Vibrational Circular Dichroism and IR Absorption Spectra of Amino Acids: A Density Functional Study

Zhi Ji, Rubén Santamaria, and Ignacio L. Garzón*

Instituto de Física, Universidad Nacional Autónoma de México, Apartado Postal 20-364, 01000 México D. F., México

Received: November 14, 2009; Revised Manuscript Received: January 28, 2010

With density functional theory, vibrational circular dichroism (VCD) and infrared absorption (IR) spectra are obtained at the B3LYP/CC-pVTZ level of theory for 20 α-amino acids. The contribution of different vibration modes to the IR and VCD spectra is analyzed. Overall agreement between calculated results for amino acids in gas phase with the available experimental VCD data for matrix-assisted amino acid films is found. The analysis of the calculated IR and VCD spectra indicates that the functional groups in the backbones and side chains of amino acids contribute differently to the spectra line shape. It is obtained that molecular torsions are the characteristic vibrations of the amino acids at the low-frequency regime, whereas the bending of bond angles, the out-of-plane wagging of individual atoms, and some stretching modes dominate the intermediate frequency range. Specific modes like NH_2 scissoring, CO bond stretching, and the (symmetric and asymmetric) stretching of the hydrogen atoms in the NH_2 and OH groups characterize the high-frequency regime. A general trend emerging from these calculations indicates that the ρ(OH) rocking and ν(C═O) stretching modes have the highest intensity in the VCD spectra of most amino acids.

1. Introduction

Infrared absorption (IR) spectroscopy is a powerful technique to understand the structure and dynamics of biomolecules since the line intensities and frequencies give direct information on the forces that hold the molecule together.[1] During the past few years this technique has been broadly utilized, in conjunction with quantum mechanical calculations, to identify the conformational structure of amino acids,[2] peptides,[3] and proteins[4] in *gas phase*. These investigations are useful not only to study the intrinsic properties of the biomolecules but also to provide insights into fundamental intramolecular interactions that can serve to calibrate theoretical methods.[5] Similarly, vibrational circular dichroism (VCD), defined as the differential absorption of left and right circularly polarized infrared light by vibrating chiral molecules, has also been shown to be particularly useful for the study of the conformational characteristics of amino acids in solution[6] and matrix isolated.[7] In fact, the higher sensitivity of VCD spectroscopy, together with first principles calculations, like density functional theory (DFT), is an excellent tool to investigate detailed changes in the three-dimensional structure of amino acids.[8,9]

As the basic building blocks of proteins, the amino acids are molecules of biological and biochemical interest since the understanding of their physical behavior is an essential first step in unraveling the biological functionality of proteins. With the exception of glycine, the amino acids are chiral, such that for the last 30 years they had been studied by VCD both experimentally and theoretically.[10] For example, VCD spectra have been reported by Nafie and co-workers[11,12] in the carbon−hydrogen stretching region for amino acids in aqueous solution. A common feature of nearly all of these spectra, considered as a marker band of the absolute configuration, is a positive VCD bias for the L-amino acids resulting from an intense band associated with the methine C−H stretching mode located near 2970 cm^{-1}.[12] A characteristic VCD couplet was also observed in the methine bending region (1200−1400 cm^{-1}) for several amino acids in aqueous solution.[13] An interpretation of the signs of these VCD features was provided using the ring current mechanism.[13] The VCD spectra of L-alanine and L-alanine-*N*-d_3 in water and deuteroxide were also reported between 900 and 1700 cm^{-1}, showing large VCD signals in vibrations associated with the C−H deformation modes of the methine hydrogen, whereas other vibrations exhibit small, but certainly observable VCD.[14] Aside from very intense methine deformation vibrations, VCD signals were also observed for rocking modes as well as skeletal stretching vibrations.[14] Although the previous mentioned studies[11−14] demonstrated the feasibility of observing VCD in aqueous solution, the strong absorption of H_2O, and the low solubility of some amino acids make it very difficult to undertake the VCD of amino acids in the mid-infrared region. Nevertheless, VCD spectra in the mid-infrared (1800−1200 cm^{-1}) region of two amino acids with low solubility, L-phenylalanine and L-tryptophan, and four amino acids with moderate solubility, L-alanine, L-proline, L-methionine, and L-histidine, were recently reported by Zhang and Polavarapu.[7] They used matrix-assisted amino acid films, eliminating the strong interfering water absorption existing in solution studies and making possible study of the VCD spectra of amino acids with low solubility.[7]

As far as we are aware, no investigations of VCD along the whole infrared spectral region for all amino acids have been reported to date due, in part, to several experimental difficulties related with the limited use of materials for the optical layout, the reduced spectral window of solvents, and other instrumental limitations.[15] On the other hand, nowadays it is recognized that VCD spectra can be theoretically predicted from first principles methods, providing confident information for the interpretation of existing experimental VCD spectra.[10] In this work, we report DFT calculations of the VCD spectra for 20 free α-amino acids:

* To whom correspondence should be addressed. E-mail: garzon@fisica.unam.mx. Phone: +52-55-56225147. Fax: +52-55-561651535.

10.1021/jp9108442
Published on Web 02/15/2010

alanine, arginine, asparagine, aspartic acid, cysteine, glutamic acid, glutamine, glycine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine, and valine. Although at present time there are experimental results of the IR absorption, in a limited range of frequencies, for some amino acids in the *gas phase*,[2,5,16–19] VCD spectra of amino acids in a solvent-free isolated *gas phase* have not yet been reported. Therefore, in this theoretical work we investigate the intrinsic relation between the chiral structure of free amino acids and their vibrational optical activity. Our objective is to provide insights into the fundamental intramolecular interactions capable of spectral manifestation existing in amino acids free of any environmental effects and its relation with their inherent VCD line intensities and frequencies. Since our calculations are performed for isolated amino acids, we use their neutral nonzwitterionic form, $NH_2-CH(R)-COOH$. In particular, we analyze the effect of the organic side chain group R, which distinguishes one amino acid from another, on the overall line shape of the VCD spectra. In this study, we first optimize the structures of all amino acids including glycine, which is an achiral amino acid. The harmonic vibrational frequencies, dipole strengths, and rotational strengths of the fundamental transitions of the optimized structures are predicted using the DFT method, and the IR and VCD spectra obtained hence.

The paper is structured as follows: section 2 contains a brief description of the theoretical−computational methodology. Calculated and experimental geometries for some amino acids are discussed in section 3.1, whereas in section 3.2 a comparison of their calculated vibrational frequencies with experimental results is reported. In section 3.3, we describe the calculated VCD spectra for all amino acids. A summary of this work is presented in section 4.

2. Theoretical Methodology

All DFT calculations are performed using the Gaussian03 program,[20] with the hybrid Becke three-parameter Lee−Yang−Parr B3LYP functional,[21,22] whose reliability in calculations of the ground-state geometries and vibrational properties of amino acids has been widely assessed.[23,24] We use the cc-pVTZ Dunning's basis set,[25] since this kind of Dunning's correlation consistent basis set is known to be adequate to describe both organic molecules and hydrogen-bonded systems.[23,24]

All the structures of amino acids are fully optimized starting from the coordinates found in the GSU Library.[26] A convergence threshold for the maximum force of 4.5×10^{-4} au was used for the geometry optimization in all calculations. In addition, vibrational frequencies, infrared intensities, and VCD spectra for the optimized structures are calculated. The procedure for calculating the VCD rotational strength is based on solving the Stephens's equation[27,28] using DFT, as implemented in the Gaussian03 program.[20] IR and the VCD spectra are simulated from the calculated dipole and rotational strengths values, assuming that each band was a Lorentzian waveform with a bandwidth of 7 cm^{-1}.

3. Results and Discussion

Results of the structure of cysteine and vibrational patterns of glycine are described and compared to available experimental data in Tables 1 and 2. The characteristic vibrations of the amino acids in the high-frequency regime are presented in Table 3. The goal of this first part of the work is (i) to assess the accuracy of the calculations by comparison with available experimental data, (ii) to point out individual vibrational fingerprints, and (iii) to determine the general trends on the vibrational frequency values of the amino acids. In the second part of this section, the VCD spectra of selected modes of different amino acids are discussed. These results appear in Tables 4−8 and Figures 3−7. The discussion attempts to correlate the molecular structure with the vibrational normal modes. As Supporting Information, the optimized structures (x,y,z coordinates) and the calculated vibrational frequencies, IR intensities, and VCD spectra of the 20 amino acids are provided.

TABLE 1: Bond Lengths, Bond Angles and Dihedral Angles of Cysteine[a]

	this work	X-ray data	diff
bond			
N(1)C(2)	1.456	1.484	−0.028
C(2)C(3)	1.525	1.531	−0.006
C(3)O(4)	1.202	1.255	−0.053
C(3)O(6)	1.350	1.254	0.096
C(2)C(9)	1.527	1.528	−0.001
C(9)S(10)	1.845	1.819	0.026
N(1)H(5)	1.010	1.030	−0.020
N(1)H(7)	1.012	1.030	−0.018
S(10)H(11)	1.345	1.340	0.005
bond angle			
N(1)C(2)C(3)	112.7	109.9	2.8
N(1)C(2)C(9)	110.5	111.3	−0.8
C(2)C(3)O(4)	124.0	116.7	7.3
C(2)C(3)O(6)	113.1	118.0	−4.9
C(2)C(9)S(10)	110.7	115.1	−4.4
C(3)C(2)C(9)	110.3	111.7	−1.4
O(4)C(3)O(6)	122.8	125.3	−2.5
C(9)S(10)H(11)	95.6	98.1	−2.5
dihedral angle			
N(1)C(2)C(3)O(6)	19.6	−4.47	
N(1)C(2)C(9)S(10)	53.4	74.39	
C(2)C(9)S(10)H(11)	142.4	−64.6	

[a] Lengths in angstroms and angles in degrees. X-ray data are from ref 29.

TABLE 2: Frequencies and Relative Intensities of Glycine[a]

freq[b]	freq[c]	int[b]	int[c]	assign[b]
74	[52.4]	0.02		τ(NCCO)
206		0.16		$\tau(CNH_2)$
261	[356]	0.08		δ(NCC)
463	463	0.00	0.990	δ(CCO(9))
468	498 [504]	0.22	1.431	ω(OH), $\tau(CCH_2)$
571	617 [548]	0.04	2.642	τ(NCCO), ω(CCO(9))
649	619 [607]	0.04	2.642	δ(CO(4)), δ(NCC)
830	801	0.07	3.948	ν(CC)
923	883	0.52	1.749	ν(CC), $\omega_s(NH_2)$
923	911	0.00	0.100	ω(OCO)
1125	1101	0.11	5.438	a(ν(NC), ν(CO(9)))
1161	1136	0.03	2.009	s(ν(NC), ν(CO(9)))
1186		0.00		$\omega_a(CH_2)$, $\omega_a(NH_2)$
1285	1210	1.00	0.092	ν(CO(9)), δ(OH)
1377	1373	0.07	0.092	ν(CC), $\omega_s(CH_2)$
1394	1427	0.00	0.341	$\omega_a(NH_2)$, $\omega_a(CH_2)$
1465	1429	0.02	0.341	scCH_2, νCN
1678	1630	0.06	0.315	sc(NH_2), ν(CN)
1856	1790	0.65	8.352	ν(CO(4))
3011	2944	0.08	0.826	$\nu_s(CH_2)$
3042	2958	0.03	0.826	$\nu_a(CH_2)$
3496	3410	0.01	0.313	$\nu_s(NH_2)$
3563		0.01		$\nu_a NH_2$
3783	3560	0.12	3.111	ν(OH)

[a] Frequencies in cm^{-1} and relative intensities in KM/mol. Nomenclature: a = asymmetric, s = symmetric, δ = bending, ω = wagging, ν = stretching, sc = scissoring, τ = torsion. [b] This work [c] Experimental measurements from ref 23.

TABLE 3: Characteristic Vibrations of the 20 Naturally Occurring Amino Acids in the High-Frequency Regime[a]

molecule	sc(NH_2)	ν(CO) + ω(OH)	ν(C_αH)	ν_s(NH_2)	ν_a(NH_2)	ν(OH)
ala	1652	1824	3071, 3101	3492	3574	3744
arg	1651	1814	2901	3508	3592	3743
asn	1663	1793	3011	3499	3581	3736
asp	1665	1803,1808	3017	3505	3582	3741
cys	1652	1819	2909	3517	3605	3743
glu	1652	1818,1820	2913	3509	3593	3742
gln	1670	1815	3075	3504	3577	3737
gly	1678	1856	3011, 3042	3496	3563	3783
his	1660	1841	3021, 3047	3495	3585	3495, 3509
ile	1653	1812	2899	3521	3599	3743
leu	1651	1815	2894	3504	3587	3743
lys	1650	1816	2904	3506	3590	3743
met	1651	1818	2909	3506	3590	3742
phe	1652	1821	3067, 3084	3497	3579	3744
pro		1811	3037		3568	3740
ser	1660	1807	3044, 3049	3495	3575	3740
thr	1662	1804	3054	3506	3589	3741
trp	1651	1819	3061, 3079	3500	3581	3744
tyr	1652	1820	3064, 3082	3496	3578	3744
val	1651	1805	2943	3510	3593	3736
average	1657	1816	3010	3503(3293)	3584(3369)	3743(3518)

[a] Frequencies in cm^{-1}. The nomenclature is the same as in Table 2. The two frequencies of histidine were avoided in the computation of the average frequency of the ν(OH) vibration mode. Numbers in parentheses represent scaled frequencies, for a scale factor of 0.94 (refer to the text).

TABLE 4: Experimental and Calculated Vibrational Frequencies, and VCD Intensities of Selected Modes of Alanine, Glycine, Isoleucine, Leucine, Valine, and Methionine, See Figure 3

		freq, cm^{-1}			
		exptl	calcd	VCD rot strength × 10^{44} esu^2 cm^2	assignment[a]
ala	1		589.1	52.026	ρ(OH)
	2		611.5	−47.782	δ(COOH)
	3	1117[b]	1022	−56.751	ρ(CH_3)
	4	1221[b]	1100	46.486	ρ(NH_2)
	5		1824	−51.389	ν(C═O)
gly	1		205.9	0.0	ρ(NH_2)
	2	504[c]	468.0	0.0	ρ(OH)
	3	893[c]	923.5	0.0	ω(NH_2)
	4		1284	0.0	ω(OH)
	5	1703[c]	1856	0.0	ν(C═O)
ile	1	628[d]	644.7	−101.87	ρ(OH)
	2	729[d]	726.2	72.428	side chain rocking δ(COOH)
	3	765[d]	774.2	72.489	side chain rocking
	4	881[d]	892.6	−54.606	ρ_{as}(CH_2CHCHN)
	5	1146[d]	1124	49.090	δ(COH); side chain rocking
leu	1	679[d]	644.3	−104.03	ρ(OH)
	2	810[d]	739.4	113.07	δ(COOH)
	3	819[d]	813.6	−69.876	δ_s(NH_2)
	4	1084[d]	1145	57.494	ν(CH_2CHNH_3)
	5	1377[d]	1423	55.874	side chain flexing
val	1	593[d]	650.4	−237.81	ρ(OH)
	2		801.0	59.781	ρ(NH_2)
	3	905[d]	877.6	−77.398	ν_s(CH_3CHCH_3); ρ(CN)
	4	1129[d]	1153	113.96	ν_{as}($CHCH_3$); δ(COOH)
	5	1709[d]	1805	88.524	ν(C═O)
met	1		630.4	−123.42	ρ(OH)
	2		731.7	96.445	δ(COOH); ν_{as}(SCC)
	3	1355[e]	1325	−69.992	ρ(OH); ν(CO)
	4	1404[e]	1426	89.713	ν_{as}(CCN)
	5		1818	321.21	ρ(OH); ν(C═O)

[a] The notations for the various motions of atoms within the normal modes are defined as follows: ν, stretching; δ, bending; ω, wagging; ρ, rocking; τ, torsion; s, symmetric; as, asymmetric. [b] Data from ref 14. [c] Data from ref 34. [d] Data from ref 35. [e] Data from ref 7.

3.1. Ground State Geometries. In order to determine the accuracy of the calculations, structural parameters like bond lengths, bond angles, and dihedral angles should be compared with experimental measurements. In particular, we pay attention to cysteine as a test molecule because it is a small compound containing the different types of atoms that appear in the amino acids (see Figure 1). The bond length parameters are gauged with respect to measurements from X-ray crystallographic experiments.[29] Our optimized structure of cysteine appears as an intermediate configuration between the gauche conformation

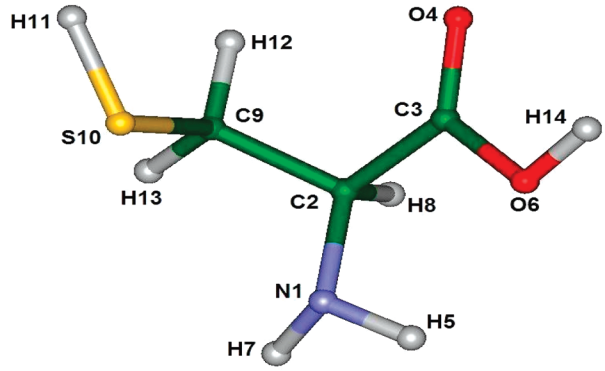

Figure 1. Molecular structure of cysteine.

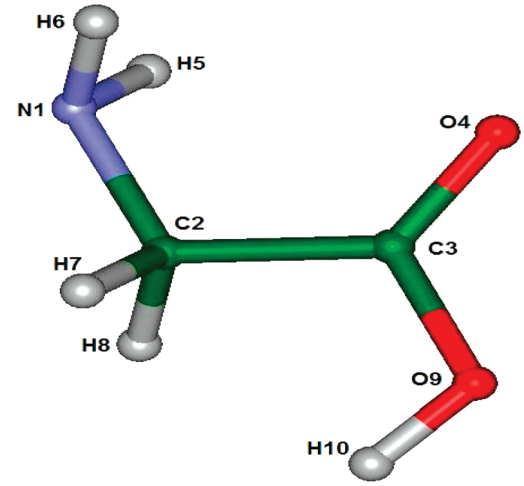

Figure 2. Molecular structure of glycine.

of L-Cys(A) and the trans conformation of L-Cys(B) presented in ref 29. These two conformers show similar bond distances and angles but different dihedral angles. The difference is attributed to the thiol groups participating in the formation of H bridges in the crystal. In Table 1, we compare the cysteine calculated structural parameters with those of the gauche conformation of L-Cys(A) because such a conformer is the predominant one in the experiment. An overall agreement between theoretical and experimental measurements is observed; in fact, the average deviations are lower than 0.03 Å for bond lengths and 3.02° for bond angles. Still, the largest deviations are observed for the dihedral angles. Since they constitute the soft degrees of freedom and require less energy than bond lengths and bond angles to be distorted, the discrepancies in dihedral angles between theory and experiment are associated to the effect of the molecular environment on cysteine in the crystal. In order to quantify the effects that different dihedral angles have in the structure of cysteine, we compare the total energies of cysteine—as obtained in the optimization process—with a cysteine structure having dihedral angles fixed to the experimental values given in Table 1. The energy of the later structure is −722.0737 au and for the former one it is −722.0921 au. The energy difference is 0.0184 au ∼ 11.6 kcal/mol, with the optimized structure being more stable. The energy difference is relatively small and indicates the great susceptibility of such degrees of freedom to be distorted by, for example, the crystal packing forces.

Despite these differences, the comparison indicates that the calculated structural parameters are in close agreement with the experimental values. Similar structural results are expected for the remaining amino acids because they all share a common backbone. The conclusion is that the DFT method at the B3LYP/cc-pVTZ level of theory provides satisfactory results in reference to amino acids geometrical parameters.

3.2. Harmonic Vibrational Frequencies. With the optimized structures, the vibrational frequencies, normal modes, and IR intensities of the amino acids in vacuum are calculated at the DFT/B3LYP/cc-pVTZ level of theory. In particular, the vibrational frequencies of glycine appear in Table 2 (Figure 2 gives the structure of glycine). The values are compared with experimental results from ref 23. At first sight, a good overall agreement is found between the calculated and experimental values; nevertheless, it is more convenient to carry a detailed analysis of frequencies by intervals.

One of the first characteristic vibrations of the amino acids that appears in the low frequency range considered from 0 up to 500 cm^{-1} corresponds to molecular torsion. The torsion is shown as an oscillation of dihedral angles formed by three main backbone bonds, where one of them is usually the N−C peptide bond. In the case of glycine (Table 2), the first few vibrations, for instance with frequencies 74, 206, and 261 cm^{-1}, illustrate the molecular torsion. Despite the common molecular torsions of the 20 amino acids in the low-frequency region, the frequencies change from one amino acid to another due to the different environments created by the R-groups. From an experimental point view, there are difficulties to determine the very low frequencies in this regime due to the low intensity of the vibrations. Still, a relatively good agreement with experimental measurements is observed in the last part of this interval.

At the intermediate frequency range (500−1600 cm^{-1}), the amino acids are subject to a great number of deformations which, in contraposition to molecular torsion, only involve a few local atoms, mostly belonging to the amino acid R groups. As an example we have the bending of bond angles and the out-of-plane wagging of individual hydrogen bonds (see Table 2). Some vibrations that call attention, especially for their simplicity, are the numerous stretching N−C, C−C, and C−O bonds. However, for the relatively similar vibration modes observed in all amino acids in this interval, the specific molecular structure of the R group greatly influences the magnitudes of the vibrational frequencies. In this regard, the intermediate frequency range is particularly attractive for the identification of individual amino acids (see tables in Supporting Information containing the frequencies in this interval). In the case of glycine, the computed frequencies and assignments in the intermediate range are in good agreement with the experimental measurements and with other DFT calculations, as well.[23a]

The last region corresponds to the high-frequency regime, above 1600 cm^{-1}. Common vibrations of the amino acids are clearly identified here (see Table 3). They correspond to the scissoring vibrations of the NH_2 hydrogen bonds, stretching of the C−O bond (with a consequent wagging of the nearby O−H bond), and the symmetric and asymmetric stretching of the hydrogen atoms in the NH_2 and OH groups. The symmetric vibrations take less energy and appear before the asymmetric ones in the spectra. They all represent characteristic vibrations of the amino acids, with similar frequency values because the vibrations involve the common parts that form the amino acid skeleton. However, some exceptions are observed (in Table 3 they appear as entries with double values or no value at all). The exceptions are associated to changes in the "usual" structure of amino acids. For instance, in the case of proline, the N atom of the peptide bond takes part in the five-membered ring. In

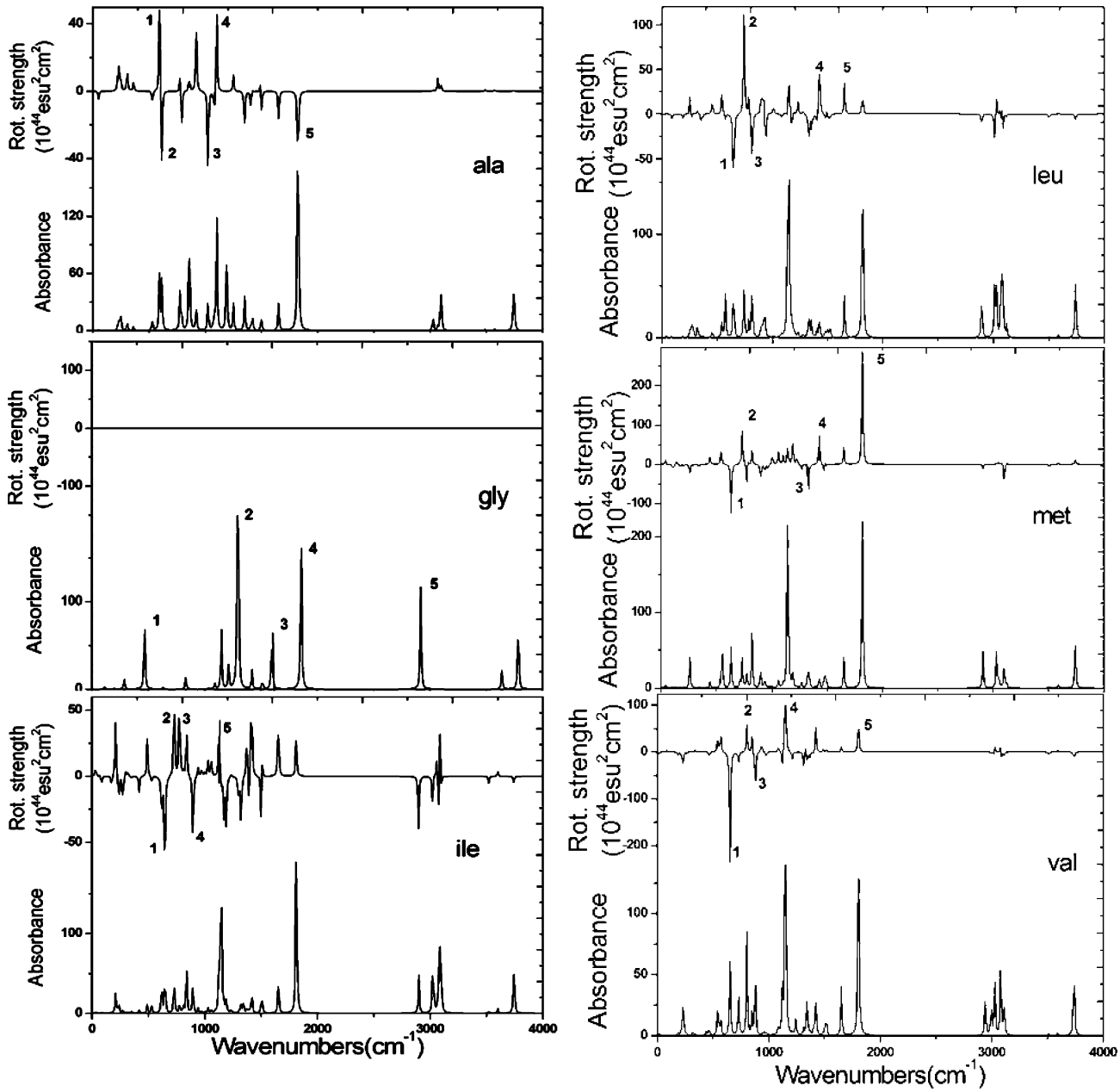

Figure 3. (a) Calculated IR (bottom) and VCD (top) spectra of alanine, glycine, and isoleucine. Labels indicate the selected bands. See Table 4. (b) Calculated IR (bottom) and VCD (top) spectra of leucine, methionine, and valine. Labels indicate the selected bands. See Table 4.

this particular case, nitrogen owns single hydrogen and no NH_2 scissoring is observed, only a single NH stretching. On the other hand, in the cases of the aspartic and glutamic acids, the presence of several oxygen atoms produces a vibrational coupling, resulting in two frequencies of the ν(CO) vibration. Other vibrational modes, like $\nu(C_\alpha H)$ in glycine, histidine, etc., show vibrational coupling attributed to nearby hydrogen atoms of the R group, which oscillate in symmetric and asymmetric manners, thus duplicating the frequencies observed of the $\nu(C_\alpha H)$ mode. In tryptophan and tyrosine the stretching vibrations are deformed as a strong component of wagging appears in these vibrations, presumably due to the presence of rings.

Despite the few exceptions discussed above, the high-frequency interval becomes ideal for the identification of the amino acids, distinguishing them from other biological compounds like the nucleic acids,[30] because it shows general vibration patterns that are specific to the peptide backbone and with frequencies that exhibit relatively little scattering with the presence of the R group of each amino acid. The similar frequencies and vibration modes in the high-frequency regime additionally point out similar bond forces of the backbone structure of all the amino acids. In this regard, average frequencies are given in Table 3 as representative frequencies of the vibrational modes described there. The $\nu(C_\alpha H)$ vibration shows the largest frequency scattering, of approximately 100 cm^{-1} with respect to the average frequency. Unfortunately, it is in the high-frequency range where the level of DFT theory used in this work overestimates the frequencies, particularly for the ν(OH) vibration. With no surprise, a similar performance of DFT theory in the high-frequency range is observed in the nucleic acid bases.[31] With a scale factor of 0.94, the ν(OH) frequency shifts from 3743 to 3518 cm^{-1}. This last result is the outcome of averaging the ν(OH) experimental values of glycine (3560 cm^{-1}[123a]), alanine (3560 cm^{-1}[124]), and *N*-methylproline (3425 cm^{-1}[132]). Interestingly, scale factors of 0.95 and 0.98 for the Watson−Crick base pairs adenine−thymine and guanine−

TABLE 5: Experimental and Calculated Vibrational Frequencies, and VCD Intensities of Selected Modes of Phenylalanine, Tyrosine, and Tryptophan, See Figure [4]

		freq, cm^{-1}			
		exptl	calcd	VCD rot strength × 10^{44} esu^2 cm^2	assignment[a]
phe	1		596.7	168.55	ρ(OH)
	2		1084	107.52	ν(CN); phenyl deformation
	3		1136	−117.50	ν(CO); ν(CN); phenyl deformation
	4		1199	−126.91	ν(CO); ν(CC); phenyl deformation
	5		1821	−619.53	ν(C═O)
tyr	1		595.3	207.77	ρ(OH)
	2		1088	107.66	$\nu(CHCH_2)$
	3		1202	−120.24	side chain flexing
	4		1820	−688.43	ν(CO)
	5		3816	158.62	ρ(OH) (phenyl)
trp	1		894.9	112.61	deformation ring
	2		1005	−66.684	$\nu(CHCH_2)$; $\nu_{as}(CH_2)$
	3		1076	58.473	$\nu CHCH_2$; deformation ring
	4	1342[b]	1350	−45.766	ν(C═O)
	5		1819	−65.131	ν(CO)

[a] See the footnote to Table 1. [b] Data from ref 7.

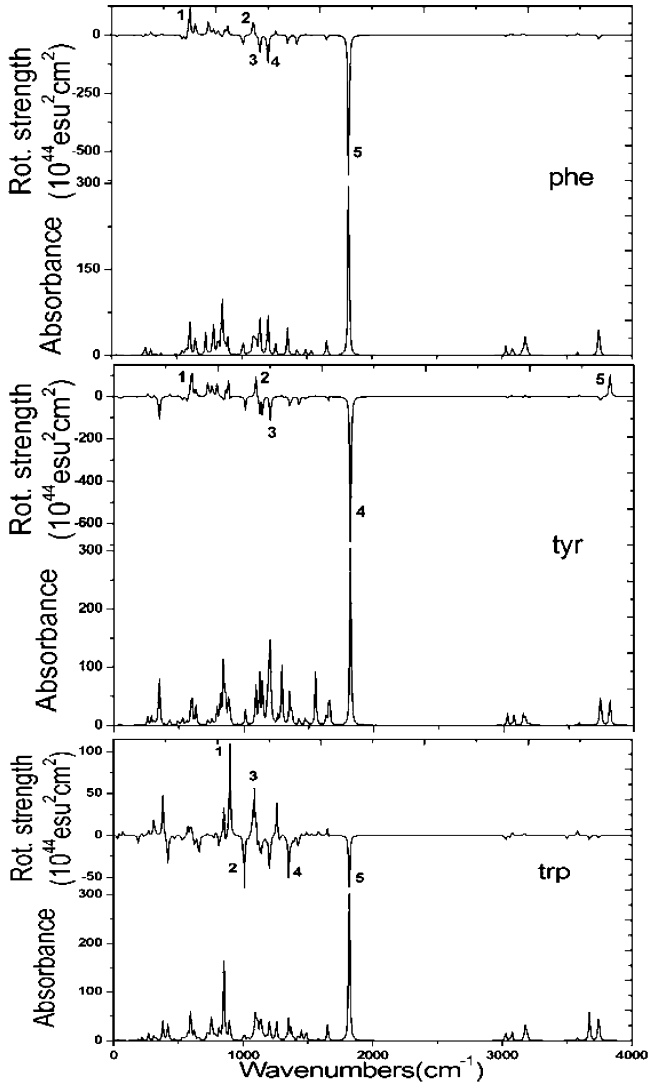

Figure 4. Calculated IR (bottom) and VCD (top) spectra of phenylalanine, tyrosine, and tryptophan. Labels indicate the selected bands. See Table 5.

cytosine were proposed, respectively, in ref 31. If we consider that in the high-frequency regime the scale factor roughly accounts for the anharmonicity effects, as well as for the inaccuracy of both the exchange-correlation functional and limited orbital basis sets that we used, then the scale factor should also be considered in the $\nu_s(NH_2)$, $\nu_a(NH_2)$, and ν(OH) vibrations. In this context, the scaled frequencies are 3293, 3369, and 3518 cm^{-1}, respectively.

3.3. VCD Spectra. The VCD spectra of the optimized structures were calculated at the B3LYP/cc-pVTZ level of theory. They are presented in Tables 4−8 and Figures 3−7 (no scale factors were used in these results). The IR and VCD spectra for each amino acid are displayed at the bottom and top halves of each figure, respectively. Most of the calculated vibrational modes have been assigned, with the exception of a few modes where skeletal stretching and rocking motions are strongly mixed. The calculated spectra may be well interpreted in terms of group frequencies and predominantly local motions. Amino acids are generally divided into groups on the basis of their side chains (R groups). The most helpful starting point, in terms of the resulting properties, is to separate amino acids into those with polar and nonpolar side chains and then to subdivide the polar amino acids according to any charge that the R group might carry. In the present study, we divide the 20 amino acids into five groups according to this criterion.

Alanine, glycine, isoleucine, leucine, valine, and methionine are nonpolar amino acids which are characterized by having only carbon and hydrogen atoms in their side chains. For alanine, the five more significant bands (those with higher positive or negative intensity) include three negative at 611.5, 1022, and 1824 cm^{-1} and two positive ones at 589.1 and 1100 cm^{-1}. The most intense band is the negative one at 1022 cm^{-1} which can be assigned to the $\rho(CH_3)$ rocking mode. The weaker band is that one at 589.1 cm^{-1}, which corresponds to the ρ(OH) rocking mode. The bands at 1022 and 1100 cm^{-1} were found in the experiments by Diem,[14] but there they were located at 1117 and 1221 cm^{-1}, respectively. Also, in that experiment, the intense methine deformation vibrations which are located at 1358 and 1306 cm^{-1} render the main contribution to the VCD spectra. However, in our study, it is unclear to find that the modes of methine deformations which locate at 1249 and 1344 cm^{-1} (without being scaled) only make a small contribution to the calculated VCD spectra. In terms of the empirical models already proposed, such as the ring current model,[13] the VCD spectra should be dominated by signals due to the methine C−H deformation vibrations, whereas the remaining vibrations are expected to exhibit much smaller VCD intensities. The present

TABLE 6: Experimental and Calculated Vibrational Frecuencies, and VCD Intensities of Selected Modes of Arginine, Histidine, and Lysine, See Figure [5]

		freq, cm^{-1}			
		exptl	calcd	VCD rot strength × 10^{44} esu^2 cm^2	assignment[a]
arg	1		641.6	−496.66	ρ(OH); $\rho(NH_2)$
	2		653.9	250.64	ρ(OH); $\rho(NH_2)$
	3		1446	−194.68	ν(CN)
	4		1622	−253.33	$_s NH_2$
	5		1815	392.31	ν(C═O)
his	1		598.2	−124.30	ω(COO)
	2		707.8	83.955	side chain flexing
	3		834.0	−124.16	ρ(OH); $\rho(NH_2)$
	4	1362[b]	1399	−95.351	ρ(OH); side chain flexing
	5	1405[b]	1408	165.90	ν(CO); ρ(OH)
lys	1		651.1	−75.029	ρ(OH)
	2		737.7	96.193	δ(COOH)
	3		850.5	−85.642	$\rho(NH_2)$
	4		1424	78.565	ρ(CH)
	5		1816	106.55	ν(C═O)

[a] See the footnote to Table 1. [b] Data from ref 7.

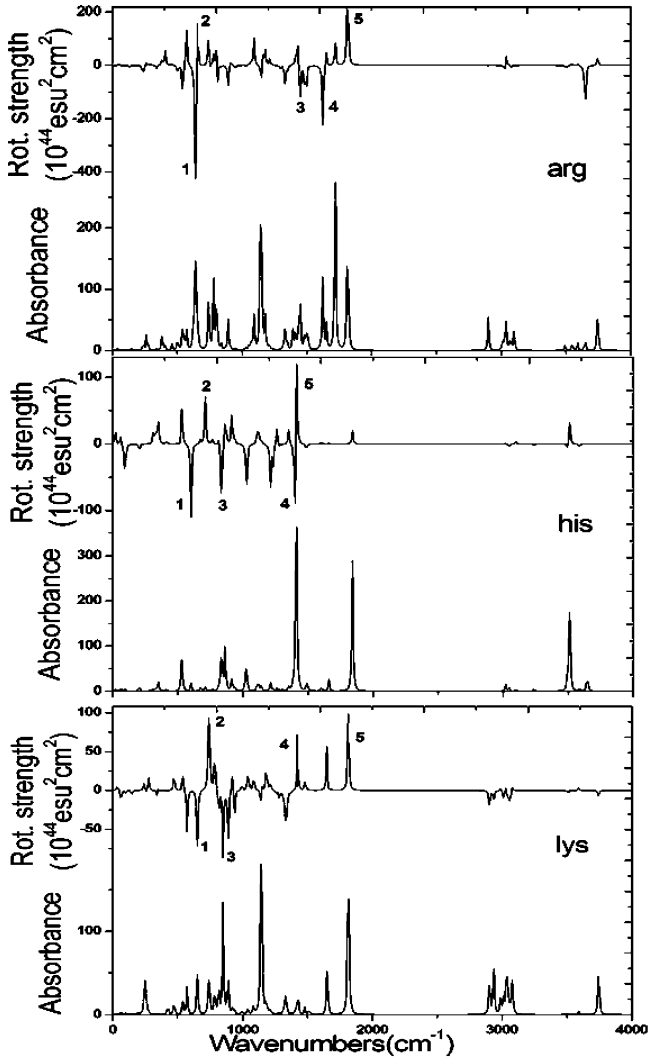

Figure 5. Calculated IR (bottom) and VCD (top) spectra of arginine, histidine, and lysine. Labels indicate the selected bands. See Table 6.

results cannot be explained on the basis of the ring current model, because it requires the formation of a closed ring due to hydrogen bonding between the charged carboxylate and amino groups, which only appears in aqueous solutions. Nevertheless, in our calculations all the amino acids are in gas phase, and therefore, it is not possible to get charged carboxylate and amino groups.

Glycine is the simplest amino acid which has a single hydrogen atom as its side chain and it is the only amino acid which has no chirality. As expected, our calculation shows that the VCD of glycine is null. On the other hand, the most significant band is the negative one at 644.7 cm^{-1} for isoleucine, which is assigned to the ρ(OH) rocking mode. The weaker band is a positive−positive doublet at the 726.2 and 774.2 cm^{-1}. They are assigned to the side chain rocking mode. Leucine is an isomer of isoleucine as its name implies. Their most significant bands are one positive at 739.4 cm^{-1} and a negative one at 644.3 cm^{-1}, which are assigned to the δ(COOH) and ρ(OH) bending and rocking modes, respectively. Methionine includes a sulfur atom in its side chain. In Table 4, it is found that the ν_{as}(SCC) asymmetric stretching mode of its side chain has a high intensity positive band at 731.7 cm^{-1}, although some contribution in this frequency range is coming from the δ(COOH) bending mode. The most significant negative band is at 630.4 cm^{-1}, which is assigned to the ρ(OH) rocking mode. For the bands at 1325 and 1426 cm^{-1}, corresponding to vibrations located at the −COOH group and ν_{as}(CCN), Zhang and Polavorapu found them at almost the same positions, in their matrix-assisted amino acid film experiments.[7] For valine, a very significant negative band at 650.4 cm^{-1} is found, which can be assigned to the ρ(OH) rocking mode, whereas a big positive band is seen at 1153 cm^{-1}, corresponding to the δ(COOH) bending mode. The $\nu_{as}(CHCH_3)$ asymmetric stretching mode also contributes to this band.

Three amino acids, phenylalanine, tyrosine, and tryptophan, found in phenylalanine proteins are aromatic (histidine, also being an aromatic amino acid, is discussed below): they have a benzene-like ring structure in their side chain. Their VCD spectra are displayed in Table 5 and Figure 4. For phenylalanine, the most intense band is a negative one at 1821 cm^{-1}, associated with the ν(C═O) stretching mode. A weaker positive band is found at 596.7 cm^{-1}, corresponding to the ρ(OH) rocking mode. Furthermore, the main VCD bands also include a positive−negative−negative triplet at 1085, 1136, and 1199 cm^{-1}, attributed to the modes of phenyl deformation. For tyrosine, the situation is similar to the one found in the case of phenylalanine, the most significant band is a negative one at 1757 cm^{-1}, and a weaker positive one is found at 575 cm^{-1}, associated with the ν(C═O) and ρ(OH) stretching and rocking

TABLE 7: Experimental and Calculated Vibrational Frequencies, and VCD Intensities of Selected Modes of Asparagine, Cysteine, Glutamine, Proline, Serine, and Threonine, See Figure 6

		freq, cm^{-1}			
		exptl	calcd	VCD rot strength × 10^{44} esu^2 cm^2	assignment[a]
asn	1		305.6	−217.76	$\rho(NH_2)$
	2		605.0	−187.51	$\tau(NH_2)$; δ(COOH)
	3		1176	−160.32	ν(CN); δ(COOH)
	4		1207	125.53	ν(CN); δ(COOH)
	5		1793	236.18	ν(C═O)
cys	1	593[b]	623.8	−130.25	ρ(OH)
	2	636[b]	733.2	133.15	δ(COOH)
	3	763[b]	790.9	−57.520	$\rho(CH_2)$
	4	1346[b]	1418	49.575	$\delta_s(NH_3)$
	5	1581[b]	1652	44.829	δ_{as} (NH_3)
gln	1		659.7	−85.4015	ρ(OH)
	2		876.4	92.31	ρ(OH)
	3		1053	−141.602	COOH
	4		1135	94.2665	ν(C═O)
	5		1770	−163.603	ν(C═O)
pro	1		546.5	97.867	ρ(NH)
	2		600.7	−108.91	ρ(OH)
	3		639.1	−154.41	ρ(OH)
	4		773.9	−149.38	ρ(NH); $\rho(CH_2)$
	5		1159	101.02	δ(COOH)
ser	1		558.2	80.683	ρ(OH)
	2		666.5	−87.416	ρ(OH)
	3		819.0	−119.93	δ(CCC)
	4	1128[c]	1161	−90.978	δ(COOH); ρ(OH)
	5	1710[c]	1807	156.31	ν(C═O)
thr	1		619.5	−149.11	δ(CCC)
	2		955.7	85.565	ν(CC)
	3	1128[c]	1159	−135.64	ρ(OH)
	4	1244[c]	1259	116.16	ν(CN)
	5	1710[c]	1804	161.95	ν(C═O)

[a] See the footnote to Table 1. [b] Data from ref 33. [c] Data from ref 36.

modes, respectively. According to their structures, tyrosine is like phenylalanine, but with an extra hydroxyl (−OH) group attached. In our calculated VCD spectra, it is found that the rocking mode ρ(OH) about this hydroxyl group also provides an important contribution at 3816 cm^{-1}. Tryptophan has two ring indole groups added in place of the single aromatic ring found in phenylalanine and tyrosine. Due to this difference, the VCD of tryptophan is also different. The most intense band is associated with the mode of the indole group deformation at 894.9 cm^{-1}. The ν(C═O) stretching mode also makes a large contribution to the VCD spectra at 1350 cm^{-1}. Zhang and Polavorapu found this band at 1342 cm^{-1}, in their matrix-assisted amino acid film experiments.[7]

Arginine, histidine, and lysine can be considered as polar amino acids. Arginine has NH and NH_2 groups in its side chain. The vibrational modes associated to such groups make the main contribution to their VCD spectra. Table 6 and Figure 5 show two intense bands, a negative one at 1622 cm^{-1} associated with $\delta_s(NH_2)$ bending mode and a positive one at 1815 cm^{-1}, associated with ν(C═O) stretching mode. Also, a negative−positive doublet at 641.6 and 653.9 cm^{-1} is seen in its VCD spectra, which can be assigned to the $\rho(NH_2)$ rocking modes. The vibrations of the ρ(OH) rocking modes also generate a large contribution to the arginine VCD spectra. For histidine, there is a negative−positive doublet at 1399 and 1408 cm^{-1}. The flexing behavior of the whole side chain, as well as the vibrations of ρ(OH) rocking modes, might be mainly responsible for such a doublet. Using matrix-assisted amino acid films, Zhang and Polavorapu found a doublet at almost the same positions and with the same sign.[7] In the case of lysine, the vibrational modes of groups on the backbone and side chain have relatively equal contribution to its VCD spectra.

Asparagine, cysteine, glutamine, proline, serine, and threonine are amino acids that possess oxygen, sulfur, and/or nitrogen in the side chain and are therefore polar, but do not carry an overall charge. Their VCD spectra are shown in Table 7 and Figure 6. Asparagine has a NH_2 group in its side chain; from Table 7 it can be seen that the vibration of this group (ρNH_2) makes the most significant negative band at 305.6 cm^{-1}. The most intense positive band is that one associated with the ν(C═O) stretching mode. Although cysteine has a sulfur atom in its side chain, the main contribution still comes from the ρ(OH) and δ(COOH) rocking and bending modes. For glutamine, the most significant contribution to the VCD spectra is from the negative−positive band at 1770 cm^{-1}, related with the ν(C═O) stretching modes. Proline is unique among the amino acids building proteins because its side chain is bonded to the backbone nitrogen as well as to the α-carbon. It is found that the groups NH and CH_2, in this ring, make important contributions to the VCD spectra at 546.5 and 773.9 cm^{-1}; nevertheless, the most significant band is at 639.1 cm^{-1}, associated with the ρ(OH) rocking mode. For serine, the most intense band is a positive one at 1807 cm^{-1}, associated with the ν(C═O) stretching mode. A weaker negative band is found at 819.9 cm^{-1}, associated with the $\rho(NH_2)$ rocking mode. The situation of threonine is similar as with serine, the most significant band is a positive one at 1804 cm^{-1}, associated with the ν(C═O) stretching mode. A weaker negative band is found at 619.5 cm^{-1}, associated with the $\rho(NH_2)$ rocking mode.

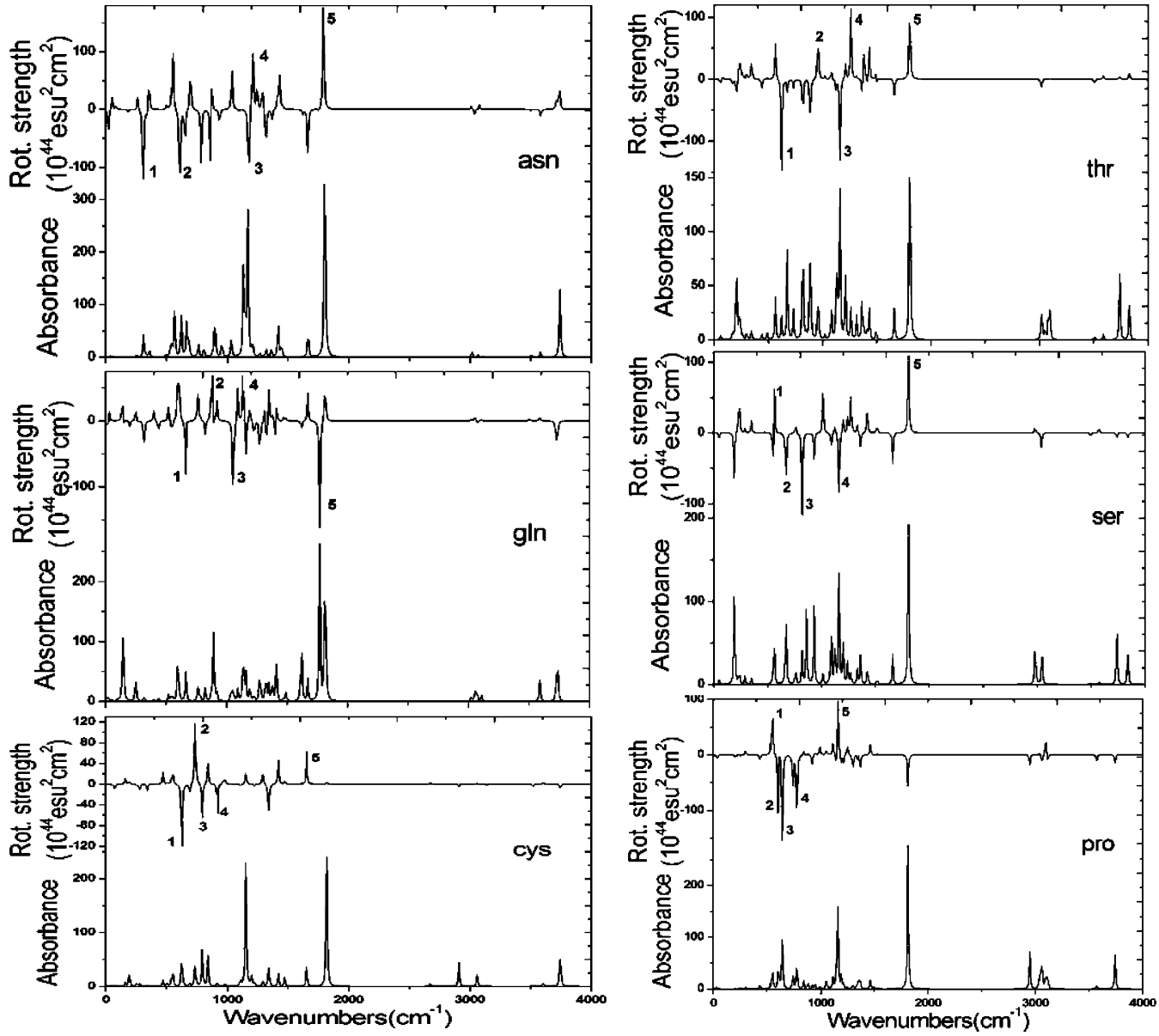

Figure 6. (a) Calculated IR (bottom) and VCD (top) spectra of asparagine, glutamine, and cysteine. Labels indicate the selected bands. See Table 7. (b) Calculated IR (bottom) and VCD (top) spectra of threonine, serine, and proline. Labels indicate the selected bands. See Table 7

TABLE 8: Calculated Vibrational Frequencies and VCD Intensities of the Selected Modes of Aspartic and Glutamic Acids, See Figure [7]

		cald freq, cm^{-1}	VCD rot strength × 10^{44} esu^2 cm^2	assignment[a]
asp	1	618.3	−204.29	ρ(OH)
	2	662.8	159.57	ρ(OH)
	3	1170	235.25	δ(COOH)
	4	1803	−496.06	ν(C═O)
	5	1808	472.99	ν(C═O)
glu	1	572.0	149.62	ρ(OH)
	2	613.2	−169.42	ρ(OH)
	3	739.1	174.93	δ(COOH)
	4	1818	−505.09	ν(C═O)
	5	1820	437.68	ν(C═O)

[a] See the footnote to Table 1.

Finally, the VCD spectra of the aspartic and glutamic acids are shown in Table 8 and Figure 7. They all have a carboxyl group in their side chain. It is found that carboxyl groups in backbones and side chains provide the main contribution to the VCD spectra in both cases. Single or negative−positive bands associated with the vibrational modes on these carboxyl groups can be seen in their VCD spectra.

From the previous analysis of the 20 calculated VCD spectra, a general emerging trend indicates that the ρ(OH) rocking and ν(C═O) stretching modes generate the highest VCD intensities in most amino acids. On the other hand, the δ(COOH) bending mode is also very intense for nonpolar amino acids, whereas the ρ(NH_2) rocking mode has relatively high intensity for polar amino acids. It is a task for future studies to gain insights into the origin of these trends.

In order to study changes in the IR and VCD spectra between charged gas phase species and neutral ones, similar computations on charged species of, for example, Asp, Glu, and His have been carried out. The charged species were built from the corresponding neutral species. The Asp and Glu charged amino acids were created by subtracting the hydrogen atom of the hydroxyl bond in the R group and assigning a net charge of −1 to the compound. In the case of the His amino acid, the charged compound was created by adding a hydrogen atom to

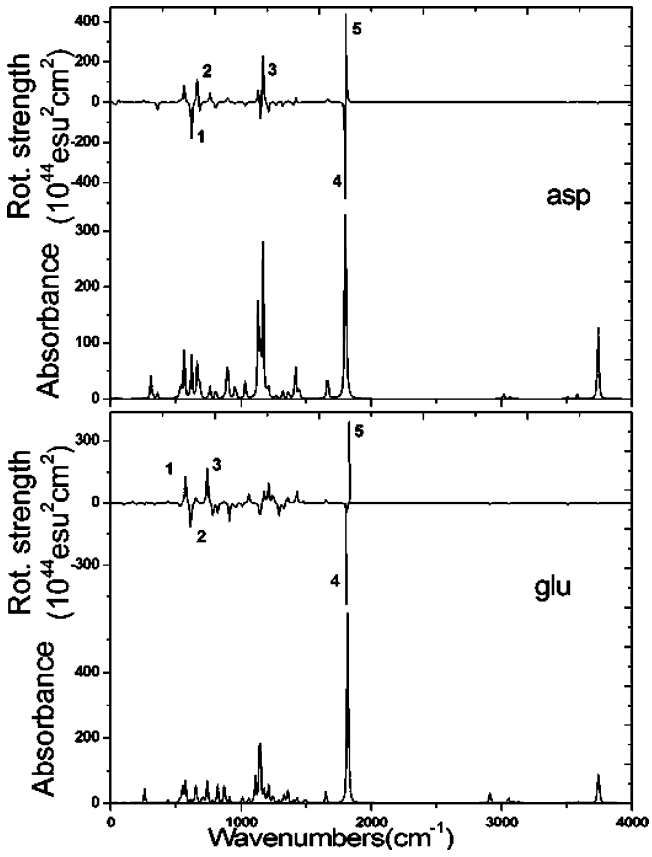

Figure 7. Calculated IR (bottom) and VCD (top) spectra of aspartic and glutamic acids. Labels indicate the selected bands. See Table 8.

the nitrogen of the five-membered ring of the R group but assigning a net charge of +1 to the compound. The results show that the charged amino acids exhibit substantial structural modifications with respect to the conformations of the neutral amino acids and with consequent changes in the IR and VCD spectra. The large number of vibrational modes in the low-, intermediate-, and high-frequency intervals makes a difficult task to establish frequency shifts with respect to the spectra of the neutral species. The complete calculated VCD spectra for the neutral 20 α-amino acids can be obtained from the Supporting Information associated to this work.

4. Summary

By using density functional theory at the B3LYP/cc-pVTZ level of theory, it was possible to determine the vibrational spectra of the 20 α-amino acids in their non-zwitterionic form. The agreement of different structural properties values with experimental measurements indicates that the level of theory used in this work is satisfactory to analyze the amino acids vibrations. In general, it is observed that molecular torsions are the characteristic vibrations of the amino acids at the low-frequency regime (0, 500 cm^{-1}). At the intermediate range (500, 1600) cm^{-1}, the bending of bond angles, the out-of-plane wagging of individual hydrogen atoms, and the stretching of NC, CC, and CO bonds dominate this part of the spectra. However, for the relatively similar vibration modes of the amino acids, the specific molecular structure of the R groups strongly characterizes the frequencies in both the low and intermediate frequency intervals. Therefore, such intervals are important for the identification of individual amino acids. On the other hand, in the high-frequency regime, above 1600 cm^{-1}, several common vibrations of the amino acids are observed, for instance, NH_2 scissoring, CO bond stretching, and the (symmetric and asymmetric) stretching of the hydrogen atoms in the NH_2 and OH groups. As a consequence, this last frequency interval is appropriate for the general identification of the amino acids and to distinguish them from other biological compounds.

In the present study, it was also found that the groups in the backbones and side chains make different contribution to the VCD spectra for the 20 amino acids. Nevertheless, some general trends emerging from the calculations indicate that specific modes, such as the ρ(OH) rocking and ν(C═O) stretching ones, are the most intense for most amino acids. It is expected that the present calculations will motivate experimental research which allow, in the near future, performing VCD spectra measurements of molecules in gas phase.

Acknowledgment. We thank DGSCA-UNAM Supercomputer Center for the computing resources used in this work. I.L.G. is grateful for support from DGAPA-UNAM under Project IN112808 and Conacyt-Mexico under Project 80610.

Supporting Information Available: Optimized structures and calculated vibrational frequencies, infrared intensities, and VCD spectra of all amino acids. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) *Infrared and Raman Spectroscopy of Biological Materials*; Gremlich, H.-U., Yan, B., Eds.; Marcel Dekker, Inc.: New York, 2001.
(2) Bakker, J. M.; Aleese, L. M.; Meijer, G.; von Helden, G. *Phys. Rev. Lett.* **2003**, *91*, 203003.
(3) Compagnon, I.; Oomens, J.; Meijer, G.; von Helden, G. *J. Am. Chem. Soc.* **2006**, *128*, 3592.
(4) Oomens, J.; Polfer, N.; van der Meer, L.; Marshall, A. G.; Eyler, J. R.; Meijer, G.; von Helden, G. *Phys. Chem. Chem. Phys.* **2005**, *7*, 1345.
(5) von Helden, G.; Compagnon, I.; Blom, M. N.; Frankowski, M.; Erlekam, U.; Oomens, J.; Brauer, B.; Gerber, R. B.; Meijer, G. *Phys. Chem. Chem. Phys.* **2008**, *10*, 1248.
(6) Zuk, W. M.; Freedman, T. B.; Nafie, L. A. *J. Phys. Chem.* **1989**, *93*, 1771. references therein.
(7) Zhang, P.; Polavarapu, P. L. *Appl. Spectrosc.* **2006**, *60*, 378.
(8) Tajkhorshid, E.; Jalkanen, K. J.; Suhai, S. *J. Phys. Chem. B* **1998**, *102*, 5899.
(9) Tanaka, T.; Kodama, T.; Morita, H. E.; Ohno, T. *Chirality* **2006**, *18*, 652.
(10) Nafie, L. A.; Freddman T. B. In *Infrared and Raman Spectroscopy of Biological Materials*; Gremlich, H.-U., Yan, B., Eds.; Marcel Dekker, Inc.: New York, 2001; pp 15−54.
(11) Diem, M.; Polavarapu, P. L.; Oboodi, M. R.; Nafie, L. A. *J. Am. Chem. Soc.* **1982**, *104*, 3329.
(12) Nafie, L. A.; Oboodi, M. R.; Freedman, T. B. *J. Am. Chem. Soc.* **1983**, *105*, 7449.
(13) Freedman, T. B.; Chernovitz, A. C.; Zuk, W. M.; Paterlini, M. G.; Nafie, L. A. *J. Am. Chem. Soc.* **1988**, *110*, 6970.
(14) Diem, M. *J. Am. Chem. Soc.* **1988**, *110*, 6967.
(15) Urbanová, M.; Maloň, P. In *Analytical Methods in Supramolecular Chemistry*; Schalley, C. A., Ed.; Wiley VCH: Weinheim, 2007; pp 265−304.
(16) Snoek, L. C.; Robertson, E. G.; Kroemer, R. T.; Simons, J. P. *Chem. Phys. Lett.* **2000**, *321*, 49.
(17) Snoek, L. C.; Kroemer, R. T.; Hockridge, M. R.; Simons, J. P. *Phys. Chem. Chem. Phys.* **2001**, *3*, 1819.
(18) Kamariotis, A.; Boyarkin, O. V.; Mercier, S. R.; Beck, R. D.; Bush, M. F.; Williams, E. R.; Rizzo, T. R. *J. Am. Chem. Soc.* **2006**, *128*, 905.
(19) Oomens, J.; Steill, J. D.; Redlich, B. *J. Am. Chem. Soc.* **2009**, *131*, 4310.
(20) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich,

S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03, Revision B.05*; Gaussian, Inc.: Pittsburgh, PA, 2003.

(21) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648. Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785. Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys* **1980**, *58*, 1200.

(22) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623.

(23) (a) Stepanian, S. G.; Reva, I. D.; Radchenko, E. D.; Rosado, M. T. S.; Duarte, M. L. T. S.; Fausto, R.; Adamowicz, L. *J. Phys. Chem. A* **1998**, *102*, 1041. (b) Matei, A.; Drichko, N.; Gompf, B.; Dressel, M. *Chem. Phys.* **2005**, *316*, 61.

(24) Stepanian, S. G.; Reva, I. D.; Radchenko, E. D.; Adamowicz, L. *J. Phys. Chem. A* **1998**, *102*, 4623.

(25) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007.

(26) The initial coordinates for the amino acids geometry optimization were obtained from http://chemistry.gsu.edu/Glactone/PDB/pdb.html.

(27) Stephens, P. J. *J. Phys. Chem.* **1985**, *89*, 748.

(28) Cheeseman, J. R.; Frisch, M. J.; Devlin, F. J.; Stephens, P. J. *Chem. Phys. Lett.* **1996**, *252*, 211.

(29) Görbitz, C. H.; Dalhus, B. *Acta Crystallogr., Sect. C* **1996**, *52*, 1756.

(30) Santamaria, R.; Vázquez, A. *J. Comput. Chem.* **1994**, *15*, 981.

(31) Santamaria, R.; Charro, E.; Zacarias, A.; Castro, M. *J. Comput. Chem.* **1999**, *20*, 511.

(32) Rosas-Acevedo, H. Private communication.

(33) Pawlukojć, A.; Leciejewicz, J.; Ramirez-Cuesta, A. J.; Nowicka-Scheibe, J. *Spectrochim. Acta, Part A* **2005**, *61*, 2474.

(34) Kumar, S.; Rai, A. K.; Singh, V. B.; Rai, S. B. *Spectrochim. Acta, Part A* **2005**, *61*, 2741.

(35) Tulip, P. R.; Clark, S. J. *J. Chem. Phys.* **2004**, *121*, 5201.

(36) Lakard, B. *J. Mol. Struct.: THEOCHEM* **2004**, *681*, 183.

JP9108442