

## Simulation of the Raman Optical Activity of L-Alanyl–L-Alanine

Petr Bouř,<sup>\*,†</sup> Josef Kapitán,<sup>†,‡</sup> and Vladimír Baumruk<sup>‡</sup>

*Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo nám 2, 16610, Prague, Czech Republic, and Institute of Physics, Charles University, Ke Karlovu 5, 12116, Prague, Czech Republic*

*Received: July 19, 2000; In Final Form: February 7, 2001*

Raman and Raman optical activity (ROA) spectra of zwitterionic dipeptide L-alanyl–L-alanine (AA) were measured and compared to previous experiments and ab initio computations. The molecule does not exist in a vacuum and classical approaches of quantum mechanics cannot be directly applied. Thus, the dipolar Onsager, ionic, and conductor-like screening (COSMO) continuum solvent models were used for calculation of geometries, harmonic vibrational frequencies, and Raman intensities. Reasonable agreement was found between experiment and the spectra, enabling one to assign most of the observed experimental bands. The results suggest that bulk water significantly increases conformational flexibility of the dipeptide by lowering the energy differences among its conformers. However, detailed conformational analysis was not attempted due to the limited accuracy of the calculations. Supposedly, the nearly extended conformation of the zwitterion prevails under experimental conditions. A simplistic polar model of ROA was used in a preliminary conformer search. The model that led to significant savings in computer time was also successfully applied for mono- and trialanine molecules.

### Introduction

Flexible peptide molecules in aqueous environment are natural targets of Raman optical activity (ROA) spectroscopy.<sup>1</sup> Data about their conformations and solvent interactions cannot be often obtained from other techniques, such as NMR, X-ray, or a low-angle neutron scattering. For example, due to the intrinsic fast response of the optical spectroscopy, the ROA spectrum is a sum of contributions from individual conformers, while the additivity is often lost in NMR due to the relaxation processes.<sup>2</sup> On the other hand, interpretations of ROA are dependent on extensive computer simulations, which restricts interpretation of the spectra for big molecules.<sup>3–6</sup> Typically, these quantum mechanical computations have been restricted to rigid molecules in a vacuum, because of computer time and memory limits. No rule between spectral shape and peptide secondary conformation has been found yet, unlike for the complementary vibrational circular dichroism (VCD) technique.<sup>7,8</sup> This is particularly challenging, since theoretical studies indicate such a dependence<sup>9</sup> and the progress in ROA instrumentation allows measurement of the spectra almost with the same sensitivity as for VCD and often in regions inaccessible by infrared techniques.<sup>10</sup> Specific difficulties arise for zwitterionic molecules, unstable in a vacuum and inaccessible by usual apparatus of quantum chemistry. More realistic solvent models are clearly necessary for reliable ROA spectra simulation. As a first step we consider the exploration of the potential and accuracy of the COSMO<sup>11,12</sup> and other solvent models as they have become available in the Gaussian<sup>13,14</sup> and Turbomole<sup>15</sup> software packages.

The L-alanyl–L-alanine (AA) molecule is a good model for more complex systems studied by ROA. Spectra of peptides and other biologically interesting polymers including sugars<sup>16</sup>

and DNA fragments<sup>17</sup> have been recorded. Also many smaller model molecules have been investigated with the purpose of explaining the relation between the structure and spectral response. Thus, the technique was found sensitive with respect to various ionic forms and isotopic isomers of alanine<sup>18,19</sup> and its oligomers.<sup>20–22</sup> Alanine spectra, as well as those of a simpler model dipeptide<sup>23</sup> were simulated with ab initio techniques on the HF/SCF level using the simplified Onsager dipole model of the solvent. A reasonable agreement with experimental data was achieved, although problems with a systematic overestimation of computed frequencies and consistent cavity definition were observed.<sup>19</sup> In this study, in an attempt to improve an overall accuracy of the results, namely, the harmonic vibrational frequencies, we use a combination of the modern BPW91<sup>24,25</sup> functional with the COSMO model instead. The BPW91 combination of functionals lead to a very good agreement between theory and experiment for many systems studied previously.<sup>26</sup> Although the ROA intensities cannot be calculated consistently in DFT due to software limitations, results obtained at various levels can be combined and realistic spectral patterns obtained.<sup>23,27</sup> We use Cartesian tensor transfer techniques<sup>27</sup> for combining various force fields and electromagnetic tensors, because of their flexibility and numerical stability. The tensors are then transferred as properties of individual atoms.

The COSMO model provided good estimates of molecular solvation energies in the past.<sup>28</sup> Unlike for the Onsager model, the molecular surface in COSMO reflects actual molecular shape; solvent electric permittivity and radius enter as parameters. A relative simplicity allows one to calculate energies and gradients analytically with a sufficient accuracy and in a feasible amount of time. Recent results for DNA fragments suggest that similar models are essential for modeling biomolecules in aqueous environment,<sup>29</sup> perhaps in combination with explicit solvent models.<sup>30</sup> However, such a combination was not feasible for our systems due to restrictions in available programs, namely for demands on the cavity surface.

<sup>†</sup> Academy of Sciences of the Czech Republic. E-mail: bour@uochb.cas.cz, jkap6258@menza.mff.cuni.cz. Fax: (420-2)-2431-0503.

<sup>‡</sup> Charles University. E-mail: baumruk@karlov.mff.cuni.cz. Fax: (420-2)-2492-2797.

According to our knowledge, computation of the ROA polarizability tensors in the presence of solvent has not been implemented yet. Even the vacuum-based calculation is severely limited by a lengthy numerical differentiation. Fortunately, trial computations suggest that the intensities can be calculated with sufficient accuracy in a vacuum only. For example, no qualitative difference in relative peak intensities was observed for Raman spectra simulated in a vacuum and with the use of the Onsager model. Thus, we combine the vacuum polarizabilities with the COSMO force field.

In the past, polar<sup>4,31</sup> and SOS<sup>5,9</sup> simplified models were proposed for computationally cheaper modeling of ROA. As shown below, the former provides a reliable tool for estimation of spectral intensities for our systems. The polar models resemble the two-group or bond polarizability models,<sup>32</sup> but a black box method is more suitable for machine computations. The SOS models are not further followed in this study because of their complexity.

## Experiment

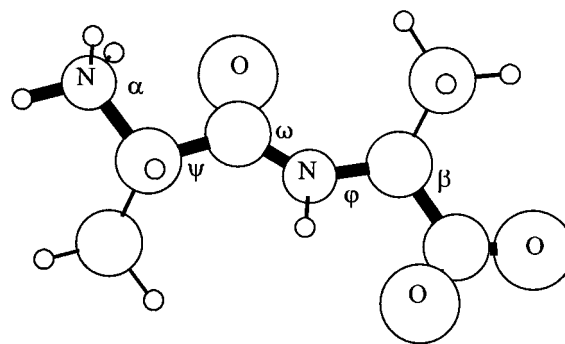
L-Alanyl-L-alanine (Sigma) was dissolved in deionized double distilled water at final concentration of about 3 mol/L. Backscattering ICP ROA and Raman spectra were collected on the instrument built at the Charles University, which has been described in detail elsewhere.<sup>4</sup> The spectra were produced with a 514.5 nm line of the argon laser with the excitation power at the sample of about 500 mW. The spectral slit width was set to 10 cm<sup>-1</sup>. Total acquisition time was approximately 15 h.

## Computation

Equilibrium geometries, harmonic force fields, and Raman intensities were calculated with the Gaussian program using the BPW91 DFT functional and the COSMO solvent model. Obtained harmonic frequencies were used without further scaling. An optimum size of the differentiation step was found for each system (typically 0.0001–0.01 Å), to minimize the number and absolute values of imaginary frequencies. These were often obtained as a result of limited accuracy of COSMO gradients<sup>12,14</sup> and their effect on the experimentally accessible region was negligible. Atomic derivatives of the electric polarizability ( $\alpha$ ) were calculated in a vacuum at the HF level. As follows from our current computations, the overall error of the computations stems almost entirely from the force field inaccuracies and the dependence of polarizabilities on solvent permittivity does not change the ROA spectral pattern. Similarly, atomic derivatives of the electric dipole–electric quadrupole (**A**) and optical activity (**G'**) tensors were obtained at the HF level, using Dalton<sup>33</sup> or Cadpac<sup>34</sup> programs in a vacuum. The frequency dependence and London atomic orbitals<sup>35</sup> implemented in Dalton were found insignificant for our systems and thus these two programs were used interchangeably. However, the Dalton package seemed more versatile and provided results slightly faster than Cadpac. The Turbomole<sup>15</sup> program was used for an additional estimation of conformer energies. For a fast estimation of ROA spectral shape for individual conformers we used the polar model, where derivatives of **A** and **G'** are obtained from  $\alpha$ :

$$\frac{\partial G'_{\alpha\beta}}{\partial R_\epsilon^\lambda} \Big|_0 \omega^{-1} = -\frac{1}{2} \epsilon_{\alpha\beta\gamma} Y_\gamma^\lambda \frac{\partial \alpha_{\alpha\delta}}{\partial R_\epsilon^\lambda} \quad (1)$$

$$\frac{\partial A_{\alpha\beta\gamma}}{\partial R_\epsilon^\lambda} \Big|_0 = \frac{3}{2} \left[ Y_\beta^\lambda \frac{\partial \alpha_{\alpha\gamma}}{\partial R_\epsilon^\lambda} + Y_\gamma^\lambda \frac{\partial \alpha_{\alpha\beta}}{\partial R_\epsilon^\lambda} \right] - Y_\delta^\lambda \frac{\partial \alpha_{\alpha\delta}}{\partial R_\epsilon^\lambda} \delta_{\beta\gamma} \quad (2)$$



**Figure 1.** Definition of the torsion angles in the AA zwitterion. Angle  $\alpha$  ( $\angle$ HNCC) defined for hydrogen closest to the NCC plane; the negative angle  $\beta$  ( $\angle$ NCCO) is given in Table 1.

Vector  $\mathbf{Y}^\lambda$  defines the equilibrium position of an atom  $\lambda$ ,  $\mathbf{R}^\lambda$  is its coordinate,  $\omega$  is the frequency of exciting light, and  $\epsilon$  is the antisymmetric tensor. Note that derivatives in the laboratory system ( $\partial/\partial R_\epsilon^\lambda|_0$ ) and in that originating at nucleus  $\lambda$  ( $\partial/\partial R_\epsilon^\lambda|_\lambda$ ) are not equal for **G'** and **A**, unlike for  $\alpha$ . Simulated spectra were obtained from calculated intensities using Lorentzian bands and a bandwidth of 5 cm<sup>-1</sup>, i.e., slightly smaller than the experimental resolution, so that individual modes could be better seen.

## Results and Discussion

**Geometry.** A detailed discussion about the conformation for zwitterionic AA can be found in ref 36. We have sampled the conformational space with respect to the torsion angles  $\alpha$ ,  $\beta$ ,  $\varphi$ , and  $\psi$ , as defined in Figure 1. The angle  $\omega$  was not fixed except for the BPW91/6-311G\*\* computation (Tables 4 and 6; Figures 2, 5, and 7), where  $\omega = 180^\circ$ . For this case also a nonstandard Gaussian option of 80 tesserae was used. Angles  $\varphi$  and  $\psi$  correspond to the key angles defining the secondary structure of long peptides.

Trial computations in a vacuum (not shown here) did not provide any stable zwitterionic forms of AA. Typically, an enolic form (CH<sub>3</sub>NH<sub>2</sub>HC-(HO)C=N-CCH<sub>3</sub>COOH) of the dipeptide was a result of such attempts. We do not expect a presence of this isomer under our experimental conditions and its Raman and ROA spectra could not be reasonably related to the measurement. Similar problems were encountered with models explicitly considering water molecules, because of the large volume needed for stabilizing the charged termini. Additionally, explicit waters contribute to the simulated ROA signal, which is then difficult to relate to the experiment. Combination of explicit water molecules with a continuum solvent model proved to be numerically unstable. Thus, we based our analysis primarily on the implicit COSMO model of solvent. Other continuum models (e.g., polarizable continuum model, PCM) as implemented in Gaussian failed for molecules of this size.

Relative energies of AA conformations are listed in Table 1, as calculated by the COSMO model. The conformations can be roughly classified as planar (a–d),  $\varphi$ -bent (e–g),  $\varphi,\psi$ -bent (h–l), conformation m corresponds to the cis form of the amide group, and finally, conformation n has the biggest dipole moment. Generally, quite shallow potential minima were found and, because of the limited accuracy of COSMO, the default convergence criteria tailored for the vacuum had to be relaxed, unlike for similar calculations reported on smaller systems previously.<sup>13</sup> Energy differences obtained for the rotation of the NH<sub>3</sub><sup>+</sup> and CO<sub>2</sub><sup>-</sup> groups are quite low and within the expected error of the calculation. Thus, detailed conformation ratios

**TABLE 1: Relative Energies (kcal/mol) and Torsion Angles for a Set of L-Alanyl-L-Alanine Conformations As Obtained by the COSMO Model**

conformation	<i>E</i>	$\alpha$	$\psi$	$\omega$	$\varphi$	$\beta$
BPW91/6-31G*						
a	0	-5	179	175	-162	-7
b	0	171	156	177	-158	-9
c	2	-10	180	-176	-151	-64
d	3	-177	175	180	-142	-66
e	6	176	-175	176	85	-19
f	6	176	-165	-178	-64	-32
g	6	-179	-158	-178	-103	-142
h	1	158	146	173	-117	-40
i	5	171	121	172	-100	-52
j	8	-179	-68	167	-89	-21
k	4	178	56	179	-129	-32
l	4	-179	60	167	-149	-158
m	4	3	165	-4	-135	-15
n	16	-178	1	175	4	-92
BPW91/6-311G**						
a	0	-3	179	175	-162	-6
b	0	164	158	169	-146	-7
c	3	-10	180	-177	-151	-64
d	0	160	153	173	-138	-41

Angles defined in Figure 1.

**TABLE 2: Relative Energies (kcal/mol) of the Four Planar AA Conformers As Calculated in Vacuum, with the Onsager and COSMO Solvent Corrections**

	vacuum <sup>a</sup>	Onsager <sup>a</sup>	COSMO	COSMO	COSMO	COSMO
	BPW916-31G*	BPW916-31G*	HF6-31G	BPW916-31G*	BPW916-311G**	B3LYPSVP
a	0	0	0	0	0	0
b	7	2	-2	0	0	0
c	12	8	3	2	3	1
d	17	11	2	3	0	2

<sup>a</sup> COSMO/BPW91/6-31G\* geometry.

cannot be predicted. Nevertheless, simulated spectra suggest that the planar conformation a prevails in the sample and thus the vibrational assignment presented below is done for the first conformer. Although the geometry parameters are similar if calculated in bigger 6-311G\*\* basis (bottom of Table 1), the energy ordering is not stable and the most stable form, if it exists, cannot be determined solely on the basis of energy calculation.

In Table 2, relative energies for the first four conformers are given for six levels of approximations. In accordance with previous computations<sup>29</sup> the solvent lowers energy differences of individual conformers if compared to vacuum. Since the differences are comparable to the temperature Boltzmann quantum, the presence of solvent has a dramatic influence on conformer ratios. On the other hand, a minor contribution of molecular relaxation is expected,<sup>29</sup> which justifies estimation of the vacuum relative energies with the COSMO geometry, as in Table 2. The Turbomole values (last column of Table 2) are consistent with the results obtained by Gaussian. The magnitude and trend of the differences between vacuum and solvent energies are also consistent with the conformational analysis of  $\beta$ -alanine, where a PCM model was compared to NMR data.<sup>37</sup>

**Vibrational Frequencies of Alanine.** We calculated harmonic frequencies of zwitterionic alanine in order to relate our modeling to previous results. The values are listed in Table 3, as calculated in the 6-31G\* basis set for the HF Onsager model with default parameters from Gaussian, a modified Onsager model with optimized cavity radius according to ref 19, the HF/COSMO, BPW91/Onsager, and finally, the BPW91/COSMO model.

**TABLE 3: Vibrational Frequencies (cm<sup>-1</sup>) of Alanine Zwitterion<sup>a</sup>**

	HF	HF modified	HF	BPW91 <sup>c</sup>	BPW91	
	Onsager	Onsager <sup>b</sup>	COSMO	Onsager	COSMO	exp <sup>b</sup>
	1939	1876	1832	1688	1678	1631
	1841	1836	1786	1642	1608	
	1801	1802	1733	1617	1580	
	1642	1652	1617	1489	1473	1503
	1639	1642	1634	1483	1475	1457
	1597	1635	1576	1402	1393	
	1564	1541	1550	1365	1362	1412
	1520	1578	1510	1299	1315	1374
	1500	1501	1471	1343	1355	1352
	1441	1452	1511	1271	1250	1302
	1314	1317	1274	1194	1193	1217
	1214	1222	1195	1096	1082	1139
	1183	1186	1195	1071	1078	1113
	1069	1060	1051	997	989	1003
	1059	1048	1082	973	971	995
	955	970	1025	851	859	922
	885	893	889	810	808	848
	840	844	837	734	740	781
	657		711	599	594	
$\alpha^d$	1.10	1.10	1.09	0.98	0.98	(1)
$\delta^d$	48	40	39	36	33	(0)

<sup>a</sup> All calculations were done with the 6-31G\* basis set. <sup>b</sup> Reference 19. <sup>c</sup> For COSMO geometry, zwitterion not stable. <sup>d</sup>  $a$ ,  $\delta$  ... slope of a linear fit  $\omega_{\text{calc}} = a\omega_{\text{exp}}$  and its mean deviation ( $\delta = [\sum_i(x_i - y_i)^2/n]^{1/2}$ , in cm<sup>-1</sup>).

The HF Hamiltonian leads to a large overestimation of vibrational frequencies, on the average by 10% in this region, cf. the fitting coefficients  $a$  in Table 3. Namely, the asymmetric C=O stretching of the carboxyl group (experimentally at 1631 cm<sup>-1</sup>), but also the NH bending modes (at about 1600 cm<sup>-1</sup>) are overestimated by more than 200 cm<sup>-1</sup>. The modified Onsager model improves this mode but does not provide statistically significant improvement. On the other hand, the COSMO frequencies (third column of the table) are more realistic; improvements of 35–107 cm<sup>-1</sup> for the four highest frequency modes can be observed. The model, however, causes reordering of several modes and the overall error remains rather high, cf. coefficients  $a$  and  $\delta$ .

A more significant improvement can be achieved with the BPW91 functional (fourth and fifth columns of Table 3). The average error of 10% for the HF calculations was reduced to 2% ( $a = 0.98$ ) and also the square root deviation is smaller by 6–15 cm<sup>-1</sup>. The BPW91/COSMO frequencies are slightly more accurate ( $\delta = 33$ ) than the BPW91/Onsager results. Thus, we can conclude that the COSMO/BPW91 model provides the most accurate frequency pattern for the alanine zwitterion.

**Raman Spectra of AA.** In Table 4 calculated harmonic frequencies for the AA zwitterion are listed and compared to experiment. The assignment is based on BPW91/Onsager/6-31G\*, BPW91/COSMO/6-31G\*, and BPW91/COSMO/6-311G\*\* calculations of harmonic force fields for conformation a. Additionally, for this molecule an ionic “KCl” model was investigated, where the zwitterionic form was stabilized by potassium and chlorine ions in a complex [Cl<sup>-</sup>...NH<sub>3</sub><sup>+</sup>-AA-CO<sub>2</sub><sup>-</sup>...K<sup>+</sup>]. The K<sup>+</sup> and Cl<sup>-</sup> ions were placed at the charged group symmetry axes; otherwise the complex geometry was fully optimized. Besides its simplicity, this model allows standard vacuum analytic computations of frequencies and Raman intensities at the same level of approximation.

Curiously, similar average frequency errors (about 1%, see the coefficient  $a$  in Table 4) of the dispersions  $\delta$  (25–40 cm<sup>-1</sup>) were obtained for all models. However, such a statistics does not capture occasional switches in the normal mode ordering,

**TABLE 4: Comparison of Calculated and Experimental Frequencies (cm<sup>-1</sup>) for the AA Zwitterion in Mid-IR Region<sup>a</sup>**

mode no.	Onsager	ionic	COSMO	COSMO	exp	
	BPW/6-31G*	BPW/6-311G**	BPW/6-31G*	BPW/6-311G**		
51	1692	1707	1669	1653	1677	CONH, C=O stretch (amide I)
50	1676	1599	1639	1591	1640	COO <sup>-</sup> , C=O stretch, out of phase
49	1633	1605	1596	1574		NH <sub>3</sub> <sup>+</sup> , NH scissoring
48	1643	1567	1576	1554	1572	NH <sub>3</sub> <sup>+</sup> , NH scissoring
47	1492	1479	1484	1471		amide II (C-N stretch) + CH bend
46	1487	1456	1474	1447		CH <sub>3</sub> scissoring
45	1482	1449	1471	1435	1459	CH <sub>3</sub> scissoring
44	1481	1455	1466	1442		CH <sub>3</sub> scissoring
43	1449	1448	1454	1434		amide II (C-N stretch)+CH bend
42	1421	1410	1398	1405	1414	CH <sub>3</sub> and NH <sub>3</sub> <sup>+</sup> umbrella modes
41	1398	1372	1388	1386		CH <sub>3</sub> and NH <sub>3</sub> <sup>+</sup> umbrella modes
40	1364	1409	1372	1372	1383	CH <sub>3</sub> and NH <sub>3</sub> <sup>+</sup> umbrella modes
39	1365	1363	1359	1348	1369	C(O)-N stretch (amide III)
38	1340	1323	1327	1340	1335	COO <sup>-</sup> , C=O stretch, in phase
37	1323	1333	1302	1313	1320	CH, NH bending
36	1288	1280	1246	1279	1266	<sup>a</sup> CH bending, amide III
35	1256	1247	1207	1241	1169	C(O)-N stretch (amide III)
34	1191	1176	1166	1164	1168	NH <sub>3</sub> <sup>+</sup> , CH <sub>3</sub> wagging
33	1159	1137	1128	1180	1120	N- <sup>a</sup> C stretch, CH bending
32	1116	1113	1103	1110	1112	N- <sup>a</sup> C stretch, deloc
31	1100	1078	1084	1089	1098	NH wagging
30	1089	1051	1079	1084	1106	<sup>a</sup> CH bending, CH <sub>3</sub> wagging, amide group def.
29	1033	1034	1037	1033		NH <sub>3</sub> <sup>+</sup> wagging
28	1064	1051	1017	1065	1047	CH <sub>3</sub> wagging
27	971	1137	989	970	1011	NH <sub>3</sub> <sup>+</sup> wagging, C- stretch
26	989	890	965	978	981	NH <sub>3</sub> <sup>+</sup> wagging, C- stretch
25	890	917	909	907	952	NH <sub>3</sub> <sup>+</sup> wagging, C- stretch
24	837	898	862	896	928	CH <sub>3</sub> wagging, deloc
23	890	858	848	856	885	C-NH <sub>3</sub> <sup>+</sup> stretch, deloc
22	808	830	815	843	855	deloc. deformation mode
21	731	756	742	751	777	C=O bend
20	774	736	724	731	770	NH out of plane, deloc
19	699	653	664	671	702	NH out of plane, deloc
18	619	632	615	621	577	NH out of plane, deloc
17	597	613	563	608	662	NH out of plane, deloc
16	554	564	554	558	583	CC(O)C bending, deloc.
<i>a</i>	0.998	0.997	0.986	0.992	1	
$\delta$	41	42	25	30	0	

<sup>a</sup>Numbering corresponds to the computed total of 3N-6 = 63 normal modes; the description is based on the dynamic visualization of their displacement. Coefficients *a*,  $\delta$  as in Table 1.

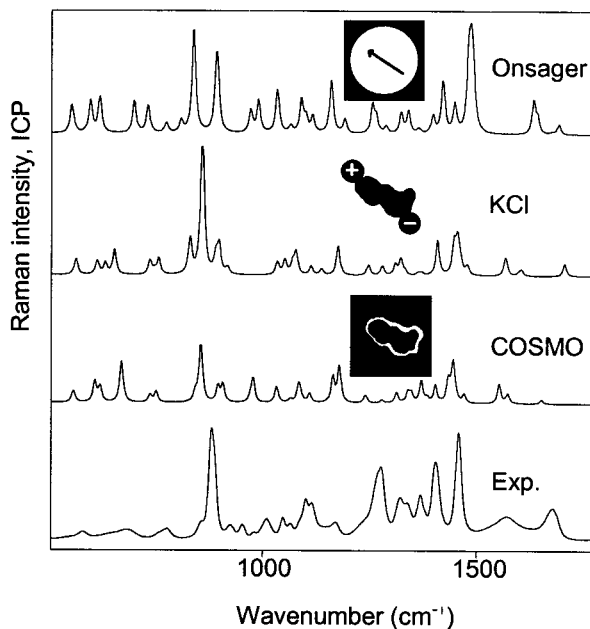
e.g., for modes 48–49, 39–40, and 28–29. Generally, satisfying agreement between calculated and experimental frequencies can be observed, namely for the highest-frequency modes, which are traditionally difficult to get precisely with ab initio simulations. This rather contradicts the popular belief of a specific role of the hydrogen binding for carbonyl group vibrations.

In Figure 2 Raman intensities are plotted and compared to experiment. Rather surprisingly, the COSMO and KCl intensity spectral patterns are similar and superior to that of Onsager. This indicates the importance of a proper localization of partial molecular charges, not possible for the Onsager dipole model. For the Onsager model, an overall increase of intensities can be observed if compared with COSMO, while only vibrations involving the charged NH<sub>3</sub><sup>+</sup> and CO<sub>2</sub><sup>-</sup> groups are more intense. Unfortunately, agreement of absolute intensities cannot be discussed since experimental values are not known. Relative intensity of C=O stretching (region above 1600 cm<sup>-1</sup>; see Figure 2) is strongly underestimated by the calculation and also rather high calculated intensity around 1200 cm<sup>-1</sup> is lacking in experiment. Since these vibrations involve movement of polar or charged groups, the solvent models may not be good enough to describe nonelectrostatic solvent influence. The two strongest experimental peaks (assigned to the CH<sub>3</sub> scissoring modes 44–47 and the delocalized modes 23–24, respectively) as well as the overall intensity pattern are reasonably well reproduced by the simulation. The relatively high intensity below 800 cm<sup>-1</sup>,

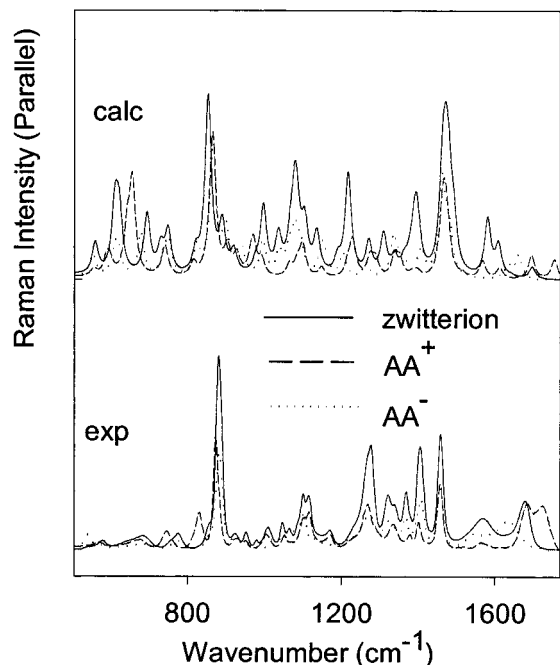
unseen experimentally, can be attributed to a tight dependence of vibrational frequencies on molecular conformation for these modes.

**Polarized Raman Spectra for AA<sup>±</sup>, AA<sup>-</sup>, and AA<sup>+</sup>.** The sensitivity of ROA on various ionic forms of peptides has been shown in several studies previously.<sup>21</sup> It is thus desirable that the simulation reproduces such changes. Additionally, the assignment of zwitterion vibrational modes (Table 4) can be confirmed because of their different behavior for cationic and anionic forms. Polarized parallel Raman spectra (which can be measured with a better resolution than unpolarized) are plotted in Figure 3. Experimental intensities were scaled arbitrarily in order to match approximately the calculated absolute intensity profile, to enhance the visual comparison.

Observed intensities in the highest frequency region (1500–1800 cm<sup>-1</sup>) are most sensitive to the molecular charge. In Table 5, calculated (BPW91/COSMO/6-31G\*) frequencies for the amide I, carboxyl group C=O stretching, and NH<sub>2</sub>/NH<sub>3</sub> hydrogen scissoring modes are compared to experiment. The experimental frequency of the C=O stretching in the CO<sub>2</sub>H group in cation AA<sup>+</sup> is by about 100 cm<sup>-1</sup> bigger than that for CO<sub>2</sub><sup>-</sup> in AA<sup>±</sup> and AA<sup>-</sup>, which could be reproduced by the calculation. The frequency of amide I is seen at 1685 cm<sup>-1</sup> for AA<sup>+</sup>, above that of zwitterion (at 1677 cm<sup>-1</sup>) and anion (1638 cm<sup>-1</sup>). The model reproduces this trend faithfully (1696, 1669, and 1664 cm<sup>-1</sup>, respectively). Also, no significant shift for the



**Figure 2.** Raman spectra of the zwitterion of AA as simulated with the Onsager, ionic, and COSMO solvent models (top) and comparison to the experimental spectrum (bottom). Theoretical spectra are plotted in the same y-scale.



**Figure 3.** Simulated (top) and experimental polarized (parallel, bottom) Raman spectra of the zwitterion, cation, and anion of AA. Experimental spectra of these three forms were obtained for water, HCl, and NaOH solutions, respectively. Theoretical spectra plotted in the same, experimental intensity scaled arbitrarily in order to enhance the comparison.

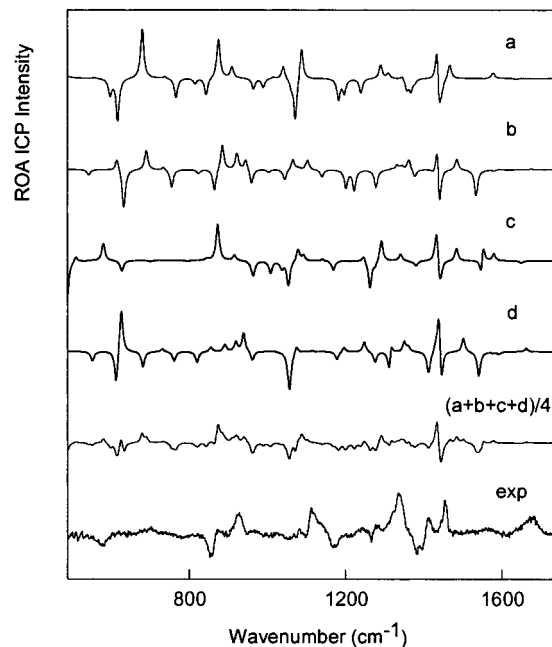
hydrogen scissoring frequencies (experimentally around 1570  $\text{cm}^{-1}$ ), but a drop in their relative intensities, was measured for the  $\text{AA}^-$  and  $\text{AA}^+$  forms, in accordance with the calculation.

Smaller changes can be seen for the lower frequency modes. The C–H bending frequencies (experimentally at 1459, predicted at about 1470  $\text{cm}^{-1}$  for  $\text{AA}^\pm$ ) are insensitive to molecular charge, but the intensity diminishes for the two charged forms. The umbrella  $\text{NH}_3$  mode calculated at 1398  $\text{cm}^{-1}$  for zwitterion (experimentally at 1414  $\text{cm}^{-1}$ ) disappears rather surprisingly

**TABLE 5: Selected Vibrational Frequencies (Calculated/Experimental) in the Three Ionic Forms of AA<sup>a</sup>**

	C=O stretch in $\text{CO}_2$	amide I	$\text{NH}_3$ scissoring
$\text{AA}^\pm$	1639/1640	1669/1677	1576–1596/1572
$\text{AA}^+$	1755/1722	1696/1685	1568–1612/1564–1612
$\text{AA}^-$	1623/1545	1664/1638	1634/1579

<sup>a</sup> BPW91/COSMO/6-31G\* calculation.



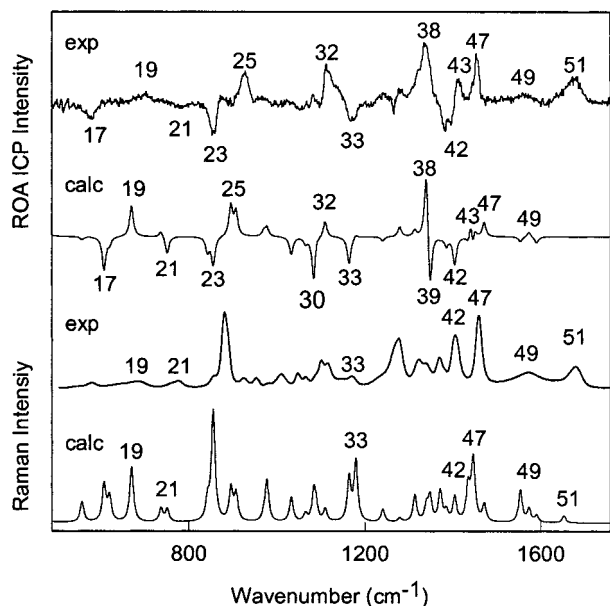
**Figure 4.** Simulated ROA spectra (BPW91/COSMO/6-311G\*\*) of the four lowest energy conformations of AA (cf. Table 1) compared with the experiment.

for both  $\text{AA}^+$  and  $\text{AA}^-$ . The fine splitting and frequency changes below 800  $\text{cm}^{-1}$  are not well reproduced.

**ROA Spectra of AA.** The dependence of ROA spectra on molecular conformation is documented in Figure 4. Simulations (BPW91/COSMO/6-311G\*\* force field) for the four “planar” conformations of AA are compared to the experiment. Although conformation a was selected for the mode assignment in Table 4, many experimental features can be explained by the presence of other conformers. This is namely apparent by broadening of ROA bands (as in the sum  $a + b + c + d$  spectrum) as well as cancelation of ROA intensities in the region below 800  $\text{cm}^{-1}$ . Thus, although this sensitivity may be regarded as an advantage of the ROA spectroscopy, limited accuracy of the calculations does not allow quantitative conformer analysis at the present stage.

In Figure 5 simulated Raman and ROA spectra for the most promising conformation (cf. column 5 of Table 4 for frequencies and Table 6 for intensities) are compared to experiment. Computation of the force field (BPW91/COSMO 80 tes./6-311G\*\*) took 74 h using four CPUs on a V-server (HP 9000/800 240 MHz) and additional 276 h of one CPU were needed on a similar computer (MIPS R10000 195 MHz) for the ROA tensor (HF/6-31G\*) computations with Dalton. Dipolar polarizability  $\alpha$  was computed at the HF/6-31++G\*\* level and combined with tensors  $\mathbf{G}'$  and  $\mathbf{A}$  obtained from Dalton.

As can be seen in Figures 2, 4, and 5, this last computation best matches shapes of experimental Raman and ROA spectra. Most ROA bands were computed with correct signs. The positive signal at 1459  $\text{cm}^{-1}$  arises from the  $\text{CH}_3$  scissoring modes, which contribute with a roughly conservative pattern and a positive signal from the amide II vibration. The negative



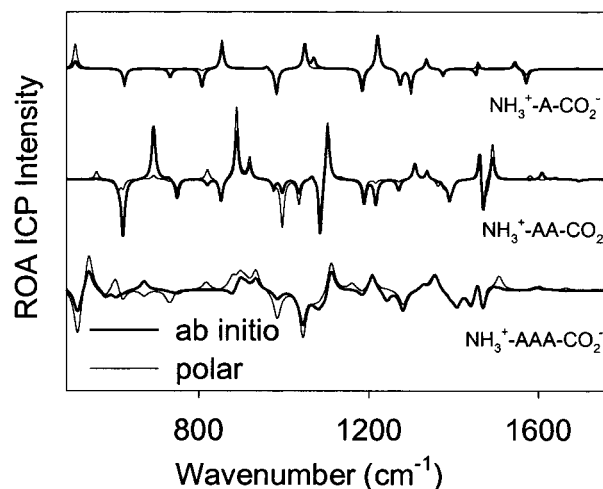
**Figure 5.** Detailed comparison of simulated (BPW91/COSMO/6-311G\*\*, see the normal mode numbering in Tables 4 and 6) and experimental ROA and Raman spectra of AA zwitterion.

**TABLE 6: Calculated Raman ( $I$ ) and ROA ( $\Delta I$ , ICP) Intensities and Depolarization Ratios ( $\sigma$ ) for AA**

mode	$I$	$\sigma$	$\Delta I$	mode	$I$	$\sigma$	$\Delta I$
51	2.4	0.15	0.5	33	8.7	0.49	-35.9
50	2.2	0.75	-12.7	32	2.4	0.56	20.6
49	4.1	0.72	10.5	31	2.3	0.42	12.2
48	9.6	0.58	-9.5	30	5.5	0.66	-56.8
47	4.8	0.21	26.7	28	4.3	0.50	-19.8
46	1.5	0.36	24.8	29	1.4	0.45	-6.9
45	14.9	0.73	-36.0	26	0.9	0.75	5.2
44	1.6	0.72	30.1	27	6.6	0.74	9.9
43	8.8	0.73	-11.7	25	3.9	0.15	23.3
42	6.3	0.73	-46.9	24	4.7	0.32	29.0
41	2.6	0.39	-14.4	23	14.8	0.20	-25.2
40	7.6	0.52	-2.3	22	2.5	0.16	-12.7
39	6.0	0.63	-99.7	21	1.5	0.43	-12.8
38	3.7	0.74	116.6	20	1.5	0.75	4.6
37	6.4	0.24	10.2	19	5.5	0.25	20.1
36	0.9	0.38	14.9	18	2.2	0.54	-3.6
35	2.9	0.73	-6.1	17	3.3	0.18	-19.0
34	12.3	0.30	4.9	16	1.6	0.09	-1.5

experimental signal around 1390 can be reproduced with correct sign as from the  $\text{NH}_3$  and  $\text{CH}_3$  umbrella modes 39–42, as well as the positive peak at  $1338\text{ cm}^{-1}$  assigned to the symmetric  $\text{C}=\text{O}$  stretching and  $^{13}\text{C}-\text{H}$  bending modes. The broadness of the beaks stemming from the terminal groups suggests a conformational equilibrium of more conformations. A negatively biased couplet is predicted for the delocalized ( $\text{C}-\text{N}$  stretch,  $\text{CH}_3$ ,  $\text{NH}_3$  wagging,  $\text{C}-\text{H}$  bending) modes 31–32, while a predominantly positive signal was observed. The predicted negative intensity of mode 30 (exp.  $1106\text{ cm}^{-1}$ ) cannot be seen in our experiment, although a weak negative signal was observed previously in this region.<sup>19</sup> The strong experimental  $\pm$  pattern for the complicated skeletal modes 22–25 is well reproduced by the calculation. The assignment for the lower frequency modes 16–21 is somewhat speculative, but the ROA sign pattern well agrees with the observation.

**Dependence on Peptide Length. The Polar Model.** Experimental ROA spectra of alanine, dialanine (AA), and trialanine peptides (AAA) have many common features<sup>19,21</sup> that can be reproduced by simulations. The spectra plotted in Figure 6 were calculated for fully extended conformations on the basis

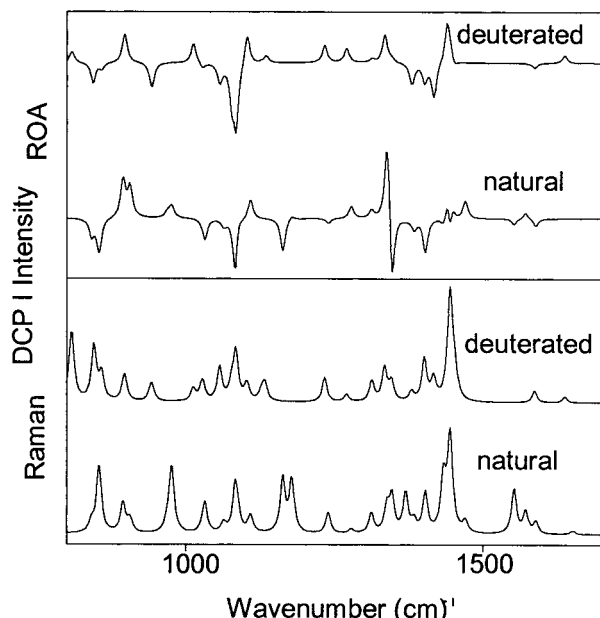


**Figure 6.** Simulated ROA spectra of mono-, di-, and trialanine zwitterions. Dotted lines correspond to the intensities obtained using the polar model (eqs 1–2).

of the BPW91/COSMO/6-31G\* force field. The overall bigger ROA intensity for AAA may suggest a long distance coupling among chromophores as well as additivity of signals from alanine residues. The intensity patterns around  $1470$  and  $1100\text{ cm}^{-1}$  are similar for AA and AAA and unseen in A, in accordance with the experiment. Supposedly, the signal arises from coupling of two vibrating methyl groups, impossible in alanine. On the other hand, the  $\pm$  patterns for the modes involving the  $\text{C}-\text{CO}_2$  and  $\text{C}-\text{NH}_3$  stretches (calculated at  $808/855$  for alanine, experimentally at  $845/918$ ; for AA calcd at  $815/862$ , obsd at  $853/926\text{ cm}^{-1}$ ; for AAA calcd at  $883/908$ , obsd at about  $850/950\text{ cm}^{-1}$ ) are present in the spectra of all three peptides. Thus, the ROA spectrum appears as a rather complicated sum of local and longer-range contributions, both from side chains and amide groups, unlike for peptide VCD governed by longer-range amide interactions.

The spectra simulated by the polar model (drawn with dotted line in Figure 6) reproduce with a high degree of fidelity the full ab initio pattern. This agreement justifies the use of the model as a fast alternative to rigorous computation of the tensors  $\mathbf{A}$  and  $\mathbf{G}'$ , reducing drastically the computational demands, as outlined above. Additionally, tensor  $\alpha$  can be computed analytically and with a much bigger basis set. Obviously, when the frequency error is reduced and intensity accuracy becomes the limiting factor of the calculations, the exact procedure will have to be used. But also in that case, as shown for the ROA spectrum of AA above, the local part of the tensors can be treated as perturbations while the polar part can be obtained more easily.

**Spectra of Isotopic Isomers of AA.** Isotopic substitution, mostly deuteration, can be done relatively easily in the laboratory and provides an indispensable tool for normal mode assignment. In Figure 7, such an effect is simulated for DCP<sub>1</sub> spectra of the AA zwitterion, since it can be compared to the experiment in ref 22. Note that in principle DCP<sub>1</sub> and ICP measurements provide the same ROA but different Raman intensities. For Raman, the deuteration leads to a strong increase of relative intensity of the highest peak (calculated at  $1447$ , experimentally at  $1453\text{ cm}^{-1}$ ) and a loss of intensity in the region above  $1500\text{ cm}^{-1}$ , in accordance with the observation. We explain the former effect by a strong coupling between the amide II and  $\text{CH}_3$  scissoring modes. For ROA, the positive signal around  $900\text{ cm}^{-1}$  (modes 24 and 25, experimentally at  $928\text{ cm}^{-1}$ ) shifts to  $898\text{ cm}^{-1}$ , somewhat underestimating the observed shift to  $915\text{ cm}^{-1}$ . The negative signal calculated at  $1164\text{ cm}^{-1}$  (exp  $1170$ ) for the



**Figure 7.** Comparison of ROA and Raman DCP<sub>1</sub> intensities for AA zwitterion, simulated for the natural and N-deuterated form.

natural form shifts to 1086 cm<sup>-1</sup>. Only a shift by about 20 cm<sup>-1</sup> was seen experimentally for this case, probably because of conformational equilibria. The deuteration also gives a positive signal in the region of 1220–1300 cm<sup>-1</sup> and splits the signal for the CH<sub>3</sub> scissoring modes, all of which were observed, although quantitative intensity agreement is not so satisfactory. We have also conducted trial simulations of deuteration on the  $\alpha$ -carbons of AA, to reproduce Raman studies given in ref 20. The level of agreement was similar to that for the DCP<sub>1</sub> spectra given above. Thus, instead of giving detailed numbers, we can conclude that while the main effects of the isotopic substitutions can be reproduced with the model, an increase in accuracy is desirable in the future. As apparent, this will be a complex task involving many fields of computational chemistry.

## Conclusions

The COSMO solvent model reduces relative conformer energies of the L-alanyl-L-alanine zwitterion by up to 15 kcal/mol, if compared with vacuum computations, and thus increases the conformational flexibility of the molecule. This is in agreement with results observed previously for other systems.

The model in a combination with the BPW91 density functional enabled a realistic computation of vibrational frequencies and Raman and ROA intensities, as well as band-to-band frequency assignment. However, accuracy increase of the calculations is still desirable in order to determine conformational equilibria more precisely.

ROA intensities are extremely sensitive on fine conformational changes, even including the terminal groups. Qualitative trends predicted for the dependence of ROA spectra on the ionic state of AA, isotopic substitutions as well as the dependence of the spectra on peptide chain length agree with those found experimentally.

**Acknowledgment.** The work was supported by the Grant Agency of the Czech Republic (grant numbers 203/97/P002 and 203/01/0031).

## References and Notes

(1) Barron, L. D.; Hecht, L. In *Circular Dichroism*; Nakanishi, K., Berova, N., Woody, R. W., Eds.; VCH Publishers: New York, 1994; pp 179–215.

- (2) Cuperlovic, M.; Meresi, G. H.; Palke, W. E.; Gerig, J. T. *J. Magn. Reson.* **2000**, *142*, 11.
- (3) Constante, J.; Hecht, L.; Polavarapu, P. L.; Collet, A.; Barron, L. D. *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 885.
- (4) Bouř, P.; Baumruk, V.; Hanzliková, J. *Collect. Czech. Chem. Commun.* **1997**, *9*, 1384.
- (5) Bouř, P. *Chem. Phys. Lett.* **1998**, *288*, 363–370.
- (6) Bouř, P.; Tam, C. N.; Shaharuzzaman, M.; Chickos, J. S.; Keiderling, T. A. *J. Phys. Chem.* **1996**, *100*, 15041.
- (7) Bouř, P.; Keiderling, T. A. *J. Am. Chem. Soc.* **1993**, *115*, 9602.
- (8) Bouř, P.; Kubelka, J.; Keiderling, T. A. *Biopolymers* **2000**, *53*, 380.
- (9) Bouř, P. *J. Comput. Chem.* **2001**, *22*, 426.
- (10) Barron, L. D.; Wilson, G.; Hecht, L.; Blanch, E. W. In *Spectroscopy of Biological Molecules: Modern Trends*; Carmona, P., Navarro, R., Hernanz, A., Eds.; Kluwer: Dordrecht, The Netherlands, 1997; pp 9–10.
- (11) Klamt, A.; Schhrmann, G. *J. Chem. Soc., Perkin Trans. 2* **1993**, *2*, 799.
- (12) Klamt, A. *J. Phys. Chem.* **1995**, *99*, 2224.
- (13) Barone, V.; Cossi, M.; Tomasi, J. *J. Comput. Chem.* **1998**, *19*, 404.
- (14) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revisions A.3 and A.7; Gaussian, Inc.: Pittsburgh, PA, 1998.
- (15) Klamt, A. In *The Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., III., Schreiner, P. R., Eds.; John Wiley & Sons: Chichester, U.K., 1998; pp 604–615.
- (16) Bell, A. F.; Barron, L. D.; Hecht, L. *Carbohydr. Res.* **1994**, *25*, 11.
- (17) Bell, A. F.; Hecht, L.; Barron, L. D. *J. Am. Chem. Soc.* **1998**, *120*, 5820.
- (18) Barron, L. D.; Gargaro, A. R.; Hecht, L.; Polavarapu, P. L. *Spectrochim. Acta* **1991**, *47A*, 1001.
- (19) Yu, G.; Freedman, T. B.; Nafie, L. A.; Deng, Z.; Polavarapu, P. L. *J. Phys. Chem.* **1995**, *99*, 835.
- (20) Oboodi, M. R.; Alva, C.; Diem, M. *J. Phys. Chem.* **1984**, *88*, 501.
- (21) Ford, S. J.; Wen, Z. Q.; Hecht, L.; Barron, L. *Biopolymers* **1994**, *34*, 303.
- (22) You, G. S.; Che, D.; Freedman, T. B.; Nafie, L. A. *Biospectroscopy* **1995**, *1* (2), 113.
- (23) Deng, Z.; Polavarapu, P. L.; Ford, S. J.; Hecht, L.; Barron, L. D.; Ewig, C. S.; Jalkanen, K. *J. Phys. Chem.* **1996**, *100*, 2025.
- (24) Becke, A. D. *Phys. Rev. A* **1998**, *38*, 3098.
- (25) Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 13244.
- (26) Bouř, P.; McCann, J.; Wieser, H. *J. Phys. Chem. A* **1998**, *102*, 102.
- (27) Bouř, P.; Sopková, J.; Bednářová, L.; Maloň, P.; Keiderling, T. A. *J. Comput. Chem.* **1997**, *18*, 646.
- (28) Gonçalves, P. F. B.; Livotto, P. R. *Chem. Phys. Lett.* **1999**, *304*, 438.
- (29) Bouř, P.; Král, V. *Collect. Czech. Chem. Commun.* **2000**, *65* (5), 631.
- (30) Bohr, H. G.; Jalkanen, K. J.; Elstner, M.; Frimand, K.; Suhai, S. *Chem. Phys.* **1999**, *246*, 13.
- (31) Freedman, T. B.; Nafie, L. A. *J. Chem. Phys.* **1983**, *78*, 27; **1983**, *79*, 1104.
- (32) Barron, L. D. *Molecular Light Scattering and Optical Activity*; Cambridge University Press: Cambridge, U.K., 1982.
- (33) Helgaker, T.; Jensen, H. J. A.; Jørgensen, P.; Olsen, J.; Ruud, K.; Ågren, H.; Andersen, T.; Bak, K. L.; Bakken, V.; Christiansen, O.; Dahle, P.; Dalskov, E. K.; Enevoldsen, T.; Fernandez, B.; Heiberg, H.; Hettema, H.; Jonsson, D.; Kirpekar, S.; Kobayashi, R.; Koch, H.; Mikkelsen, K. V.; Norman, P.; Packer, M. J.; Saue, T.; Taylor, P. R.; Vahtras, O. *Dalton*, Release 1.0, 1997.
- (34) Amos, R. D. Cadpac 5, 1990, Darresbury Laboratory, Darresbury (in-house modified version of the program).
- (35) Helgaker, T.; Ruud, K.; Bak, K. L.; Jørgensen, P.; Olsen, J. *Faraday Discuss.* **1995**, *99*, 121.
- (36) Mohammady, K. M.; Jalkanen, K. J.; Nardi, F.; Wade, R. C.; Suhai, S. *Chem. Phys.* **1999**, *240*, 63.
- (37) Nielsen, P. A.; Norrby, P. O.; Liljefors, T.; Rega, N.; Barone, V. *J. Am. Chem. Soc.* **2000**, *122*, 3151.