AIP | The Journal of Chemical Physics

Dielectric and vibrational properties of amino acids

P. R. Tulip and S. J. Clark

Citation: The Journal of Chemical Physics **121**, 5201 (2004); doi: 10.1063/1.1781615
View online: http://dx.doi.org/10.1063/1.1781615
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/121/11?ver=pdfcov
Published by the AIP Publishing

Dielectric and vibrational properties of 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 23 April 2004; accepted 18 June 2004)

We calculate polarizability tensors and normal mode frequencies for the amino acids alanine, leucine, isoleucine, and valine using density functional perturbation theory implemented within the plane wave pseudopotential framework. It is found that the behavior of the electron density under external fields depends to a large extent on the geometrical structure of the molecule in question, rather than simply on the constituent functional groups. The normal modes are able to help distinguish between the different types of intramolecular hydrogen bonding present, and help to explain why leucine is found in the zwitterionic form for the gaseous phase. Calculated IR spectra show a marked difference between those obtained for zwitterionic and nonzwitterionic molecules. These differences can be attributed to the different chemical and hydrogen bonds present. Effective dynamical charges are calculated, and compared to atomic charges obtained from Mulliken population analysis. It is found that disagreement exists, largely due to the differing origins of these quantities. © *2004 American Institute of Physics.* [DOI: 10.1063/1.1781615]

I. INTRODUCTION

The Hohenberg-Kohn formulation of density functional theory (DFT) (Refs. 1 and 2) provides a powerful technique for the accurate computation of the ground state material properties of condensed matter systems. In recent years the development of density functional perturbation theory (DFPT) (Ref. 3) has allowed the response of condensed matter systems to a range of perturbations to be calculated within a density functional framework. Notable examples include the calculation of macroscopic dielectric constants in semiconductors such as silicon,[4] phonon spectra,[3] Raman scattering, and nuclear magnetic resonance chemical shift tensors.[5]

There is almost no work concerning the application of DFPT to molecular systems;[6] however, the response of molecular systems to perturbations can yield useful physical insights, as well as provide a rigorous testing ground for *ab initio* calculations. For example, polarizabilities may be determined by computing a system response to an electric field perturbation; these are of intrinsic interest due to their importance in electro-optical experiments, and are important in understanding intermolecular interactions. Similarly, calculating vibrational properties, i.e., the system response to atomic displacements, is useful in helping to understand the nature of the chemical bonds present, in characterizing and identifying molecules via vibrational spectroscopy, and in providing information about chemical reaction pathways.

The 20 naturally occurring α-amino acids are obvious candidates for the application of *ab initio* calculations. As the building blocks of proteins, they are molecules of biological and biochemical interest: understanding their physical behavior is an essential first step in unravelling both the biological functionality of proteins, and in expressing this functionality in terms of quantum mechanics. A body of *ab initio* work concerning amino acids does exist,[7–11] but this is concerned, in the main, with conformational analyses of only a small subset of these molecules. In this work, we apply DFPT to the amino acids alanine, leucine, isoleucine, and valine, in order to determine their dielectric and vibrational properties. These calculations are compared to finite difference calculations of the same in order to demonstrate the efficiency of the DFPT framework, and its applicability to problems in molecular physics and quantum chemistry. The chemical structures of the four systems under consideration are given in Fig. 1.

The paper is structured as follows: the following section provides an overview of the theoretical methods involved, while in Secs. III and IV we present the dielectric and vibrational properties, respectively. In Sec V we discuss our results, and draw our conclusions in Sec. VI.

II. THEORY

A. Polarizabilities

It is useful to elucidate the connections between the polarizability tensor and derivatives of the Kohn-Sham energy. If one considers an electric field perturbation, then the molecular Hamiltonian may be written

$$H' = H - \vec{\mu} \cdot \vec{\varepsilon}, \tag{1}$$

where $\vec{\mu}$ is the dipole moment of the molecule and $\vec{\varepsilon}$ is the applied perturbing field. Evidently

$$\mu_\alpha = -\frac{\partial E}{\partial \varepsilon_\alpha}, \tag{2}$$

where E is the molecular energy, and $\alpha = x,y,z$, where x, y, and z are Cartesian directions. For a molecule, the dipole moment may be written as a function of the applied field

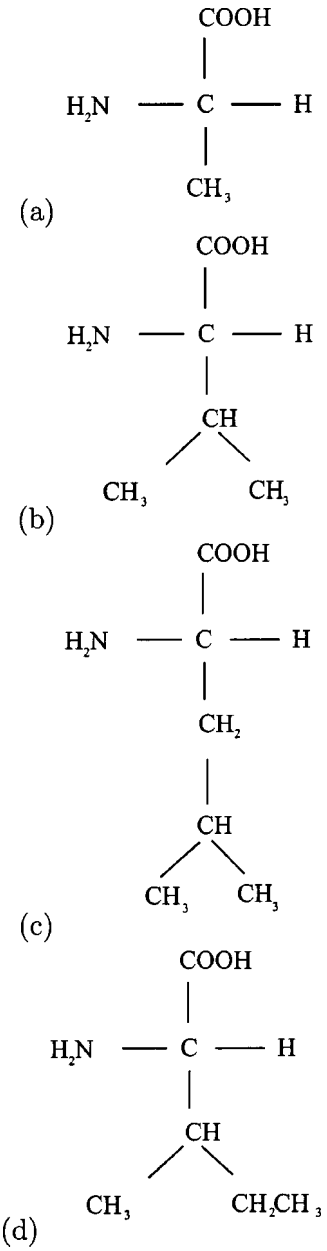

FIG. 1. Structures of (a) alanine, (b) valine, (c) leucine, and (d) isoleucine.

$$\mu_\alpha(\vec{\varepsilon}) = \sum_\beta \alpha_{\alpha\beta}\varepsilon_\beta + \cdots . \tag{3}$$

It then follows that one can write

$$\alpha_{\alpha\beta} = -\frac{\partial^2 E}{\partial\varepsilon_\alpha \partial\varepsilon_\beta}, \tag{4}$$

giving a direct link between the polarizability tensor and the second order derivatives of the total energy. One can then calculate the polarizability by two methods: by computing the dipole moment as a function of electric field, and extracting the polarizability tensor components as the proportionality coefficients; or by directly evaluating the second order derivaive of the energy. The former method is the finite difference method. The objective of DFPT is to enable the second order energy, and thus the polarizability, to be evaluated directly.

DFPT presents a number of advantages over finite difference approaches. The finite difference approach, although trivial to implement, is aesthetically unappealing; furthermore, if one calculates the molecular dipole moment at N different field strengths, say, then one needs to carry out $3N$ single point calculations, in order to obtain the full polarizability tensor. This is in contrast to DFPT, where the full polarizability tensor can be obtained from one calculation, which in terms of computational cost, is comparable to the single point calculation. Thus DFPT is signficantly more efficient. A further advantage is provided by the ease with which, within the DFPT framework, molecular repsonses to mixed perturbations may be considered. This becomes especially significant when we consider the vibrational properties in the following section. DFPT also allows access to quantities that are difficult to determine using finite difference methods: for example, electric fields in crystalline materials can be treated with care using DFPT, but cannot be implemented using finite difference methods. DFPT is thus more powerful and flexible than simple finite difference methods.

B. Vibrational properties

The vibrational frequencies, or normal modes, may be obtained directly from the second order derivative of the Kohn-Sham energy with respect to atomic displacements. In the harmonic approximation, it is assumed that for small displacements from equilibrium, the molecular energy may be expressed as

$$E_{Tot} = E_{Tot}^{(0)} + \sum_{\kappa,\alpha}\sum_{\kappa',\beta}\frac{1}{2}\left(\frac{\partial^2 E_{Tot}}{\partial\tau_{\kappa\alpha}\partial\tau_{\kappa'\beta}}\right)\Delta\tau_{\kappa\alpha}\Delta\tau_{\kappa'\beta}, \tag{5}$$

where $\Delta\tau_{\kappa\alpha}$ is the displacement of atom κ in direction α from the equilibrium position τ_κ. It is then easy to show that the vibrational frequencies are obtained from solutions of the eigenvalue problem

$$\sum_{\kappa',\beta}\frac{1}{(M_\kappa M_{\kappa'})^{1/2}}\frac{\partial^2 E_{Tot}}{\partial\tau_{\kappa\alpha}\partial\tau_{\kappa'\beta}}e(\kappa'\beta) = M_\kappa\omega^2 e(\kappa\alpha), \tag{6}$$

where M_κ is the mass of ion κ, the $e(\kappa\alpha)$'s are atomic displacement eigenvectors, and ω is the vibrational frequency.

Displacing an atom in direction α under zero electric field conditions may also cause a change in the polarization of the molecule, i.e., its dipole moment, in direction β. This is equivalent to the change in force upon an atom in direction α caused by a macroscopic field applied in direction β with clamped nuclei, and may be described by the coefficient

$$Z^\star_{\kappa,\beta,\alpha} = \Omega_0\frac{\partial P_{mac,\beta}}{\partial\tau_{\kappa\alpha}} = \frac{\partial F_{\kappa,\alpha}}{\partial\varepsilon_\beta}, \tag{7}$$

where Ω_0 is the volume of the supercell used in the calculation, and κ labels the atoms. Such a quantity is referred to as the *Born effective charge*, and can be used to describe the well-known phenomenon of LO-TO splitting in ionic semiconductors.[12] Here, we will use such a concept as it is physically useful: it yields information on the change in dipole moment under atomic displacements. Further, it is a chargelike quantity derived from the dynamics of the system;

as such, it may prove more useful and convenient in assigning charges to atoms in molecules than conventional Mulliken analysis,[13] which is based upon a rather arbitrary partitioning of electronic charge based on atomic orbitals.

Given the effective charges and the normal mode eigenvectors, one can calculate IR absorption spectra from the oscillator strength,

$$I_m \propto \sum_{\alpha} \left| \sum_{\kappa,\beta} Z^{\star}_{\kappa,\alpha\beta} e_m(\kappa\beta) \right|^2, \tag{8}$$

where m is the mode of vibration. Calculation of *ab initio* absorption spectra allows one to unambiguously assign spectral features to particular molecular motions. This allows for transparent interpretation of vibrational spectra, particularly those of complex systems, allowing the use of spectra as molecular ''fingerprints.'' This is important for both medical applications[14] and astrophysics, where the identification of amino acids in the interstellar medium has profound implications for the origin and likelihood of life on other planets.[15] Furthermore, a knowledge of the vibrational spectrum of a molecule allows one to deduce important information about the symmetry and character of the normal modes.

III. COMPUTATIONAL APPROACH

We carry out our calculations within the plane wave pseudopotential (PWP) code CASTEP.[16] Plane wave basis sets are advantageous for a variety of reasons: they are, in principle, complete, and convergence may be improved by adjustment of one parameter, namely, the cutoff energy; they cover all space equally, and are not biased towards certain chemical bonds; in contrast to localized basis sets, perturbations leave the basis set unaltered, and thus there is no need for the additional complexities of calculating Pulay-type correction terms. They thus provide a natural basis set in which to implement DFPT calculations. Of course, such a basis set requires a periodic system; this though can be imposed artificially by placing the molecule under consideration in a periodic ''supercell;'' provided that the supercell is sufficiently large, interaction between the molecule and its fictitious periodic images will be minimized. In this work a cubic supercell of 10 Å^3 is used; this is found to have sufficient vacuum region to ensure that the energy levels are nondispersive. Plane wave basis sets used in conjunction with the generalized gradient approximation proposed by Perdew and Wang[17] for exchange and correlation are known to provide a good description of hydrogen bonded systems[18,19] and are thus suitable for high-quality *ab initio* calculations of molecular properties.

The electronic ground state is determined using a preconditioned conjugate gradients minimizer. The electronic energy levels are nondispersive, and therefore only the Γ point need be considered in carrying out integrations over the Brillouin zone. Interactions betwen the valence electrons and the atomic nuclei are treated using norm conserving pseudopotentials with cutoff energies of 1000 eV; such high cutoffs are required because of the presence of oxygen and nitrogen atoms in these systems and ensure convergence to better than 0.001 eV/atom. This adds considerably to the computational effort of the calculations.

Our calculations fall into three distinct stages: first, the geometrical structures at equilibrium are obtained by minimizing all internal degrees of freedom using the well-known Hellmann-Feynman theorem. We then obtain the polarizabilities by finite difference methods, and by application of DFPT. Last, the molecular normal modes and effective charges are determined using DFPT. A thorough discussion of DFPT may be found in Refs. 20 and 21, while an in-depth discussion of our implementation is provided in Ref. 22.

IV. ELECTRIC FIELD RESPONSE

A. Polarizabilities

In the gaseous phase amino acids are found to be neutral, although both in solution and in the solid state they are commonly found in a so-called *zwitterionic* form,[7] whereby the amino group gains a hydrogen atom donated by the carboxy group, leading to two oppositely ionized functional groups on the same molecule. Although overall the molecule is still neutral, we shall continue to use the term to describe the nonzwitterionic form. Further, the isolated molecules are known to possess a degree of conformational freedom, whereby a large number of conformers at closely separated energy levels may exist. It is well known[7] that DFT commonly obtains ground state geometries that are at odds with those obtained experimentally.

In order to validate the accuracy and efficicacy of DFPT when applied to molecular problems, we have calculated the polarizability tensors by two independent methods: by DFPT itself, and by finite differences. Finite difference PWP-DFT calculations of molecular polarizabilities have been shown[23] to be capable of yielding values in better agreement with experiment than more traditional quantum chemical methods such as Hartree-Fock with the MP2 level of correlation. They thus represent an appropriate test of the accuracy of DFPT calculations.

As can be seen in Table I, good agreement between the two methods is found not only for the average of the trace but also for individual diagonal components; such agreement is edifying, and demonstrates convincingly that DFPT can obtain molecular polarizabilities. It is likely that errors originate mainly in the finite difference treatment, as one is effectively taking a numerical derivative; the quality of such a result will depend upon the number of points used to evaluate the polarizability. It can also be seen that the agreement is at times much less satisfactory for the off-diagonal elements. This is because the off-diagonal elements, in contrast to the diagonal elements, are obtained via a nonvariational calculation. As such, they will be extremely sensitive to the quality and convergence of the calculation. It should be noted that the DFPT off-diagonal elements are constrained by symmetry, i.e., element $\alpha\beta$ is the same as $\beta\alpha$. The finite difference calculations do not have such a constraint, and thus, although in general one finds that it is satisfied to within the second decimal place, it is not always. In this case, one can consider the DFPT result to be the more reliable value. In Fig. 2 we

TABLE I. Amino acid polarizability components: finite differences (FD) vs DFPT; all values are in Å^3.

Molecule	α^{DFPT} (Å^3)			α_{av}^{DFPT} (Å^3)	α^{FD} (Å^3)			α_{av}^{FD} (Å^3)
	9.30	−0.12	−0.86	8.67	9.62	−0.1	−0.87	9.14
Alanine	−0.12	9.50	0.29		−0.16	10.28	0.33	
	−0.86	0.29	7.23		−0.89	0.25	7.53	
	18.50	−0.21	−0.12	15.37	17.95	−0.1	−0.19	15.49
Leucine	−0.21	13.03	0.63		−0.19	13.77	0.61	
	−0.12	0.63	14.57		−0.22	0.64	14.75	
	17.19	−0.61	0.01	14.73	18.02	−0.57	0.11	15.33
Isoleucine	−0.61	13.71	1.72		−0.59	14.19	1.77	
	0.01	1.72	13.29		0.1	1.80	13.77	
	14.34	0.05	−0.55	12.72	14.27	0.04	−0.52	13.03
Valine	0.05	13.18	1.36		0.26	13.75	1.40	
	−0.53	1.36	10.64		−0.55	1.45	11.06	

show typical curves displaying dipole moment as a function of applied field for alanine.

The first order electronic density gives the linear variation in the electronic density caused by the external perturbing field. In Figs. 3 and 4 the first order densities obtained by DFPT calculations for alanine and valine are shown. We present these as representative of the types of behavior found. The contributions of the carboxy and amino functional groups may be clearly seen; what is equally clear is that these functional groups are not equally polarizable in different molecular environments; we shall now discuss these contributions in more detail, and attempt to elucidate possible connections between molecules in terms of shared geometric and structural factors.

B. Relation to geometric structure

The most noteworthy structural difference between the molecules under consideration is that leucine forms a zwitterion in the gaseous state, in contrast to the other three molecules. The reason for this is connected to intramolecular hydrogen bonds. In the neutral molecules, there is a hydrogen bond that forms between the hydrogen atom of the carboxy group and the nitrogen of the amine group. The bond lengths, and populations as obtained from a Mulliken population analysis, can be seen in Table II. These suggest that the three bonds are of approximately the same strength, which is perhaps not unexpected. These intramolecular bonds will act so as to stabilize this conformer. However, the position of the amine group with respect to the carboxy group in leucine is such that it is not possible to form a hydrogen bond between these two groups of the same strength; this is essentially because the amine group does not lie in the same plane as the carboxy group. Instead, by zwitterionizing, it is possible for the amine group to form two hydrogen bonds, each bond involving different oxygen atoms in the carboxy group. It is thus energetically favorable for zwitterionization to occur. As can be seen in Table II, in leucine these hydrogen bonds are longer, suggesting that they are weaker, a conclusion that is reinforced by the substantially smaller populations of these bonds.

In support of such a hypothesis, we note that for the neutral molecules alanine and valine, a depletion/augmentation of charge may be seen in the region between the amino and carboxy groups [Figs. 3 and 4(b)], suggesting the existence of a hydrogen bond between these two functional groups. A similar feature is observed for leucine.

For valine, leucine, and isoleucine, the behavior of the electron density appears to be broadly similar in each case, which, considering the zwitterionization of leucine, is an interesting point. Given the geometric similarities between valine, leucine, and isoleucine, this is perhaps expected, but it does suggest that the geometric structure plays a part in helping to determine the electronic response to external fields.

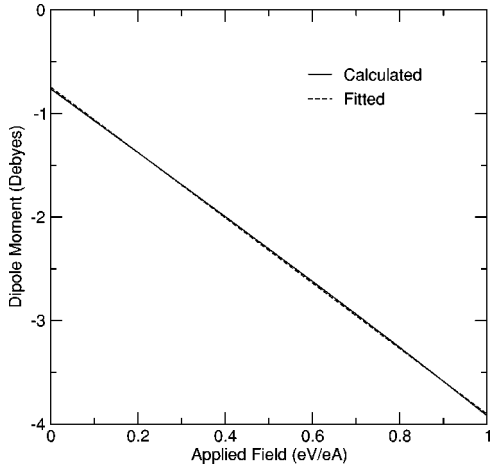

FIG. 2. The variation in the dipole moment $\mu(x)$ as a function of applied field in the x axis for alanine. The fit used is a linear fit.

TABLE II. Hydrogen bond lengths and populations. Bond 1 refers to the bond between the deprotonated oxygen atom and its nearest neighbor hydrogen; similarly, bond 2 refers to the bond between the remaining oxygen atom and its nearest neighbor hydrogen.

	Length (Å)	Population (e)
Alanine	1.81	0.12
Leucine (1)	2.49	0.03
Leucine (2)	2.71	0.01
Isoleucine	1.72	0.14
Valine	1.74	0.13

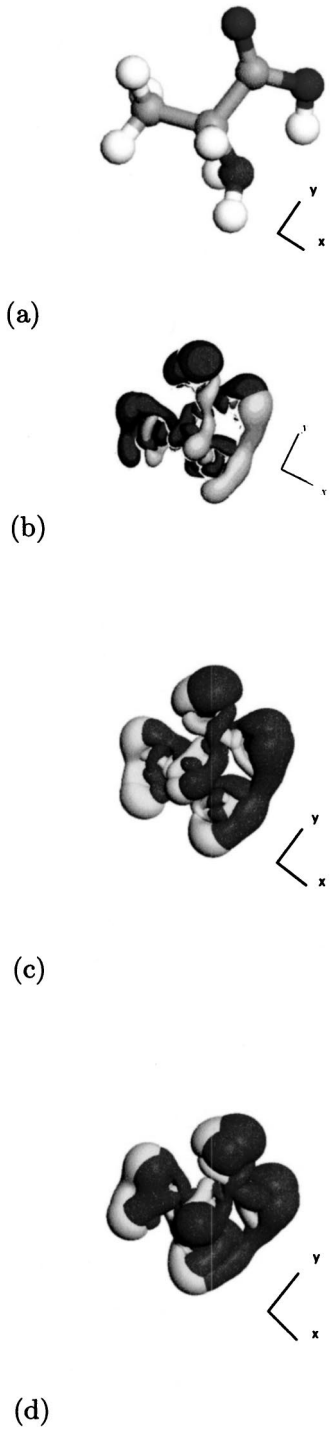

FIG. 3. Alanine: first order electron densities. (a) shows the molecular geometry used, while (b), (c), and (d) show responses to perturbations in the x, y, and z axes, respectively. The x and y axes are marked, and the z axis is out of the plane of the page. Dark gray represents where charge is being displaced to, while light gray represents a depletion of charge caused by the applied field.

Reinforcing this belief is the observation that in alanine, although neutral, as are isoleucine and valine, the electron density response is rather different. This is significant, and suggests that schemes such as that of Bader[24] based upon the partitioning of the molecular polarizability into transferrable environment- and geometry-independent functional group contributions are inappropriate for these molecules.

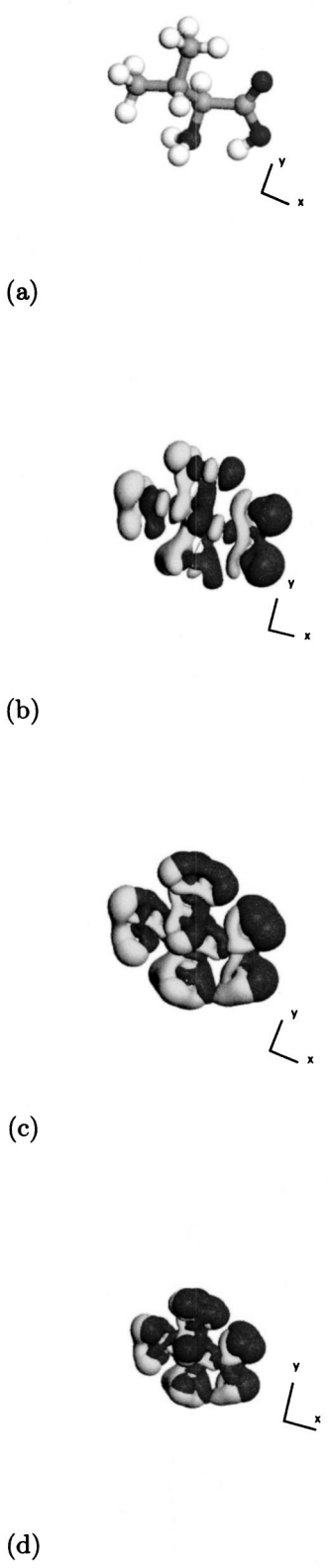

FIG. 4. Valine: first order electron densities. (a) shows the molecular geometry used, while (b), (c), and (d) show responses to perturbations in the x, y, and z axes, respectively. The colors have the same signficance as in the previous diagrams.

V. RESPONSE TO ATOMIC DISPLACEMENT

A. Vibrational properties

In Table III, we show the normal mode frequencies and assignments as calculated using DFPT. DFPT is known to

TABLE III. Amino acid normal modes; all frequencies are in cm^{-1}. The notation for the various motions of atoms within the normal modes is defined as follows: ν, stretching; δ, bending; ω, wagging; ρ, rocking; τ, torsion; s, symmetric; as, asymmetric.

Alanine		Isoleucine		Leucine		Valine	
Freq.	Assign.	Freq.	Assign.	Freq.	Assign.	Freq.	Assign.
188	τCOOH	100	Sidechain rocking	92	Sidechain twist	110	Sidechain rocking
264	τCH_3	127	Sidechain ''breathing'' mode	140	Sidechain rocking	159	''Breathing'' mode
278	δCCC	144	τCOOH	146	Sidechain rocking	231	Sidechain flexing
385	δCOH; δCNN	193	Sidechain flexing	199	Sidechain flexing	265	Sidechain flexing
433	ρNH_2	232	Sidechain flexing	254	ρCH_3; ρNH_3	272	δCHCHCOOH
438	δCCN	286	δCCH_3	265	tw CH_3; δCN	324	Sidechain flexing
529	δCCO	312	Sidechain flexing	274	tw CH_3	355	δC-O; $\delta CHCH_3$
581	δCN; δCOH	326	ρCH_3	310	tw CH_3	365	δCH-NH_2; δC-O
747	ωCOO	365	Sidechain rocking; ρCOOH	329	ρCH	443	As above
804	νCCOOH	383	δCNH_2	350	ρCH_2	449	ρNH_2
879	δCN	433	νNCOOH	407	$\delta_s CH_3CHCH_3$	467	Sidechain twisting
939	δCN	464	ρNH_2	441	tw NH_3	547	δC=O
995	δOH	495	Sidechain flexing	468	$\delta_s CH_3CHCH_2$	593	ρCOOH
1018	δCCC	555	δCOOH	547	$\delta CHCH_2$	715	δCOOH
1077	νCH_3CH	628	δCOH	679	ρCNH_3	835	$\nu_s CH_3CHCH_3$; νC-C
1128	νCN	729	$\delta COOH_s$	780	νCNH_3	838	δCCOOH
1180	νCOH	765	$\delta_{as} CH_2$	810	ρCH_2; δCCOO	905	$\nu_s CH_3CHCH_3$; δCN
1220	δCH_3CH	827	Sidechain stretching	819	ν_sCH; νCCOO	932	νCHCH
1285	δOH	857	Sidechain rocking	855	$\nu_s CH_3CHCH_3$	967	$\nu_{as} CH_3CHCH_3$
1362	τCN	881	$\rho_{as} CH_2CHCHN$; νCNH_2	887	τCCOO	998	νCCCC
1386	νCCH_3	950	Sidechain flexing; δCNH_2	923	$\nu_{as} CHCH_2CHNH_3$	1033	δOH
1423	νC-O; ρOH	976	ν_{as}CHCHNCOOH	934	$\nu_{as} CH_3CHCH_2$	1053	νCN
1463	$\delta_{as} CH_3$	1001	$\nu CHCH_3$; δCH_2CH_3	966	$\nu CHCH_3$	1093	νCHCH
1471	$\delta_{as} CH_3$	1040	δOH	1000	νCH_2CHNH_3	1129	$\nu_{as} CHCH_3$; δCCCOOH
1635	δNH_2	1053	$\rho CHNH_3$	1048	νCH_2CHNH_3	1160	ν_sCCC
1723	νC=O	1073	νCH_2CH_3	1084	νCH_2CHNH_3	1189	ν_{as}CCC
2969	$\nu_s CH_3$	1104	νCN	1121	$\nu CHCH_3$	1202	δCHCH
2976	$\nu_s CH_3$	1136	ρCH_2CHCHN	1175	$\nu_s CHCH_2CHNH_3$	1255	ρCH
3042	$\nu_{as} CH_3$	1146	νCOH; sidechain rocking	1191	$\nu CHCH_2$	1319	ρCH
3069	$\nu_{as} CH_3$; νOH	1163	νCOH; sidechain rocking	1218	ρCH	1340	δCH
3075	$\nu_{as} CH_3$; νOH	1201	τCCOOH	1257	ωCH_2; ρCH	1353	δCH
3315	$\nu_{as} NH_2$	1235	$\rho CHCHCH_2$	1266	νCCOO	1383	$\delta_s CH_3$
3400	$\nu_{as} NH_2$	1277	$\delta CHCHCH_2$	1307	$\delta CHNH_3$	1407	$\delta_s CH_3$
		1318	νCHCH	1331	ρCH	1431	νO-H; δOH
		1327	$\nu CHCH_3$ δC-H	1350	ρCH	1458	$\delta_{as} CH_3$
		1354	Sidechain stretching	1368	$\nu CHNCH_2$; νCH_3CHCH_3	1461	$\delta_{as} CH_3$
		1358	$\tau CHNH_2$	1377	Sidechain flexing	1476	$\delta_{as} CH_3$
		1388	νCCH_3; νCH_2CH_3	1401	νCH_3CH	1490	$\delta_{as} CH_3$
		1399	νCCH_3; νCCH_2CH_3	1412	$\delta_s NH_3$	1630	δNH_2
		1436	νOH; ρOH	1439	$\delta_s CH_2$	1709	νC=O
		1455	δCH_2	1462	$\delta_s CH_3$	2911	νCH
		1469	$\delta_s CH_3$	1467	$\delta_{as} CH_3$	2947	νCH
		1471	$\delta_s CH_3$	1484	$\delta_{as} CH_3$	2963	$\nu_s CH_3$
		1480	$\delta_{as} CH_3 \delta CH_2$	1505	$\delta_{as} CH_3$	2967	$\nu_s CH_3$
		1496	$\delta_{as} CH_3$	1578	νC-O	2988	νO-H
		1630	δNH_2	1626	δNH_3	3025	$\nu_{as} CH_3$
		1705	νC=O	1633	νC=O	3031	$\nu_{as} CH_3$
		2910	νCH	2927	$\nu_s CH_3$	3048	$\nu_{as} CH_3$
		2933	$\nu_s NH_2$; νCH	2935	νCH	3091	$\nu_{as} CH_3$
		2954	νCH	2951	$\nu_s CH_2$	3319	$\nu_s NH_2$
		2960	νOH	2962	$\nu_s CH_3$	3413	$\nu_{as} NH_2$
		2963	$\nu_s CH_3$	2995	$\nu_{as} CH_3$		

TABLE III. (*Continued.*)

Alanine		Isoleucine		Leucine		Valine	
Freq.	Assign.	Freq.	Assign.	Freq.	Assign.	Freq.	Assign.
		2969	$\nu_s CH_3$	3007	$\nu_{as} CH_2$		
		2984	$\nu_{as} CH_2$	3023	$\nu_{as} CH_3$		
		3031	$\nu_{as} CH_3$	3025	νCH		
		3036	$\nu_{as} CH_3$	3028	$\nu_{as} CH_3$		
		3050	$\nu_{as} CH_3$	3042	$\nu_{as} CH_3$		
		3081	$\nu_{as} CH_3$	3160	$\nu_s NH_3$		
		3319	$\nu_s NH_2$	3241	$\nu_{as} NH_3$		
		3409	$\nu_{as} NH_2$	3282	$\nu_{as} NH_3$		

provide accurate vibrational frequencies for both crystalline and molecular systems[6,25] without the necessity for any systematic scaling of the normal mode frequencies. Although one would expect anharmonicity to be important for the large-amplitude, low-frequency modes, along with the -CH, -NH, and -OH vibrations, in practice we calculate the response to infinitesimal displacements, for which the harmonic approximation is adequate. Further, DFPT is a ground state theory, and does not compute the response of the higher excited vibrational states for which anharmonicity would be more important. Based upon this, the accuracy and the reliability of the assignments presented here will be good. It should be borne in mind that when comparing with experimental results, the harmonic approximation may not be valid unless the results correspond to low-temperature measurments. Further, as alluded to above, the ground-state conformers as determined by DFT are often at odds with experimentally determined conformers; this makes direct comparison between experimental and theoretical values of the normal modes difficult. The results presented are therefore predictive rather than comparative.

The normal modes may be used to investigate the nature of the intramolecular hydrogen bonds present in these systems. It is interesting to note that in alanine and isoleucine, there are modes in the range 380–465 cm^{-1} that correspond to oscillations of the hydrogen bond existing between the amine and carboxy groups. The fact that these oscillations occur in the same frequency range for both molecules suggests that this particular hydrogen bond is of around the same strength in both of these molecules. This is in agreement with the population analysis and bond lengths described in the preceding section. In valine the same oscillatory motions are observed in the frequency range 350–450 cm^{-1}. Examining Table II indicates that in leucine, oscillations in this frequency range do not correspond to hydrogen bond distortions, rather they consist of oscillations of the carbon sidechain, although some of these oscillations suggest the existence of a weak hydrogen bond between the hydrogens on the C_β atom and the carboxy group. In leucine we observe oscillations of the three-center hydrogen bond between the amine and carboxy groups in the frequency range 198–266 cm^{-1}; this suggests a weaker hydrogen bond than in the neutral molecules, in agreement with our previous discussion. The identification of these hydrogen bonds is consistent with the discussion in the preceding section regarding the reasons for the zwitterionization of leucine.

Knowledge of the normal modes and the effective charges for the molecules under consideration allows the IR absorption spectra to be calculated. In Figs. 5 and 6 we present the spectra of isoleucine and leucine, respectively. The IR-active modes are obtained using the oscillator strengths calculated according to Eq. (8). A Gaussian thermal

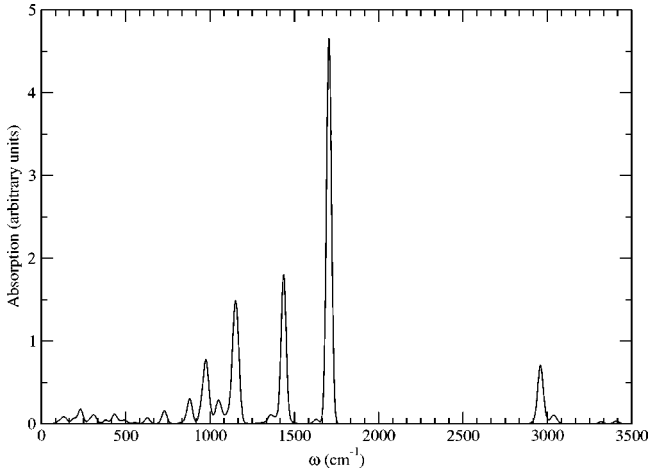

FIG. 5. IR absorption spectrum for isoleucine. Thermal Gaussian broadening corresponding to $T=300$ K.

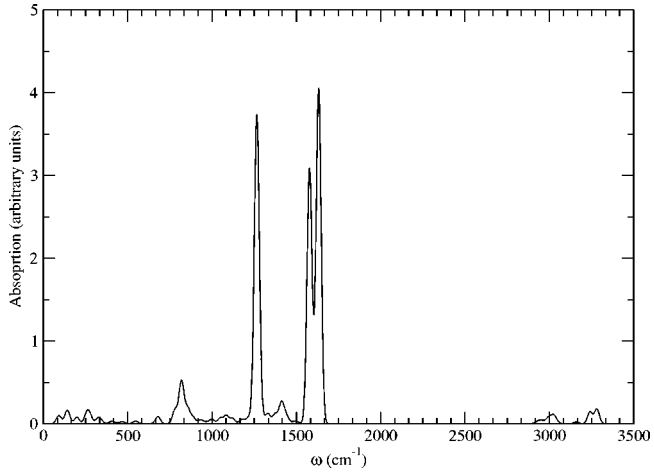

FIG. 6. IR absorption spectrum for leucine. Thermal Gaussian broadening (300 K).

TABLE IV. Effective charges and nominal atomic charges from population analysis for the molecules considered in this work. By average effective charge, we mean the average of the trace. $O(1)$ is the oxygen atom bonded to a hydrogen; $O(2)$ is double-bonded to the carbon chain. For the zwitterionic leucine, $O(1)$ is the deprotonated oxygen.

Molecule		$Z_{\kappa,\beta\alpha}$(e)			Z_{av}(e)	Z_{nom}(e)
		−0.61	−0.03	0.15	−0.49	−0.92
	N	0.03	−0.58	−0.69		
		0.21	−0.06	−0.27		
		−1.09	0.15	0.39	−0.78	−0.66
Alanine	*O*(1)	−0.25	−0.77	0.075		
		0.39	−0.06	−0.49		
		−0.80	0.25	0.20	−0.74	−0.57
	O(2)	0.37	−0.94	−0.19		
		0.20	−0.11	−0.46		
		−0.67	0.10	−0.01	−0.52	−0.90
	N	0.15	−0.49	−0.27		
		0.02	−0.26	−0.40		
		−0.92	0.37	0.14	−0.83	−0.66
Isoleucine	*O*(1)	0.12	−0.94	0.48		
		−0.19	−0.35	−0.63		
		−0.88	−0.17	−0.29	−0.77	−0.57
	O(2)	−0.06	−0.70	−0.33		
		0.23	−0.39	−0.75		
		−0.35	−0.02	−0.05	−0.26	−0.85
	N	−0.12	−0.13	0.08		
		0.04	0.12	−0.30		
		−0.89	−0.08	−0.13	−0.90	−0.70
Leucine	*O*(1)	0.11	−1.06	−0.46		
		−0.02	−0.42	−0.76		
		−1.34	0.45	0.22	−0.94	−0.69
	O(2)	0.39	−0.84	−0.27		
		0.21	−0.30	−0.63		
		−0.63	0.01	0.10	−0.50	−0.91
	N	0.05	−0.61	−0.22		
		0.12	−0.18	−0.27		
		−1.00	0.37	0.24	−0.83	−0.66
Valine	*O*(1)	0.04	−1.04	−0.30		
		0.02	−0.34	−0.44		
		−0.77	−0.29	−0.09	−0.75	−0.57
	O(2)	−0.14	−1.01	−0.29		
		−0.04	−0.29	−0.47		

broadening using a Boltzmann factor is then be applied in order to produce a full spectrum.

The spectrum for isoleucine is representative of that of the nonzwitterionic molecules, and therefore we only present this one, and discuss its features with regard to those exhibited in the zwitterionic case. It is interesting to note the contrast between this and the spectrum calculated for leucine, which is zwitterionic. Examination of the phonon eigenvectors for the individual peaks allows one to assign peaks to actual atomic motions. The peak at around 3000 cm^{-1} corresponds to oscillation of the O—H bond. As expected, no such peak appears for the zwitterion leucine. The large peak present at around 1600 cm^{-1} is due to stretching of the C═O bond. This peak does occur for leucine, but it is striking that in this case, a doublet forms, with a slightly smaller, but closely separated peak present. This doublet is due to the oscillation of the bond between the carbon atom and the deprotonated oxygen atom and the C═O bond, and it is thus not surprising that they are almost equally intense. Correspondingly, one would expect that this peak is absent from the spectra of the neutral molecules, and instead, one observes a peak at around 1400 cm^{-1} due to the stretching of the C—O bond and "rocking" of the O—H bond. This peak is around half as intense as that of the C═O bond stretching. In the zwitterionic case, one observes another strong peak at 1266 cm^{-1}; this is due to simultaneous excitation of the C_α—C bond and the carbon-oxygen bonds within the carboxylyic acid group. In the nonzwitterionic case, a peak around half as small is observed at around 1150 cm^{-1} corresponding to oscillations of the carbon sidechain, including stretching of the C—O bond.

B. Effective and Mulliken charges

In Table IV, we present the effective charges for the nitrogen and oxygen atoms in each molecule. It is immediately noticeable that all four molecules possess effective charges that display marked off-digaonal components. It is noteworthy that in all cases except leucine, the average effective charges are approximately the same. The off-diagonal components depend upon the orientation of the molecule concerned with respect to the Cartesian axes. This is not constant from moleucle to molecule, and therefore it is only possible to compare the traces meaningfully. It is not entirely surprising, in light of the discussion above regarding the zwitterionization of leucine in the gaseous phase, that leucine displays markedly different effective charges, and that this is particularly true for the nitrogen atom. It is easy to understand this: the effective charge is directly related to the vibrational properties of the molecule, and quantifies how the electronic structure changes under atomic displacement; it is therefore related directly to the chemical bonds present, and in leucine, the nitrogen atom bonds to three hydrogen atoms, rather than two, as is true for the other three molecules. The oxygen atoms also possess effective charges that are closer in value to each other than is the case in the other molecules. Again, this can be explained with reference to the zwitterionization of the moleucle, which results in both oxygen atoms forming double bonds with the C_α atom, rather than just one as in the neutral case. These bonds will be identical and thus one would expect to see similar, if not identical, effective charges. In support of this, it should perhaps be noted that in leucine, the nominal charges as calculated by population analysis for both oxygen atoms are almost identical (−0.69 versus −0.70); no such agreement is found for the other three neutral molecules.

It is also interesting to note that the nominal atomic charges determined by Mulliken population analysis[13] display a constancy from molecule to molecule. This is suggestive that the electronic structure of each molecule is the same. This is interesting, as it implies that the geometric structure and the sidechain of the amino acid under consideration have minimal impact upon the electronic structure of

the molecule, and that further, that this is dictated in main by the functional groups present.

It also suggests that zwitterionization does not substantially alter the population of the nitrogen atom involved, but rather only alters the population of the oxygen atom that donates a proton, as discussed above.

Discrepancies exist between the effective charges and the nominal charges. It is instructive to consider these discrepancies and to attempt to understand their origin. The discrepancies are most marked in dealing with the nitrogen atoms, as the effective charge is almost always around half of the value of the nominal charge; for leucine the disagreement is even more marked. Further, the nominal charges show, in all cases, that the nitrogen has more negative charge than the oxygen atoms. This is in agreement with nitrogen being more electronegative than oxygen. It is noticeable though that this trend is not followed by the effective charges, which in all cases show the reverse, that the effective charge of nitrogen is less than that of the oxygen atoms. Of course, it should be understood that such discrepancies may arise from the fact that the effective charge and the nominal charge encapsulate complementary, but different, aspects of the physics involved: the nominal charge is based upon a decomposition of the Kohn-Sham eigenstates into localized atomic orbitals, and as such yields information on the electronic density; the effective charge, on the other hand, yields information on the vibrational properties, and hence chemical bonding of the molecule concerned. The two are not therefore directly comparable, but rather, are more usefully used as complementary techniques. It can therefore be said that the fact that the effective charges for nitrogen are consistently less negative than those of the oxygen atoms present is due to the differences in chemical bonding.

The discrepancies occurring between the nominal charges and the effective charges do raise some interesting questions: often in molecular mechanics simulations, atomic charges obtained from Mulliken population analysis are used. This though relies upon an arbitary partitioning of charge; other partitioning techniques are equally valid. More importantly, perhaps, the results discussed, particularly leucine, seem to indicate that Mulliken population analysis does not always reflect the physics in the system under consideration: examining the nominal charges alone, one would not be able to identify leucine as being zwitterionic, but this does appear to be reflected in the effective charges calculated. Further, the effective charge is obtained from the dynamics of the system: in a simulation where atoms are subjected to external fields, it would perhaps be more appropriate to use effective charges rather than charges obtained from population analyses. This is an interesting point and deserves further investigation.

VI. CONCLUSIONS

We have calculated the dielectric and vibrational properties of the amino acid molecules alanine, valine, leucine, and isoleucine using density functional perturbation theory. We find that the polarizability tensors show good agreement with values determined by finite difference calculations, while analysis of the first order densities indicates that the contributions of the various functional groups to the polarizability are influenced by the geometrical structure of the molecule. It would thus seem that the polarizabilities of each functional group are not transferrable, rendering an additive approach to molecular polarizabilties difficult to justify.

The normal modes are broadly similar for the nonzwitterionic molecules, and differ significantly from the behavior observed in the zwitterionic case. This is most clearly seen when one examines the IR spectra for the two distinct cases. In contrast to the dielectric properties, it seems that the most important factor in determining the vibrational behavior is whether the molecule in question is zwitterionic or not; the different chemical bonds present and the resultant differences in the nature of the hydrogen bonding between the two cases result in very different spectra. Given that the normal modes are a direct reflection of the nature of the chemical bonding present, this is not surprising. The differences in the hydrogen bonding explain the zwitterionization of leucine.

We have also considered the use of the effective charge as a measure of the charge possessed by the constituent atoms of each molecule, and compared this to charges derived from Mulliken population analysis. We have found that the Mulliken charges are approximately constant from molecule to molecule, independent of zwitterionization; thus it seems to fail to reflect changes in electronic structure upon zwitterionization that one would perhaps expect. The effective charges do show a marked change upon zwitterionization, which reflects the different chemical bonds formed. It seems that agreement between the two is not always obtained, probably because each is obtained from different physics.

Understanding the properties of individual molecules is a necessary prerequisite for understanding the properties of molecular crystals. It is hoped that this work will form the basis of a study of the behavior of the amino acids in the solid state, in order to attempt to elucidate the connections between the two.

ACKNOWLEDGMENT

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

[1] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[2] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[3] S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi, Rev. Mod. Phys. **73**, 515 (2001).
[4] S. Baroni and R. Resta, Phys. Rev. B **33**, 7017 (1986).
[5] A. Putrino, D. Sebastiani, and M. Parrinello, J. Chem. Phys. **113**, 7102 (2000).
[6] P. Giannozzi and S. Baroni, J. Chem. Phys. **100**, 8537 (1994).
[7] A. G. Csaszar, J. Phys. Chem. **100**, 3541 (1996).
[8] A. G. Csaszar and A. Perczel, Prog. Biophys. Mol. Biol. **71**, 243 (1999).
[9] A. G. Csaszar, J. Mol. Struct. **346**, 141 (1995).
[10] M. Cao, S. Q. Newton, J. Pranata, and L. Schäfer, J. Mol. Struct.: THEOCHEM **332**, 251 (1995).
[11] V. Barone, C. Adamo, and F. Lelj, J. Chem. Phys. **102**, 364 (1995).
[12] Ph. Ghosez, X. Gonze, Ph. Lambin, and J.-P. Michenaud, Phys. Rev. B **51**, 6765 (1995).
[13] R. S. Mulliken, J. Chem. Phys. **23**, 1833 (1955).
[14] X. Cao and G. Fischer, Chem. Phys. **255**, 195 (2000).
[15] M. T. S. Rosado, M. L. R. S. Duarte, and R. Fausto, J. Mol. Struct. **410–411**, 343 (1997).

[16] 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).
[17] J. P. Perdew and Y. Wang, Phys. Rev. B **46**, 12947 (1992).
[18] D. R. Allan and S. J. Clark, Phys. Rev. B **60**, 6328 (1999).
[19] D. R. Allan and S. J. Clark, Phys. Rev. Lett. **82**, 3464 (1999).
[20] X. Gonze, Phys. Rev. A **52**, 1096 (1995).
[21] X. Gonze, Phys. Rev. B **55**, 10337 (1997).
[22] K. Refson, P. R. Tulip, and S. J. Clark (unpublished).
[23] S. J. Clark, G. J. Ackland, and J. Crain, Europhys. Lett. **44**, 578 (1998).
[24] K. E. Laidig and R. F. W. Bader, J. Chem. Phys. **93**, 7213 (1990).
[25] X. Gonze, D. C. Allan, and M. P. Teter, Phys. Rev. Lett. **68**, 3603 (1992).