

CHIROPTICAL PROPERTIES OF *S*-ALANINE CONFORMATIONAL ISOMERS¹

J. WEBB, R. W. STRICKLAND and F. S. RICHARDSON*

Department of Chemistry, University of Virginia, Charlottesville, Virginia 22901

(Received USA 24 October 1972; Received in UK for publication 29 March 1973)

Abstract—The chiroptical properties of several *S*-alanine (zwitterionic and nonzwitterionic), *S*-alanine cation, and *S*-alanine anion structural isomers are calculated from LCAO-MO-SCF-CI wave functions constructed on the semiempirical CNDO MO model. Electronic excited states are constructed in the virtual orbital-configuration interaction approximation. We calculate and report the transition energies, dipole strengths, reduced rotatory strengths, and dissymmetry factors for the four lowest energy singlet-singlet transitions as a function of a single conformational variable (the angle between the $C_{\alpha}COO'$ plane and the $CC_{\alpha}N$ plane, where C \equiv carboxylate C atom, N \equiv ammonium or amino N atom, and C_{α} \equiv α -carbon atom) for zwitterionic and nonzwitterionic *S*-alanine, the *S*-alanine anion, and the *S*-alanine cation. The ground state dipole moments and first three ionization potentials are also computed and reported for the various *S*-alanine species. Electric dipole transition integrals are calculated in the dipole velocity formalism and all two-center contributions to these and the magnetic dipole transition integrals are included. The calculated rotatory strengths are compared with experimental circular dichroism spectra reported for *S*-alanine and with predictions based on sector rules which have been proposed for α -amino acids.

1. INTRODUCTION

Natural optical activity has proved to be of enormous value in the study of both the static and the dynamical aspects of protein structure.² To be sure, much of the early optimism that was expressed concerning the reliability and general applicability of ORD and CD for elucidating the details of secondary structure (e.g., percent of alpha helical structure versus random coil) has been tempered by the more recent structural information made available from X-ray diffraction data and NMR spectra.^{2a,i} However, it remains clear that ORD and CD are exceedingly useful tools for probing protein structure and for following changes in structure.

Over the past 15 years several powerful and elegant quantum theoretical models have been developed for the purpose of quantitatively or semi-quantitatively relating the experimental observables of ORD and CD to the underlying spectroscopic processes and, ultimately, to specific stereochemical features in complex molecular systems.³ For homopolymeric systems, the input parameters for each of these models are the structure variables, spectroscopic properties and, in some instances, the ground state electronic charge distributions of the monomer units. In the language of perturbation theory, it is assumed to zeroth-order that the spectroscopic properties and structure of the monomeric units are known. Protein molecules, of course, are not homopolymers in the chemical

sense, but in the spectral region, $230 \text{ nm} > \lambda > 150 \text{ nm}$, their optical properties are most likely dominated by the chromophoric behavior of the repeating amide groups in the backbone.^{4,5} Assuming then that the spectroscopic contributions of side-chain moieties (such as the indole, phenolic, phenyl, sulfhydryl, disulfide, and imidazole groups) can be treated separately, the usual procedure is to consider a protein, optically, as a homopolymer comprised of amide monomer units. For this reason, complete and accurate characterization of the spectroscopic properties of amino acids and model amide systems is essential for optimal and reliable application of theoretical models to the interpretation of protein spectra.

The spectroscopic properties of model amides have been thoroughly examined both experimentally^{6,7a} and theoretically.^{6b,7} Although a characterization of these properties in terms of electronic structural details is far from complete, the present state-of-knowledge allows for useful spectra-structure correlations. The spectroscopic properties of amino acids have also been subject of numerous experimental studies, but their chiroptical and electronic absorption properties are not nearly as well-characterized as are those of simple amide systems.

In the present paper we examine, on a semi-empirical MO model, the chiroptical properties of the simplest optically active aliphatic amino acid, *S*-alanine. The model and calculational methods employed in this work are similar to those used previously by us in studies on chiral cyclopentanones,⁸ cyclic amides and diamides,^{7c} and

*Fellow of The Camille and Henry Dreyfus Foundation.

disulfide systems.⁹ The only previous calculation of the optical rotatory properties of *S*-alanine reported in the literature was carried out by Vol'kenshtein and Kruchek.¹⁰ Their calculation was based on Kirkwood's polarizability model and they considered just one structural isomer of the zwitterionic form (the carboxylate group and the ammonium N atom coplanar), and two structural isomers of the cation, *S*-alanine-H⁺ (CO'OH or COO'H coplanar with the ammonium N atom).

Experimental investigations of the ORD and CD spectra of aliphatic amino acids are quite numerous.¹¹ Until recently it was generally believed that all naturally occurring, and most synthetic, aliphatic *S*- α -amino acids exhibit an initial (lowest energy) positive Cotton effect with the first extremum at 223–227 nm in ORD spectra and at 208–211 nm in CD spectra. However, more recent CD measurements have revealed an additional Cotton effect in the near UV located at 235–250 nm.^{13–15} For *S*-alanine in 50% ethanol at pH 11, Craig and Pereira¹⁴ reported the following data on the first two Cotton effects, $\lambda(\text{max}) = 235$ nm, $[\theta]_{\text{max}} = -28$; $\lambda(\text{max}) = 211$ nm, $[\theta]_{\text{max}} = 1261$. These same authors reported no CD in the 235–250 nm region for *S*-alanine in ethanol at pH 1; the first CD band appears at $\lambda(\text{max}) = 209$ nm with a molar ellipticity of $[\theta]_{\text{max}} = 1277$. For the methyl ester of *S*-alanine in 95% ethanol, the first three CD bands were reported as follows: $\lambda(\text{max}) = 236$ nm, $[\theta]_{\text{max}} = -395$; $\lambda(\text{max}) = 209$ nm, $[\theta]_{\text{max}} = 3227$; $\lambda(\text{max}) = 199$ nm, $[\theta]_{\text{max}} = 3564$. The methyl ester hydrochloride of *S*-alanine, however, shows only one CD band at wavelengths >200 nm; this band was reported at $\lambda(\text{max}) = 208$ nm with $[\theta]_{\text{max}} = 2987$.¹⁴ Toniolo¹⁵ reported a negative CD band at ~ 240 nm ($\Delta\epsilon = -0.0031$) for *S*-alanine in alkaline solution (pH 13, 0.1 N NaOH), a very small negative CD band at ~ 250 nm ($\Delta\epsilon = -0.0001$) for *S*-alanine in acidic solution (pH 1, 0.1 N HCl) and no CD bands at wavelengths >215 nm for *S*-alanine in buffered solutions (phosphate buffer) at pH 7.

In summary, *S*-alanine in the zwitterionic form gives only one CD band for $\lambda > 200$ nm; this band is positive and is centered at about 204 nm ($\Delta\epsilon = 0.68$).^{12f} In alkaline solution, two bands are observed for $\lambda > 200$ nm; these bands are located at ~ 240 nm ($\Delta\epsilon = 0.0031$)¹⁵ and at ~ 214 nm ($\Delta\epsilon = 0.33$).^{12f} In acidic solution, at least one positive band is present at ~ 208 nm ($\Delta\epsilon = 1.04$),^{12f} and possibly a negative band is located at ~ 250 nm ($\Delta\epsilon = -0.0001$).¹⁵ The presence of a long wavelength negative band in acid media is, however, not certain.

The simplest optically active α -hydroxy acid, *S*(*R*)-lactic acid, also exhibits two CD bands at wavelengths >200 nm under certain solvent conditions.^{12e, 14–16} For *S*-lactic acid in 95% ethanol, Craig and Pereira¹⁴ reported the following data:

$\lambda(\text{max}) = 244$ nm, $[\theta]_{\text{max}} = -14.6$; $\lambda(\text{max}) = 210$ nm, $[\theta]_{\text{max}} = 2727$. For *S*-lactic acid in water at pH 1, these same authors reported: $\lambda(\text{max}) = 246.5$ nm, $[\theta]_{\text{max}} = -17.4$; $\lambda(\text{max}) = 212$ nm, $[\theta]_{\text{max}} = 2157$. They found only one CD band at $\lambda > 200$ nm for *S*-lactic acid in water at pH 9; this band was located at 214 nm with a molar ellipticity of 744. Anand and Hargreaves^{16a} assigned the weak, lower energy CD band in lactic acid to an $n \rightarrow \pi^*$ carboxylate transition, and the more intense, higher energy band (~ 210 – 215 nm) was ascribed to a $\pi \rightarrow \pi^*$ carboxylate transition. These assignments are in direct conflict with the more conventional view that the Cotton effect associated with the 210 nm absorption band of carboxylic acids and esters is of $n \rightarrow \pi^*$ origin.^{17, 18} Barth *et al.*,^{16b} studied the temperature dependence of the CD spectra of several lactic acid derivatives, and concluded that both CD bands at $\lambda > 200$ nm could be assigned to the carboxylate $n \rightarrow \pi^*$ transition. They attributed the presence of two bands either to specific solvational effects or to conformational equilibria between conformers with different chiroptical properties.

The problem of CD band assignments at $\lambda > 200$ nm for aliphatic α -amino acids also remains unsolved. Is the presence of two CD bands diagnostic of two different electronic excitations or of two conformational isomers which are relatively stable under certain solvent and temperature conditions? It has been proposed that analysis of NMR spectra can provide information on rotamer populations involving rotation about the C_α – C_β bond of various amino acids.¹⁹ However, this method is not suitable for assessing the predominant conformations achieved by rotation about the $\text{OOC}-C_\alpha$ bond. Crystal structures of several amino acids are now available. However, for these compounds it is unlikely that the molecular conformations assumed in the crystalline environment are fully preserved in solution. It would appear, therefore, that accurate and reliable CD spectra-structure relationships applicable to amino acids in solution must be obtained solely from thorough empirical studies and theoretical examination of the chiroptical properties of these systems. Other spectroscopic and nonspectroscopic methods are of little help in this case.

Recently, Johnson *et al.* have completed a study in which they measured the absorption and CD spectra of several alkyl amino acids to 160 nm and calculated the rotatory strengths of these systems as a function of the angle $\text{OCC}_\alpha|\text{CC}_\alpha\text{N}$ (where N \equiv the ammonium group N atom, $C_\alpha \equiv$ the α C atom, C \equiv the carboxylate C atom, and O \equiv carboxylate O atom).²⁰ Their spectral measurements were made on samples in which the amino acids were dissolved in hexafluoroisopropanol (HFIP). All the amino acids studied had a $\text{p}K_1$ of 2.4 or lower, and it was assumed that each existed

predominately in the zwitterionic form. Alanine, valine, isoleucine, and leucine each gave a single positive CD band between 180 and 160 nm (centered around 170 nm). Proline exhibited a somewhat more complex CD spectrum with a weak positive band at 210 nm, a weak negative band around 190 nm, and a stronger negative band centered around 165 nm. For *S*-alanine, the band maxima have the following characteristics: $\lambda(\text{max}) = 196 \text{ nm}$, $\Delta\epsilon(\text{max}) = 1.00$; $\lambda(\text{max}) = 168 \text{ nm}$, $\Delta\epsilon(\text{max}) = -1.00$. The absorption spectrum of alanine exhibits a single band between 220 and 160 nm, with $\lambda(\text{max}) = 165 \text{ nm}$ and $\epsilon = 7800$.

Johnson employed the independent systems model of Tinoco²¹ for calculating the rotatory strengths of the two lowest-lying transitions in the amino acid molecules as functions of the $\text{OCC}_\alpha\text{CC}_\alpha\text{N}$ angle. For each amino acid, the molecule was partitioned into a chromophoric group (the carboxylate anion) and perturbing groups (all chemical bonds external to the carboxylate group), and it was assumed that electron exchange between the chromophoric group and the perturbing groups is negligible. On this model the spectroscopic states responsible for the spectral properties observed at $\lambda > 160 \text{ nm}$ are constructed from electron configurations "localized" on the carboxylate group. The influence of the extrachromophoric environment on these states is introduced via non-exchange perturbative mechanisms. Johnson assumed that only three carboxylate excited states were required to represent the CD spectrum at $\lambda > 160 \text{ nm}$. Two of these states are expected to be nearly degenerate and can be characterized in terms of $n\pi^*$ configurations. The third state can be characterized best in terms of a $\pi\pi^*$ configuration. According to Johnson, transitions to the $n\pi^*$ states account for the lower energy ($\lambda(\text{max}) \sim 196 \text{ nm}$) CD band in *S*-alanine, and the higher energy ($\lambda(\text{max}) \sim 168 \text{ nm}$) CD band is assigned to a transition from the ground state to the $\pi\pi^*$ excited state. Furthermore, Johnson concludes that the angle $\text{OCC}_\alpha\text{CC}_\alpha\text{N}$ (henceforth referred to as ϕ) must have a value between 0° and -70° (see Fig 1 for definition and sign convention for ϕ) for the *S*-alanine zwitterion in HFIP. In crystalline *S*-alanine, $\phi = -19^\circ$.

2. METHODS OF CALCULATION

The methods of calculation employed in the present study are essentially the same as those described in two previous reports from this laboratory.^{8,9} The LCAO-MO-SCF calculations were carried out in the CNDO approximation utilizing "standard" parameters.²² Electronic excited states were constructed in the virtual orbital-configuration interaction approximation. That is, excited states are represented as linear combinations of singly excited determinantal wave functions which are obtained by replacing an orbital in the ground

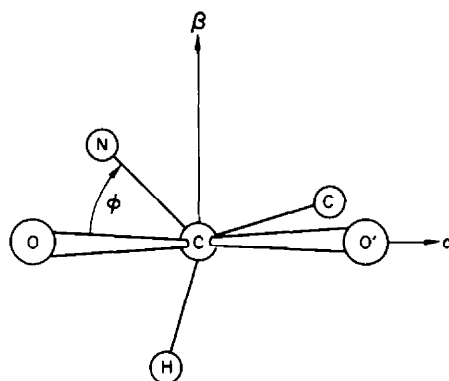


Fig 1. Definition of the angle $\phi(\text{OCC}_\alpha/\text{CC}_\alpha\text{N})$. View in the $\text{C}_\alpha\text{—OO}'$ plane.

state determinant with one of the virtual orbitals generated in the SCF calculation. In the present study only singlet excited states were calculated.

The spectroscopic quantities that we calculate and report for excitations from the ground state to singlet electronic excited states are transition energies, oscillator strengths, reduced rotatory strengths, and dissymmetry factors. For a transition, $0 \rightarrow a$, these quantities are defined as follows:

(a) oscillator strength,

$$f_{0a} = (4\pi m \nu_{0a} / 3\hbar e^2) |\langle \psi_0 | \hat{\mu} | \psi_a \rangle|^2 = (4\pi m \nu_{0a} / 3\hbar e^2) D_{0a} \quad (1)$$

(b) reduced rotatory strength,

$$[R_{0a}] = (100/\beta \mathcal{D}) \text{Im}[\langle \psi_0 | \hat{\mu} | \psi_a \rangle \cdot \langle \psi_a | \hat{m} | \psi_0 \rangle] = (100/\beta \mathcal{D}) R_{0a} \quad (2)$$

(c) dissymmetry factor,

$$g_{0a} = 4 \text{Im}[\langle \psi_0 | \hat{\mu} | \psi_a \rangle \cdot \langle \psi_a | \hat{m} | \psi_0 \rangle] / |\psi_0 | \hat{\mu} | \psi_a \rangle|^2 = 4R_{0a}/D_{0a} \quad (3)$$

where $\beta = \text{Bohr magneton} = (eh/2mc)$, $\mathcal{D} = \text{Debye unit} = 10^{-18} \text{ esu-cm}$, $\nu_{0a} = \text{transition frequency} = (E_a - E_0)/h$, $\hat{\mu}$ is the electric dipole operator, \hat{m} is the magnetic dipole operator, D_{0a} is the dipole strength, and R_{0a} is the rotatory strength.

Prior to calculating the electric dipole and magnetic dipole transition integrals, the CNDO wave functions are renormalized including overlap ("deorthogonalized").^{8,23} Using these renormalized wave functions, both the one-center and two-center terms in the transitions dipole matrices are calculated and both one-center and two-center contributions are included in the computed rotatory strengths. The electric dipole transition integrals are calculated in the dipole velocity formalism and are then transformed into the dipole

length representation according to,

$$\langle \psi_i | \hat{\mu} | \psi_j \rangle = (e\hbar^2/m) \langle \psi_i | \hat{\nabla} | \psi_j \rangle (E_j - E_i)^{-1}, \quad (4)$$

where

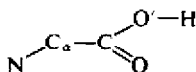
$$\hat{\mu} = e \sum_{\epsilon} \hat{r}_{\epsilon}, \quad \hat{\nabla} = \sum_{\epsilon} \left(i \frac{\partial}{\partial x_{\epsilon}} + j \frac{\partial}{\partial y_{\epsilon}} + k \frac{\partial}{\partial z_{\epsilon}} \right),$$

and \sum_{ϵ} is taken over all electrons ϵ .

3. STRUCTURE PARAMETERS

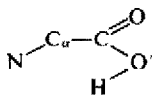
Four structural isomers of the *S*-alanine zwitterion were included in the present study. The atomic coordinates for one of these isomers (structure 1) were taken directly from the X-ray diffraction data reported by Simpson and Marsh²⁴ for crystalline *S*-alanine. In this structure, the $C_{\alpha}COO'$ group is planar and the angle $OCC_{\alpha}/CC_{\alpha}N$ is -19° (see Fig 1 for the definition of this angle, henceforth called ϕ). Structures 2, 3, and 4 are also zwitterionic forms of *S*-alanine and are generated by rotations about the $C-C_{\alpha}$ bond axis, holding all bond distances and other structure parameters fixed. For structure 2, $\phi = -60^{\circ}$; for structure 3, $\phi = 30^{\circ}$; for structure 4, $\phi = 0^{\circ}$.

Structure 5 is an anionic form of *S*-alanine ($CH_3CH(NH_2)COO^-$) in which $\phi = 0^{\circ}$ (i.e., the amino N atom is coplanar with $C_{\alpha}COO'$). Structure 6 is a cationic form of *S*-alanine ($CH_3CH(NH_3^+)COOH$) in which $\phi = 0^{\circ}$ and the carboxylate proton is cis to the carbonyl oxygen. That is, the atoms of the $NC_{\alpha}COO'H$ group are coplanar and are arranged as follows:



Furthermore, one of the ammonium group hydrogens is in this plane and is directed away from the CO group.

Structures 7-10 are nonzwitterionic neutral species of *S*-alanine ($CH_3CH(NH_2)COOH$). In structure 7, $\phi = 0^{\circ}$ and the carbonyl group is trans to the amino group. The $NC_{\alpha}COO'H$ group is planar and has the geometry,



where $\angle CO'H = 120^{\circ}$. The amino group hydrogens are symmetrically disposed above and below the plane of $NC_{\alpha}COO'H$ and are directed away from the carboxylate group (the lone pair orbital of the sp^3 hybridized N atom lies in the plane and is directed towards the carboxylate proton). Structures 8-10 are generated from 7 by rotations about

the $C-C_{\alpha}$ bond axis. The values of ϕ for each of these structures are: structure 8, $\phi = -19^{\circ}$; structure 9, $\phi = 60^{\circ}$; structure 10, $\phi = 90^{\circ}$.

The ammonium and Me groups are isoelectronic and, on this basis, structures 2 and 3 might be considered the most symmetrical $C-C_{\alpha}$ rotamers of the *S*-alanine zwitterion. Assuming the groups CH_3 and NH_3^+ to be identical, structure 2 has C_s point-group symmetry with a mirror plane which includes the COO' group. Making the same assumption for structure 3, this isomer also has C_s point-group symmetry, but in this case the mirror plane is perpendicular to and bisects the COO' group. The chiroptical properties of structures 2 and 3, then, arise solely from dissimilarities between the Me and ammonium groups.

4. RESULTS

The optical properties calculated for the first four spin-allowed transitions of structures 1-10 are presented in Table 1. In Table 2 are listed the ground state dipole moment and binding energy calculated for each structure. The energies computed for the three highest occupied molecular orbitals (HOMO's) are displayed in Table 3. The orbitals designated (\perp) in Table 3 are essentially antisymmetric combinations of oxygen p functions oriented perpendicular to the $-COO'$ plane. In the ensuing discussion these approximately non-bonding molecular orbitals are referred to as π_o . In the zwitterionic (1-4) and anionic (5) structures, the orbitals designated (\parallel) are either symmetric or antisymmetric combinations of oxygen p functions which lie in the $-COO'$ plane but which are oriented perpendicular to the $C-O$ ($C-O'$) bond axes. These molecular orbitals are referred to, henceforth, as either n (symmetric combination) or n' (antisymmetric combination) functions. In the cationic (6) and nonzwitterionic neutral (7-10) structures the two carboxylate O atoms are, of course, nonequivalent. In these structures, the n' orbital is localized almost entirely on the carbonyl oxygen and the n orbital includes substantial oxygen $2s$ and hydrogen $1s$ character. The orbitals designated as (N) in Table 3 can be characterized as lone pair functions (sp^3 hybrids) on the amino group N atom.

Given the low symmetry of all the structures studied and the extensive mixing of singly-excited configurations in our CI calculations, it is difficult to characterize the transitions listed in Table 1 in terms of one-electron excitations between specific pairs of SCF and virtual orbitals. The lowest energy transition calculated for each zwitterion structure (1-4) can be described best as a charge transfer transition in which an electron is transferred from the n' orbital (which is localized on the carboxylate oxygens) to an ammonium group orbital. The second lowest energy transition in each of the zwitterion structures can be described

Table 1. Calculated properties of four lowest energy singlet-singlet transitions

Structure ^a	ϕ	ΔE (eV)	D (debye ²)	$[R]$	$ g $
1 (Z)	-19°	7.00	2.11	31.4	0.006
		8.02	1.29	-38.2	0.012
		8.17	0.80	-3.6	0.002
		8.70	4.38	36.9	0.003
2 (Z)	-60°	6.33	1.41	36.8	0.011
		8.04	1.30	-17.0	0.005
		8.24	0.48	29.9	0.025
		8.80	0.20	-1.8	0.005
3 (Z)	30°	6.85	2.14	-27.6	0.005
		7.85	1.62	34.7	0.010
		8.15	0.80	9.0	0.004
		8.86	1.75	17.7	0.004
4 (Z)	0°	6.87	1.96	22.1	0.005
		7.77	1.20	-22.8	0.008
		8.14	0.82	-3.7	0.002
		8.22	2.50	16.9	0.003
5 (A)	0°	7.14	1.33	-24.4	0.007
		8.15	0.20	26.0	0.072
		10.41	1.38	23.2	0.010
		10.82	1.01	-41.0	0.016
6 (C)	0°	6.29	0.67	17.5	0.011
		8.90	0.35	-54.5	0.063
		10.72	5.22	36.0	0.003
		11.55	3.47	-22.2	0.003
7 (NZ)	0°	5.99	1.52	-23.7	0.006
		10.35	0.14	16.7	0.048
		11.49	1.50	-4.7	0.001
		11.85	0.94	11.8	0.006
		6.01	1.54	-25.1	0.006
		10.29	0.50	-2.9	0.003
		11.59	1.24	-24.4	0.008
		11.86	0.25	16.9	0.027
9 (NZ)	60°	5.88	1.56	-13.5	0.004
		10.31	0.72	92.6	0.052
		11.46	1.09	42.3	0.016
10 (NZ)	90°	11.78	0.59	-30.4	0.021
		5.82	1.33	-6.4	0.002
		10.06	0.29	67.4	0.093
		11.58	0.66	38.2	0.023
		11.76	0.69	-11.0	0.007

^aZ \equiv zwitterionic species; A \equiv anion; C \equiv cation; NZ \equiv nonzwitterionic neutral species.

approximately as an $n' \rightarrow \pi^*$ excitation localized on the carboxylate group. The third and fourth transitions calculated for structures 1-4 have predominantly $n \rightarrow \pi^*$ and $\pi_0 \rightarrow$ ammonium group character.

The lowest and second lowest energy transition calculated for structures 7-10 can be characterized as $n' \rightarrow \pi^*$ and $n \rightarrow \pi^*$ carboxylate transitions, respectively. The third and fourth transitions for these structures have predominantly $\pi_0 \rightarrow \pi^*$ and $n' \rightarrow \sigma_{OH}^*$ character, where σ_{OH}^* is an antibonding sigma orbital on the -OH group.

In the anionic structure 5, the first two transitions are best represented as carboxylate $n' \rightarrow \pi^*$ and

$n \rightarrow \pi^*$ excitations according to our calculations, and the third and fourth transitions have predominantly $\pi_0 \rightarrow \pi^*$ and $n' \rightarrow \sigma_{COO}^*$ character. The first four transitions of the cationic structure 6 may be described in the same way as the analogous transitions in the zwitterionic structures 1-4.

5. DISCUSSION

By Koopman's theorem,²⁵ the first three ionization potentials of each structure may be set equal to the negatives of the orbital energies listed in Table 3. Since the orbital energies given in Table 3 are relatively insensitive to the angle ϕ , we may

Table 2. Calculated dipole moments and binding energies

Structure ^a	ϕ	Dipole moment ^b	Binding energy ^c
1 (Z)	-19°	12.92	-4.513
2 (Z)	-60°	12.86	-4.506
3 (Z)	30°	12.89	-4.510
4 (Z)	0°	13.21	-4.514
5 (A)	0°	1.41	-4.504
6 (C)	0°	10.76	-4.484
7 (NZ)	0°	6.62	-4.669
8 (NZ)	-19°	6.52	-4.668
9 (NZ)	60°	6.85	-4.660
10 (NZ)	90°	7.18	-4.658

^aZ ≡ zwitterionic species; A ≡ anion; C ≡ cation; NZ ≡ nonzwitterionic neutral species.

^bIn Debye units (10^{-18} cgs).

^cIn atomic units.

list the calculated ionization potentials (in eV) as follows:

	IP_1	IP_2	IP_3
Z	~9.2	~9.7	~10.6
A	3.4	3.5	4.7
C	19.3	19.8	21.7
NZ	~12.5	~13.8	15.3

for the zwitterionic (Z), anionic (A), cationic (C), and nonzwitterionic neutral (NZ) structures, respectively. Sweigart and Turner²⁶ recently have reported the photoelectron spectra of several carboxylic acids and their derivatives. In formic acid, acetic acid, and the esters of these com-

pounds, the highest occupied MO was identified as a carbonyl oxygen lone-pair orbital (analogous to the n' orbital in our calculations) and the second highest occupied orbital was assigned as an out-of-plane nonbonding function analogous to our π_0 molecular orbital. These assignments correspond directly to our calculated results for the NZ structures of *S*-alanine. For acetic acid, Sweigart and Turner reported the first ionization potential (vertical) to be 10.87 eV, and the second ionization potential (vertical) to be 12.05 eV. For trifluoroacetic acid the first two ionization potentials were reported as 11.77 eV and 12.92 eV. These experimental values are similar in magnitude to those we calculate for the nonzwitterionic forms of *S*-alanine, and the difference, $IP_1 - IP_2 = -1.3$ eV, computed for *S*-alanine is about the same as the experimentally determined difference reported for acetic acid.

The absorption and CD spectra of amino acids are most commonly obtained in solution using ethanol or water as solvent and HCl, NaOH, or phosphate buffer to adjust the pH. It is clear that for such systems solute-solvent interactions are substantial and are probably specific. For this reason comparisons between experimental spectra and the spectroscopic properties calculated for isolated solute molecules must be made with some care and qualification. In discussing the computed properties which appear in Table 1 we shall emphasize the dependence of $[R]$ upon angle ϕ and the sensitivity of $[R]$ to protonation or deprotonation of the amino and carboxylate groups.

In section IV it was noted that the lowest energy transition computed for each zwitterionic structure involves a transfer of negative charge from the

Table 3. Energies of three highest occupied molecular orbitals^a

Structure ^b	ϕ	Orbital energies (in au)		
		1	2	3
1 (Z)	-19°	-0.339 (⊥)	-0.358 ()	-0.391 ()
2 (Z)	-60°	-0.334 (⊥)	-0.358 ()	-0.391 ()
3 (Z)	30°	-0.337 (⊥)	-0.360 ()	-0.391 ()
4 (Z)	0°	-0.338 (⊥)	-0.359 ()	-0.391 ()
5 (A)	0°	-0.124 ()	-0.127 (⊥)	-0.171 ()
6 (C)	0°	-0.708 ()	-0.727 (⊥)	-0.797 ()
7 (NZ)	0°	-0.461 ()	-0.508 (⊥)	-0.563 (N)
8 (NZ)	-19°	-0.460 ()	-0.509 (⊥)	-0.563 (N)
9 (NZ)	60°	-0.461 ()	-0.515 (⊥)	-0.567 (N)
10 (NZ)	90°	-0.458 ()	-0.514 (⊥)	-0.566 (N)

^aIn the table these orbitals are designated as (||), (⊥), or (N) according to whether they place maximum electron density: (||)—in the plane of the carboxylate group; (⊥)—perpendicular to the plane of the carboxylate group; (N)—on the amino (or ammonium) group N atom.

Z ≡ zwitterionic species; A = anion; C ≡ cation; NZ ≡ nonzwitterionic neutral species.

carboxylate group to the ammonium group. This process leads to an excited state (Born-Oppenheimer or vibrationally relaxed) with a dipole moment considerably smaller than that of the ground state. In polar media (especially hydroxylic solvents) the absorption frequency of this type of transition should be blue shifted from the vapor phase value. The second lowest energy transition calculated for structures 1-4 has been described in section IV as a $n' \rightarrow \pi^*$ carboxylate excitation. This type of transition also will be blue shifted in going from the vapor phase to a polar solvent medium. There is little question, however, that the intergroup charge transfer transition will suffer a significantly larger blue shift than will the more localized $n' \rightarrow \pi^*$ transition. This suggests that, in aqueous or ethanolic solution, the first two transitions computed for the zwitterionic structures will have similar energies or their energy ordering will be inverted from that predicted for isolated molecules.

We note from Table 1 that the sign and magnitude of $[R]$ are quite sensitive to the angle ϕ for both the $n' \rightarrow \pi^*$ and the low energy charge-transfer (CT) transitions in the zwitterion structures. In fact, $[R_{n,\pi^*}] > 0$ at $\phi = 30^\circ$, but $[R_{n,\pi^*}] < 0$ at $\phi = 0^\circ, -19^\circ$, and -60° . Similarly, $[R_{CT}] < 0$ at $\phi = 30^\circ$, and $[R_{CT}] > 0$ at $\phi = 0^\circ, -19^\circ$, and -60° . For *S*-alanine in buffered aqueous solution at pH7, the lowest energy CD band is positive and exhibits a maximum at 204 nm ($\Delta\epsilon = 0.68$).^{12f} If this band is assigned to the $n' \rightarrow \pi^*$ transition, then our results suggest that zwitterionic structures with $0 \geq \phi \geq -60$ are not present in any significant concentration. An alternative interpretation, however, is that the first band is due to the low energy charge-transfer transition and that the rotamer equilibrium mixture in solution favors the structure found in crystalline *S*-alanine ($\phi \sim -19^\circ$).

In comparing the calculated values of $[R_{n,\pi^*}]$ for structures 4 (*Z*, $\phi = 0^\circ$), 5 (*A*, $\phi = 0^\circ$), 6 (*C*, $\phi = 0^\circ$), and 8 (*NZ*, $\phi = 0^\circ$) we find that the signs are identical. That is, $[R_{n,\pi^*}] < 0$ for the zwitterion, non-zwitterion, anion, and cation species at $\phi = 0$. Furthermore, $[R_{CT}] > 0$ for the lowest energy charge transfer transition in both the zwitterion and the cation species at $\phi = 0^\circ$.

Jorgensen^{12b} has proposed a sector rule for α -amino acids which is based on the octant rule for ketones.²⁷ This rule is directly applicable to structures with symmetrical carboxylate groups (e.g., the ionized carboxylate group in the zwitterionic species and in the anionic species), and purportedly "relates the sign and amplitude of the first Cotton effect to the conformation and absolute configuration of α -amino acids." Implicit in Jorgensen's proposal is the assumption that the first Cotton effect is centered around 205-215 nm (apparently correct for aqueous solutions at pH7) and that this Cotton effect can be assigned to a

carboxylate $n' \rightarrow \pi^*$ transition. Jorgensen first considers the $-\text{COO}'$ moiety as two equivalent ketone groups separated by an angle of 120° . He then partitions the space about the carboxylate group by superimposing the nodal patterns of the two functions, XYZ and $X'Y'Z'$, where the (X, Y, Z) and (X', Y', Z') coordinate systems are centered on the carboxylate carbon atom, Z and Z' are directed along the $\text{C}-\text{O}$ and $\text{C}-\text{O}'$ bond axes, respectively, the Y and Y' axes are coincident and are directed perpendicular to the plane of the $-\text{COO}'$ group and X and X' are defined such that (X, Y, Z) and (X', Y', Z') form right-handed coordinate systems. The sector rule is defined, therefore, by the nodal surfaces of the function, $F = (XYZ) + (X'Y'Z')$. If we transform the coordinates (X, Y, Z) and (X', Y', Z') to the coordinates (α, β, γ) , defined in Fig 2, then Jorgensen's sector rule may be expressed in terms of the function $\theta = (\alpha\beta\gamma)$. This yields an octant rule with nodal planes defined as $\alpha\beta$, $\alpha\gamma$, and $\beta\gamma$.

Although our discussion of the amino acid sector rule differs in several respects from Jorgensen's original description, the essential features of the rule have been retained. How do our calculations for structures 1-4 compare with the predictions of this sector rule? In Fig 3 are plotted the contributions to $\theta(\alpha, \beta, \gamma)$ made by the methyl carbon atom, the ammonium nitrogen atom, and the hydrogen atom substituents on the α -C atom of *S*-alanine zwitterion structures. The signs and relative magnitudes of the θ vs ϕ curves shown in Fig 3 should approximate those for the $-\text{CH}_3$, $-\text{NH}_3^+$, and $-\text{H}$ substituent groups in zwitterionic *S*-alanine structures. Below, we list the values of θ (in \AA^3)

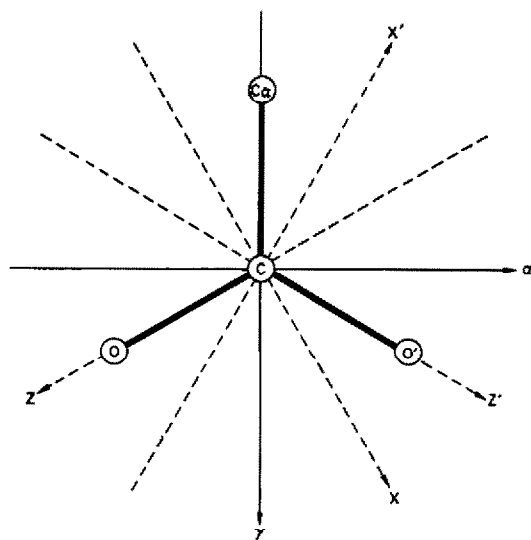


Fig 2. Amino acid sector rule.^{12b} View of $\text{C}_\alpha-\text{COO}'$ plane from the top. The coordinate axes Y, Y' , and β are coincident and are directed above the plane of the figure.

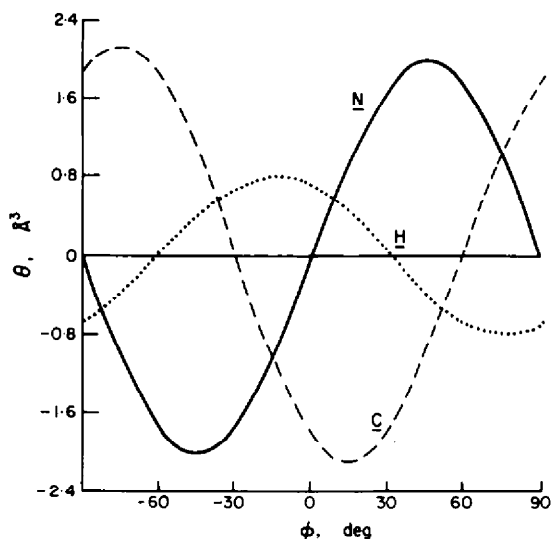


Fig. 3. Plot of θ versus ϕ for the C, N, and H substituents on the α -carbon atom in zwitterionic *S*-alanine, $\theta = (\alpha\beta\gamma)$, where α , β , and γ are defined in Fig. 2.

for the C, N, and H α -substituents in our structures 1-4, along with the calculated values of $[R_n\pi^*]$.

	N	C	H	$[R_n\pi^*]$
1	-1.24	-0.89	0.80	-38.2
2	-1.73	1.82	0	-17.0
3	1.73	-1.82	0	34.7
4	0	-1.82	0.69	-22.8
5	0	-1.82	0.69	-24.4

Note that the sign of $[R_n\pi^*]$ correlates with the sign of $\theta(N)$ for structures 1-3. In the context of the one-electron perturbation model of molecular optical activity, this implies that the $-\text{NH}_3^+$ group provides a stronger dissymmetric perturbative influence on the $n' \rightarrow \pi^*$ transition than do the $-\text{CH}_3$ and $-\text{H}$ α -substituents. In structures 4 and 5, the ammonium N atom lies on a nodal plane of the $(\alpha\beta\gamma)$ function, and the sign of $[R_n\pi^*]$ should, therefore, be determined by the positions of the $-\text{CH}_3$ and $-\text{H}$ groups. For these structures, the computed values of $[R_n\pi^*]$ have the same sign as $\theta(C)$.

How can the CD spectrum obtained by Johnson²⁰ for zwitterionic *S*-alanine in HFIP solvent be interpreted in terms of our calculated results? Recall that Johnson reports two bands in the 220 to 160 nm region, one centered at 196 nm with $\Delta\epsilon(\text{max}) \cong 1.0$ and one centered at 168 nm with $\Delta\epsilon(\text{max}) \cong -1.0$. Furthermore, the rotatory strengths of the long wavelength and short wavelength CD bands are reported to be $\sim 2.8 \times 10^{-40}$ cgs and $\sim -2.3 \times 10^{-40}$ cgs, respectively. If the

first band is assigned to an intergroup charge-transfer transition and the second band is assigned to a localized $n \rightarrow \pi^*$ carboxylate transition, then our results are consistent with $0^\circ > \phi > -60^\circ$ for the predominate conformational isomers in solution. However, if the first band is assigned to the $n \rightarrow \pi^*$ transition and the second band is assigned to an intergroup charge-transfer transition (shifted to a higher energy than the $n \rightarrow \pi^*$ transition by solvent perturbations), then the calculated results for $\phi = 30^\circ$ are in qualitative agreement with the observed CD spectrum. As was pointed out earlier in this section, the charge transfer transition should be substantially more sensitive to solvent perturbations than the $n \rightarrow \pi^*$ transition.

Our computed rotatory strengths for the zwitterionic species are an order-of-magnitude larger than those determined experimentally by Johnson.²⁰ However, Johnson's finding that the net rotatory strength of the first two CD bands is very small agrees with our calculation for the net rotatory strength of the two lowest-energy transitions. For the γ th component of the $n \rightarrow \pi^*$ magnetic dipole transition moment (see Fig. 2 for definition of α, β, γ coordinate system) we calculate a value of $|M_{n\pi^*}^\gamma| = 0.288$ Bohr magneton at $\phi = 0^\circ$. Generally, our model yields quite reasonable values for the magnetic dipole transition moments. However, our calculated electric dipole transition moments are an order-of-magnitude larger than those deduced from experimentally determined dipole strengths. This accounts for our computed rotatory strengths being too large by a factor of ten.

The sign pattern and net rotatory strength exhibited by the CD spectrum of zwitterionic *S*-alanine in the region $160 < \lambda < 220$ nm can be rationalized on our model. However, the interpretation requires band assignments which differ from the conventional assignments for carboxylate compounds. The independent systems model has also been used to rationalize the CD spectrum, but in this case conventional band assignments are adopted.²⁰ Clearly, the most serious sources of error in our method are the approximations inherent in the CNDO molecular orbital model and, of course, the construction of excited states in the virtual orbital approximation. Application of the independent systems model to zwitterionic *S*-alanine suffers from: (1) uncertainties regarding the nature of the spectroscopic states in an isolated carboxylate chromophore; (2) the rather drastic assumption of group separability, especially since the α -substituents are only about 2.5 \AA removed from the centroid of the carboxylate group (the hydrogen atoms are much closer); and, (3) the absence of accurate bond polarizability data. Neither our model nor the independent systems model can provide a *conclusive* description of the detailed relationships between the CD observables and specific structural features of amino acids.

Such a description must await more accurate electronic structure calculations or a more complete characterization of the spectroscopic states responsible for absorption and CD in the 160–220 nm spectral region.

Acknowledgements—We express our gratitude to the donors of the Petroleum Research Fund, administered by the American Chemical Society, for financial support of this research. We are also grateful to Dr. W. C. Johnson for making the results of his studies available to us prior to publication, and to a referee for bringing Johnson's work to our attention.

REFERENCES

- ¹This work was supported by a grant from the Petroleum Research Fund (PRF No. 2022-G2), administered by the American Chemical Society
- ²See, for example: ^aP. J. Urnes and P. Doty, *Advan. Protein Chem.* **16**, 401 (1961); ^bW. F. Harrington, R. Josephs and D. M. Segal, *Ann. Rev. Biochem.* **35**, 599 (1966); ^cS. Beychok, *Science*, **154**, 1288 (1966); ^dJ. T. Yang, in *Poly- α -Amino Acids* (Edited by G. D. Fasman) pp. 239–91. Marcel Dekker, New York (1967); ^eS. Beychok, *Poly- α -Amino Acids* (Edited by G. D. Fasman) pp. 293–337. Marcel Dekker, New York (1967); ^fS. N. Timasheff and M. Gorbunoff, *Ann. Rev. Biochem.* **36**, 13 (1967); ^gS. Beychok, *Ibid.* **37**, 437 (1968); ^hC. W. Deutsche, D. A. Lightner, R. W. Woody and A. Moscowitz, *Ann. Rev. Phys. Chem.* **20**, 407 (1969); ⁱS. N. Timasheff, *The Enzymes* (Edited by P. D. Boyer) Vol. II, pp. 381–407. Academic Press, New York (1970)
- ³See, for example: ^aD. D. Fitts and J. G. Kirkwood, *Proc. Natl. Acad. Sci. U.S.A.* **42**, 33 (1956); ^bW. Moffitt, *J. Chem. Phys.* **25**, 467 (1956); ^cW. Moffitt, *Proc. Natl. Acad. Sci. U.S.A.* **42**, 736 (1956); ^dW. Moffitt, D. D. Fitts and J. G. Kirkwood, *Ibid.* **43**, 723 (1957); ^eI. Tinoco and R. W. Woody, *J. Chem. Phys.* **32**, 461 (1960); ^fJ. Schellman and P. Oriol, *Ibid.* **37**, 2114 (1962); ^gI. Tinoco, R. W. Woody and D. F. Bradley, *Ibid.* **38**, 1317 (1963); ^hD. F. Bralday, I. Tinoco and R. W. Woody, *Biopolymers* **1**, 239 (1963); ⁱE. S. Pysh, *Natl. Acad. Sci. U.S.A.* **56**, 825 (1966); ^jR. W. Woody and I. Tinoco, *J. Chem. Phys.* **46**, 532 (1967); ^kR. W. Woody, *Ibid.* **49**, 4797 (1968); ^lW. Rhodes, *Ibid.* **53**, 3650 (1970); ^mE. S. Pysh, *Ibid.* **52**, 4723 (1970); ⁿC. W. Deutsche, *Ibid.* **52**, 3703 (1970); ^oW. Rhodes and D. G. Barnes, *Ibid.* **48**, 817 (1968); ^pW. Rhodes and D. G. Barnes, *J. Chim. Phys.* **65**, 78 (1968); ^qF. Loxsom, L. Tterlikkis and W. Rhodes, *Biopolymers* **10**, 2405 (1971)
- ⁴W. B. Gratzler, *Poly- α -Amino Acids* (Edited by G. Fasman) Marcel Dekker, New York (1967)
- ⁵J. A. Schellman and C. Schellman, *The Proteins* (Edited by H. Neurath) 2nd edition Vol. II, pp. 1–137. Academic Press, New York (1964)
- ⁶H. D. Hunt and W. T. Simpson, *J. Am. Chem. Soc.* **75**, 4540 (1953); ⁷E. E. Barnes and W. T. Simpson, *J. Chem. Phys.* **39**, 670 (1963); ⁸D. L. Peterson and N. T. Simpson, *J. Am. Chem. Soc.* **77**, 3929 (1955); ⁹*Ibid.* **79**, 2375 (1957); ¹⁰B. J. Litman and J. A. Schellman, *J. Phys. Chem.* **69**, 978 (1965); ¹¹E. B. Nielson and J. A. Schellman, *Ibid.* **71**, 2297 (1965); ¹²H. Basch, M. B. Robin and N. A. Kuebler, *Ibid.* **47**, 1201 (1967); ¹³H. Basch, M. B. Robin and N. A. Kuebler, *Ibid.* **47**, 1201 (1967); ¹⁴*Ibid.* **49**, 5007 (1968)
- ¹⁵D. W. Urry, *Ann. Rev. Phys. Chem.* **19**, 477 (1968); ¹⁶J. A. Schellman and E. B. Nielson, *J. Phys. Chem.* **71**, 3914 (1967); ¹⁷D. Stigter and J. A. Schellman, *Ibid.* **51**, 3397 (1969); ¹⁸P. M. Bayley, E. B. Nielson and J. A. Schellman, *Ibid.* **73**, 228 (1969); ¹⁹F. S. Richardson, R. Strickland and D. D. Shillady, *Ibid.* in press
- ²⁰F. S. Richardson, D. D. Shillady and J. E. Bloor, *Ibid.* **75**, 2466 (1971)
- ²¹J. Webb, R. Strickland and F. S. Richardson, *J. Am. Chem. Soc.* submitted for publication
- ²²M. V. Vol'kenshtein and M. P. Kruchek, *Zh Strukturnoi Khimii* **2**, 59 (1961); Russian
- ²³For surveys of the literature, see: ^aC. Djerassi, *Optical Rotatory Dispersion. Applications to Organic Chemistry* Chapter 15. McGraw-Hill, New York (1960); ^bP. Crabbe, *Optical Rotatory Dispersion and Circular Dichroism in Organic Chemistry*, Chapter 11. Holden-Day, San Francisco (1965); ^c*Optical Rotatory Dispersion and Circular Dichroism in Organic Chemistry* (Edited by G. Sznatzke) Chap 12. Heyden, London (1967); ^dP. Crabbe, *An Introduction to the Chiroptical Methods in Chemistry* Chapter II. Syntex Corp., Mexico (1971)
- ²⁴See, for example: ^aJ. P. Jennings, W. Klyne and P. M. Scopes, *J. Chem. Soc.* 294 (1965); ^bI. P. Dirks and F. L. J. Sixma, *Rec. Trav. Chim.* **83**, 522 (1964); ^cW. Gaffield, *Chem. & Ind.* 1460 (1964); ^dJ. C. Craig and S. K. Roy, *Tetrahedron* **21**, 391 (1965); ^eL. I. Katzin and E. Gulyus, *J. Am. Chem. Soc.* **90**, 247 (1968); ^fM. Legrand and R. Viennet, *Bull. Soc. Chem. Fr* 679 (1965); ^gE. C. Jorgensen, *Tetrahedron Letters* 863 (1971)
- ²⁵R. D. Anand and M. K. Hargreaves, *Chem. Commun.* 880 (1968)
- ²⁶J. C. Craig and W. E. Pereira, Jr., *Tetrahedron* **26**, 3457 (1970)
- ²⁷C. Toniolo, *J. Phys. Chem.* **74**, 1390 (1970)
- ²⁸R. D. Anand and M. K. Hargreaves, *Chem. Commun.* 421 (1967); ²⁹G. Barth, W. Voelter, E. Bennenberg and C. Djerassi, *Ibid.* 355 (1969)
- ³⁰D. W. Urry and H. Eyring, *J. Am. Chem. Soc.* **86**, 4574 (1964)
- ³¹W. D. Closson and P. Haug, *Ibid.* **86**, 2384 (1964)
- ³²See, for example: G. C. K. Roberts and O. Jardetzky, *Advan. Protein Chem.* **24**, 447 (1970)
- ³³W. C. Johnson, Jr., private communication
- ³⁴I. Tinoco, *Advan. Chem. Phys.* **4**, 113 (1962)
- ³⁵J. A. Pople and D. L. Beveridge, *Approximate Molecular Orbital Theory*. McGraw-Hill, New York (1970)
- ³⁶A. Imamura, T. Hirano, C. Nagata and T. Tsuruta, *Bull. Chem. Soc. Japan* **45**, 396 (1972)
- ³⁷J. J. Simpson and R. E. Marsh, *Acta Cryst.* **20**, 550 (1966)
- ³⁸T. Koopmans, *Physica* **1**, 104 (1933)
- ³⁹D. A. Sweigart and D. W. Turner, *J. Am. Chem. Soc.* **94**, 5592 (1972)
- ⁴⁰W. Moffitt, R. B. Woodward, A. Moscowitz, W. Klyne and C. Djerassi, *Ibid.* **83**, 4013 (1961)