups irrespective of whether the functional groups are ionized or not; however, the hydrogen bond that forms is stronger if the zwitterionization occurs. This can be understood if one examines the geometry of the molecules involved: donation of a proton to the amino group from the carboxy group allows a relatively straight hydrogen bond to form: this is characteristic of a strong bond, whereas such a straight and hence strong bond could not form if the zwitterionization did not occur.

A Mulliken population analysis,[31] although not based upon a unique decomposition of the electronic charge density, allows us to place this argument upon a more quantitative basis: the nitrogen atoms have a negative overall charge of $-0.77e$, as opposed to a charge of $-0.68e$ on the oxygen atoms. Such a result is indicative of the nitrogen atom being more electronegative than the oxygen atom, and suggests that the proton bonded to the nitrogen atom will be deshielded more than it would be if it were covalently bonded to the oxygen atom in the carboxy group. This

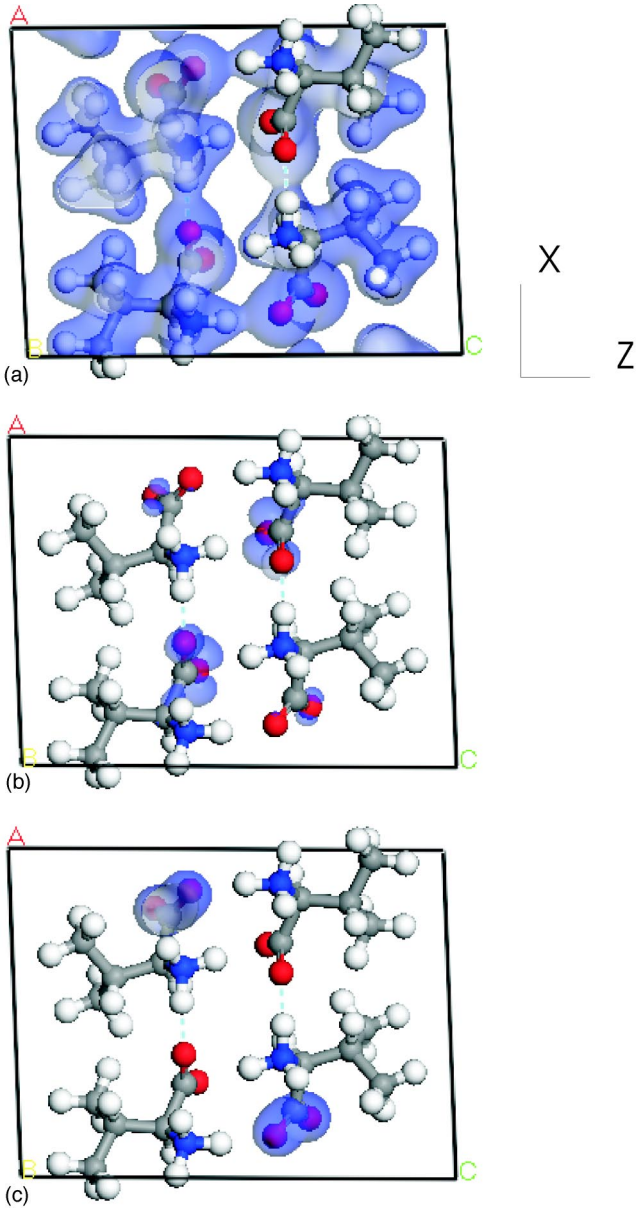

FIG. 5. The ground-state density of valine (a), highest occupied orbital (b), and lowest occupied orbital (c).

deshielding then can account for the formation of a stronger hydrogen bond. It is further interesting to note that by zwitterionization, the possibility occurs for the carboxy group to take part in more hydrogen bonds than would otherwise be the case, as can be seen in Fig. 1. This in itself will act as a further stabilizing influence.

Examination of the electron charge density of L-valine is supportive of the above conclusions. Figure 5(a) shows a constant density isosurface, and as expected, it displays a marked localization on individual molecules. It may also be noted that some level of charge overlap does occur between the amide and carboxy functional groups of neighboring molecules. This is suggestive of a hydrogen bond forming between these molecules, via these oppositely ionized functional groups, and is consistent with the discussion above on the reasons for zwitterionization occurring in the solid state.

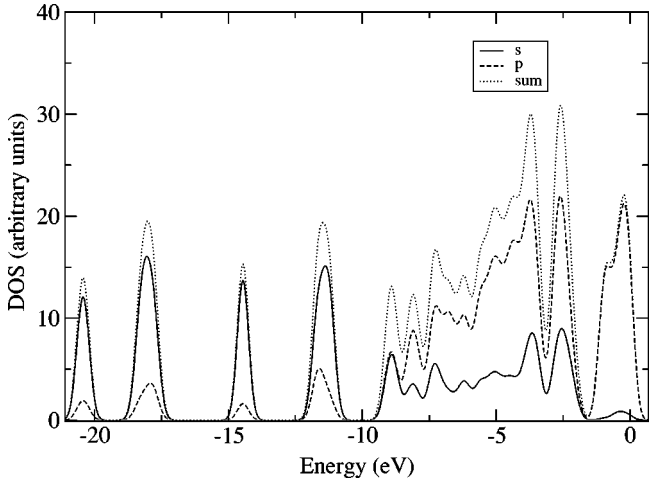

FIG. 6. The partial density of states for L-valine; Gaussian smearing width of 0.2 eV applied.

We can turn to a bond population analysis of the O-H bonds to strengthen the above analysis. The shortest bonds have a length of 1.68 to 1.7 Å; if we guardedly use as a criterion for hydrogen bonding the requirement that the sum of the van der Waals radii of the atoms involved must be greater than the separation of the atoms, then this indeed does qualify as a hydrogen bond. The bond contains a population of 0.1*e*–0.14*e*.

In Figs. 5(b) and 5(c) we show the density corresponding to the highest and lowest occupied orbitals of the system respectively. It can be seen that the highest occupied orbital appears to be "*p*-like" in nature, which is consistent with the partial density of states shown in Fig. 6, and that this orbital is primarily localized on the oxygen atoms. These are lone pair electrons on the oxygen atoms that take place in hydrogen bonding, in line with the discussion above on hydrogen bond formation.

We mention here that the band structures for isoleucine (Fig. 7) and leucine (Fig. 8) are largely similar to that of valine. Band gaps are 4.68 and 5.05 eV, respectively. The partial densities of states and orbitals for these systems ex-

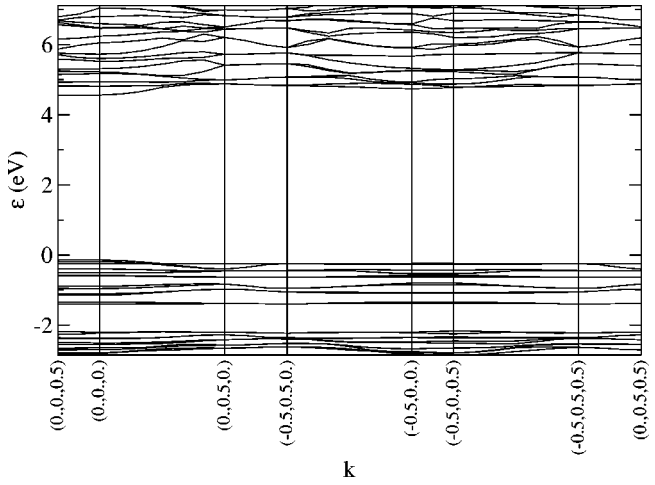

FIG. 7. The electronic band structure of L-isoleucine.

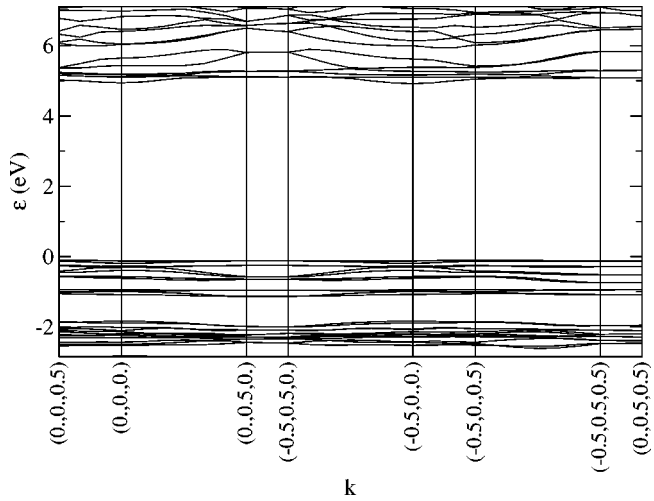

FIG. 8. The electronic band structure of L-leucine.

hibit similar behaviors to those of L-valine; for this reason, we do not present these. Given the similarity between these molecules in terms of their geometric structures and in terms of the functional groups present, these similarities are not surprising.

L-alanine (Fig. 9) deserves more detailed attention, as its band structure is markedly different. It possesses a wide band gap, of 5.07 eV, as do the other three systems examined. The band structure shows more dispersion in both the valence and conduction band manifolds possibly because of the closer packing that alanine takes part in. Examining the density in Fig. 10 does not show a noticeable degree of delocalization, which one would expect in order to produce dispersion; rather, the density is largely localized to individual molecules, with electron overlap confined to the region of the intermolecular hydrogen bonds. The density of states shown in Fig. 11 reinforces this conclusion. The carbon-oxygen bonds are slightly asymmetric, having lengths of 1.24 and 1.26 Å; suggesting that the two hydrogen bonds are not of the same strength. Population analysis suggests that the closest distance O-H bonds are of length 1.74 Å, a characteristic hydrogen bond length, and that these bonds have

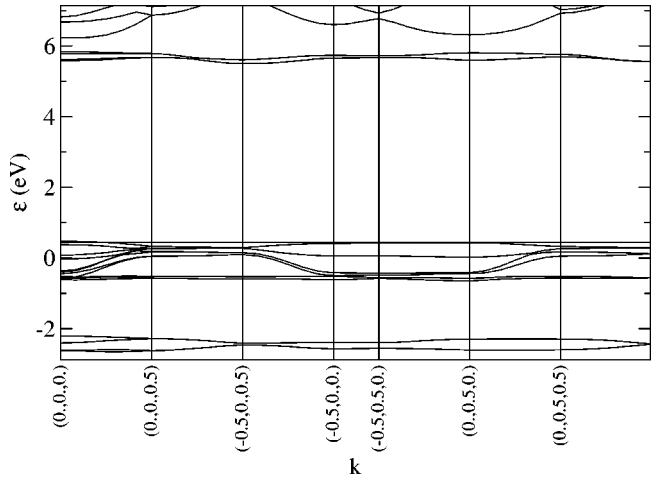

FIG. 9. The electronic band structure of L-alanine.

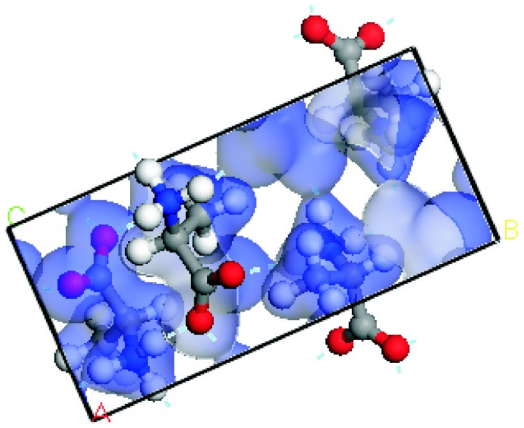

FIG. 10. The ground-state density of alanine. Note the localized nature of the density, with significant overlap only occurring in the vicinity of the amino and carboxy groups involved in hydrogen bonding.

net positive populations of 0.12*e*. It is worth noting here that this compares to hydrogen bond distances of around 1.7 Å for L-valine: this suggests that the strength of the hydrogen bonds is approximately constant from system to system. Examination of the orbitals, along with the partial density of states, reveals that they are similar to those of the other systems, with the highest orbitals being primarily oxygen *p*-type orbitals; this is again consistent with hydrogen bonding, and means that the electronic structure of L-alanine is broadly similar to those of the other systems considered.

V. CONCLUSIONS

In conclusion, this work has investigated the structural and electronic structures of the four amino acids alanine, valine, leucine and isoleucine in the solid state. Although these systems have been extensively studied by a variety of experimental techniques such as x-ray diffraction, and are broadly understood to be zwitterionic molecular crystals, stabilized largely by hydrogen bonding between the amino and

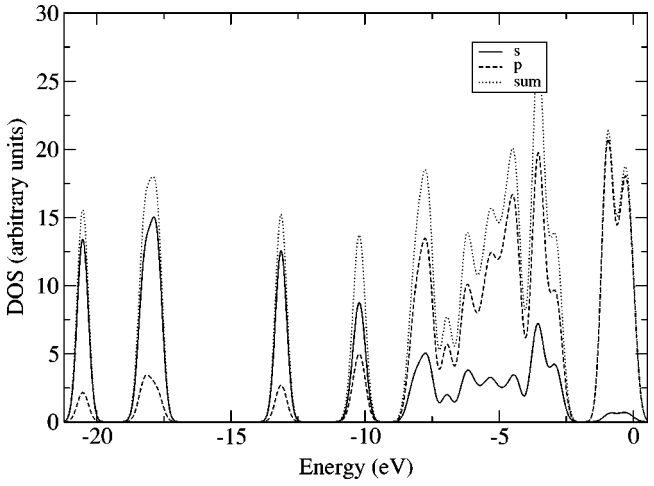

FIG. 11. Partial density of states of alanine. Gaussian smearing of width 0.2 eV applied.

carboxy functional groups, little theoretical work has been carried out to verify the validity of this picture. In this work, we have calculated, using first-principles methods, the geometric and electronic structures of these four crystals. In general, the geometric structures have been found to be in good agreement with the experimentally determined structures allowing for finite-temperature effects and van der Waals bonding. We find that the intermolecular interactions that stabilize the crystals must be primarily electrostatic in nature, for it is well known that DFT accounts well for such interactions: this has implications too for the hydrogen bonding, suggesting a primarily electrostatic nature, which is consistent with the current understanding of the nature of the typical bonds that predominate in these crystals, i.e., C═O—H—N.[25] The cases where agreement is poor between theoretical and experimentally determined lattice constants are most likely due to van der Waals forces being responsible for binding: these cases almost always involve interactions between widely separated CH_3 groups, for which other forms of interaction are unlikely to occur. A note of caution should perhaps be made here: these molecular crystals are perhaps unrepresentative of this type of crystal as a whole because, in general, dipolar and hydrogen bond interactions dominate over van der Waals interactions; this is due to the polar nature of the constituent molecules. However, in general, many molecular crystals are stabilized by van der Waals interactions;[32] the nonexistent to poor treatment of such nonlocal interactions within GGA DFT will ensure that such crystals are problematical.

The gross features of the electronic structure appear to be largely the same from system to system; this is perhaps not entirely surprising, given that they all share the same functional groups; rather the only differences appear to stem mainly from the differing side chains. As an example, the band structures, densities of states, both complete and partial, and electronic orbitals appear broadly similar. However, the differences manifest themselves in how these molecules pack together to form crystals; alanine is notably different, in the main because its much shorter side chain allows a greater degree of close packing than is possible in the other three systems, all of which possess much longer side chains. However, in all cases the bonding mechanisms appear to be the same. It is interesting to note that the structures of individual molecules within the crystal structure are so different from the gas phase conformers; this is largely due to the zwitterionization that occurs; however, it does have implications for how calculations are carried out on these systems: it is clearly not appropriate to use the isolated gas phase molecular conformers and subject these to the crystal field of the surrounding molecules, for this completely neglects the structural and electronic changes that take place upon zwitterionization.

This work has produced results that are broadly consistent with interpretative models. It is to be hoped that such first-principles techniques will find wider application to these important systems, and help to place our understanding of them and the interactions responsible for their formation on a much firmer theoretical footing. It is to this end that this work is intended to serve as a foundation for further investigations such as the vibrational and dielectric properties of these important crystals.

ACKNOWLEDGEMENT

P.R.T. would like to acknowledge the EPSRC for providing financial assistance.

[1] P. Puschnig, K. Hummer, C. Ambrosch-Draxl, G. Heimel, M. Oehzelt, and R. Resel, Phys. Rev. B **67**, 235321 (2003).
[2] W. F. Perger, Chem. Phys. Lett. **368**, 319 (2003).
[3] R. C. Haddon *et al.*, J. Phys. Chem. B **106**, 8288 (2003).
[4] M. Springborg and M. Lev, Phys. Rev. B **40**, 3333 (1989).
[5] M. S. Miao, V. E. Van Doren, and J. L. Martins, Phys. Rev. B **68**, 094106 (2003).
[6] A. G. Csaszar, J. Phys. Chem. **100**, 3541 (1996).
[7] V. Barone, C. Adamo, and F. Lelj, J. Chem. Phys. **102**, 364 (1995).
[8] A. G. Csaszar, J. Mol. Struct. **346**, 141 (1995).
[9] M. Cao, S. Q. Newton, J. Pranata, and L. Schäfer, J. Mol. Struct.: THEOCHEM **332**, 251 (1995).
[10] A. G. Csaszar, Prog. Biophys. Mol. Biol. **71**, 243 (1999).
[11] G. Zheng *et al.*, Acta Hydrochim. Hydrobiol. **55**, 729 (1997).
[12] Y. He *et al.*, Magn. Reson. Chem. **33**, 701 (1995).
[13] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[14] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[15] M. D. Segall, P. J. D. Lindan, M. J. Probert, C. J. Pickard, P. J. Hasnip, S. J. Clark, and M. C. Payne, J. Phys.: Condens. Matter **14**, 2717 (2002).
[16] L. Kleinman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).
[17] J. P. Perdew and Y. Wang, Phys. Rev. B **46**, 12 947 (1992).
[18] P. R. Tulip and S. J. Clark, J. Chem. Phys. **121**, 5201 (2004).
[19] R. O. Jones and O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989).
[20] H. J. Simpson and R. E. Marsh, Acta Crystallogr. **20**, 550 (1966).
[21] M. Coll *et al.*, Acta Crystallogr., Sect. C: Cryst. Struct. Commun. **42**, 599 (1986).
[22] C. H. Gorbitz and B. Dalhus, Acta Crystallogr., Sect. C: Cryst. Struct. Commun. **52**, 1464 (1996).
[23] B. Dalhus and C. H. Gorbitz, Acta Chem. Scand. **50**, 544 (1996).
[24] R. J. Gdanitz, Curr. Opin. Solid State Mater. Sci. **3**, 414 (1998).
[25] G. R. Desiraju and T. Steiner, *The Weak Hydrogen Bond in Structural Chemistry and Biology* (Oxford University Press, Oxford, 1999).
[26] G. A. Jeffrey, *An Introduction to Hydrogen Bonding* (Oxford University Press. Oxford, 1997).
[27] P. P. Rushton, Ph.D. thesis, University of Durham, 2002.
[28] A. Gavezzotti, Modell. Simul. Mater. Sci. Eng. **10**, R1 (2002).
[29] P. Puschnig and C. Ambrosch-Draxl, Phys. Rev. B **60**, 7891 (1999).
[30] G. Bussi, A. Ruini, E. Molinari, M. J. Calder, P. Puschnig, and C. Ambrosch-Draxl, Appl. Phys. Lett. **80**, 4118 (2002).
[31] R. S. Mulliken, J. Chem. Phys. **23**, 1833 (1955).
[32] Edgar A. Silinsh and V. Capek, *Organic Molecular Crystals: Interaction, Localization and Transport Phenomena* (AIP Press, New York, 1994).