Time Dependent Density Functional Theory Modeling of Chiroptical Properties of Small Amino Acids in Solution

Matthew D. Kundrat and Jochen Autschbach*

Department of Chemistry, 312 Natural Sciences Complex, The State University of New York at Buffalo, Buffalo, New York 14260-3000

Received: November 18, 2005; In Final Form: January 12, 2006

Time dependent density functional theory (TDDFT) and the conductor-like screening model (COSMO) of solvation were used to model the specific rotation and optical rotatory dispersion (ORD) of alanine, proline and serine solutions. Zwitterionic, cationic and anionic forms of amino acids were investigated and the results compared with experimental literature data obtained in neutral, acidic and basic conditions, respectively. It was found that TDDFT consistently underestimated the electronic excitation energies of the molecules, leading to calculated optical rotations that are of the correct sign but somewhat larger in magnitude than those of experiment. An additional challenge was encountered in the modeling of serine, an amino acid with a strong tendency to form intramolecular hydrogen bonds. The model used overestimated the extent of such hydrogen bonding for the zwitterions while possibly underestimating such bonding for the cationic form. This effect on the calculated mole fractions of the different conformers had an impact on the specific rotation.

Introduction

Any molecule whose structure lacks a plane of symmetry, center of inversion or any other improper axis of rotation is chiral and can be made as two stereoisomers, enantiomers that are mirror images of one another. Such enantiomers and identically prepared solutions thereof will rotate a plane of polarized light in equal an opposite directions. However, without additional information, the sign of the rotation alone is not enough to assign the absolute configuration of an enantiomer. Computational chemistry can provide this information.[1−4] Early computational benchmarking on modeling of optical rotation has been performed on molecules in the absence of solvent. From a computing perspective gas phase measurements are the easiest to model. However, experimentally, most measurements are carried out in solution. This complicates the work of the computational chemist, as solvent effects may have a significant impact on the observed optical rotatory dispersion of a solution. Recent advances in this field have been made toward modeling chiroptical properties under the influence of solvent.[5,6] Improvements in the accuracy of these modeling techniques have prompted an increasing number of chemists to rely on computational methods to assign absolute configurations.[7−10] Confidence in the computational methods has grown to the point that it has been used to "correct" older experimentally derived absolute configurations.[11]

Most of the molecules studied thus far can be found in approximately the same geometries in solution as in the gas phase. With such molecules one may save computational costs by optimizing the molecule in the gas phase and then treating solvent effects on the energy and response properties afterward. But not all molecules are so well behaved. Some molecules adopt significantly differing geometries in vapor and in solution. In the most difficult cases the solution phase geometry is not even stable in the gas phase. The amino acids fall into this category of molecules, forming zwitterions in aqueous solution but reverting to their neutral form upon evaporation.

In 2004 Pecul et al. conducted a study of the conformational effects on the optical activity of two amino acids in the gas phase.[12] When comparing these calculations to experimental data gathered from aqueous solutions, they found that the specific rotations calculated at the Hartree−Fock (HF) level of theory were closer to experiment than those obtained at the density functional theory (DFT) level despite the fact that DFT should and generally does yield optical rotations that are closer to experiment.[13] In particular, the authors found DFT to be "clearly unreliable" for the proline anion. The zwitterions were not studied. The authors did note that amino acids tend to be found in zwitterionic form in neutral aqueous solution and that the bulk of the optical activity caused by amino acid solutions is attributable to these forms. However, this issue was not investigated further.

Part of the purpose of our paper is to demonstrate that hybrid-DFT calculations employing large diffuse basis sets and an appropriate solvent model can provide reasonable optical rotation values for neutral, acidic and basic amino acids solutions. First we will test our computational model on glycine, an achiral amino acid that we know should have an optical rotation of exactly zero. Next we will study the optical rotation of solutions of the smallest chiral amino acid alanine, which we will use to show some of the merits and shortcomings of DFT modeling of optical activity and show how through a cancellation of errors sometimes HF results can be closer to experiment than those of DFT when one looks at an optical rotation measured at a single frequency instead of a range of frequencies. Then we will model the solution phase optical activity of the slightly larger and more conformationally complex proline, where we will begin to discuss the importance of Boltzmann-averaging optical rotations from different conformers to obtain results that best agree with those of experiment. Finally, we will attempt to extend our model to serine, an amino acid whose −OH functional group makes it especially prone to intramolecular hydrogen bonding, and we will show the challenges such interactions pose to our current method of modeling chiroptical response properties.

* Corresponding author. E-mail: jochena@buffalo.edu.

Computational Methods

All data were computed with the Turbomole[14] quantum chemical software, version 5.7. Except where otherwise noted all calculations were performed with B3-LYP[15] hybrid functional. All molecular geometries were optimized with the aug-cc-pVDZ basis set from the Turbomole library. All response calculations were with the d-aug-cc-pVDZ set,[16] which previously has been shown to work quite well for TDDFT calculations of optical rotations.[17]

All optimizations and response calculations were performed with the conductor-like screening model (COSMO)[18] of solvation applied to the ground state. Solvent model parameters were configured using the cosmoprep program of the Turbomole package. The dielectric constant of the solvent was set to 78; all other solvent parameters were left at program default values. Default atomic radii (Bondi radii x 1.17)[19] were used.

Initial geometrical parameters were set using the Molden[20] graphical interface program and its default parameters. First the alanine zwitterion structure was drawn and optimized; afterward Molden was used to modify this template into various conformations of the other amino acids that were subsequently optimized. Cationic, anionic and neutral amino acid structures were derived from their corresponding optimized zwitterionic structures by adding, removing or changing the location of a hydrogen atom as appropriate followed by re-optimization. No symmetry restrictions could be imposed on the structure of glycine during optimization because the Turbomole 5.7 code does not support symmetry with the COSMO solvent model.

Accuracy limitations[21] of COSMO apparently kept some of our geometries from meeting our convergence criteria during optimization with the doubly augmented basis set; thus all reported geometries and zero-point energies were calculated with the singly augmented basis. All structures were confirmed to be local minima having no imaginary vibrational frequencies, as calculated with the NumForce program. This numerical method of frequency computation was used because the analytical frequency module of the software was incompatible with the COSMO solvation method. In this paper the energy, "D", of a particular conformer is defined as the sum of the electronic energy, the solvation energy from the COSMO model and the zero point energy calculated by NumForce. "ΔD" is defined as the energy of a particular conformer relative to the lowest energy conformation. Boltzmann factors were calculated on the basis of this relative energy at the temperature of 293 K. Different conformations of the same molecule generally had zero point energies and solvation energies that were within $1-2$ kJ/mol of one another.

Because the Turbomole 5.7 code does not calculate optical rotations on the basis of gauge including atomic orbitals (GIAOs, also referred to as London Atomic Orbitals) or another distributed gauge-origin method, strictly speaking all calculated optical rotations are gauge origin dependent. Our gauge origin is defined as the center of mass in each molecule. Such gauge-origin dependency is known to diminish as the basis set size increases, and from a practical standpoint, reasonably reliable results for small molecules are obtained by using the large augmented basis sets that are always needed to calculate reliable optical rotations.[9,22] Data in the literature support this conclusion. Using the aug-cc-pVDZ basis set, Pecul et al. reported that for proline in its neutral and cationic form, optical rotations varied by less than 5 deg·cm^3/(g·dm) when computed with and without GIAOs.[12] Earlier Ruud and Helgaker used the larger d-aug-cc-pVTZ basis set to ensure "near gauge-origin independence" for their non-GIAO calculations.[23] As we use the d-aug-cc-pVDZ
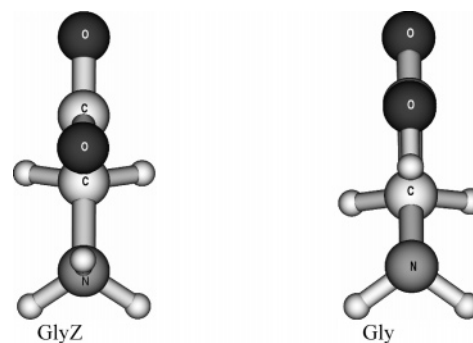

Figure 1. Glycine neutral and zwitterionic forms optimized at the B3LYP + COSMO/aug-cc-pVDZ level of theory. The mirror plane is perpendicular to page.

basis here, we can infer that the gauge-origin dependence on our computed optical rotations should be small compared to other errors inherit in the calculations.

Except where otherwise noted specific rotations were calculated at the wavelength of the sodium D line (589.3 nm). All specific rotations are reported in units of deg·cm^3/(g·dm). Computed optical rotatory dispersion (ORD) curves were calculated at 10 nm intervals from 600 to 220 nm. Experimental ORD plots were scanned from their respective graphics in the literature, digitized using the WinDIG program,[24] then converted from molecular rotation to specific rotation, and plotted alongside the calculated curves.

Results and Discussion

Glycine. Glycine is the smallest genetically encoded amino acid, and as the only achiral one, it is the only one with a vanishing optical rotation. Much computational work has been published on glycine, and a good part of this work was directed toward correctly predicting the stability of the glycine zwitterion in solution. For example, Jensen and Gordon have published the results of computational studies on the stabilization of the glycine zwitterion with explicit water molecules, what is sometimes referred to as a "discrete solvent model".[25] However, without a thorough isotropic sampling of many solvent−solute configurations, an arbitrary addition of solvent molecules around even an achiral solute can lead to an asymmetrical system that exhibits a significant optical rotation. Continuum solvent models[26] such as COSMO do not have this drawback.

We performed some calculations with the COSMO solvation method to determine if it could stabilize the glycine zwitterion without the need for the addition of explicit water molecules. These attempts were successful, and we obtained the glycine zwitterion structure shown on the left side of Figure 1. By this method the zwitterionic form of glycine is calculated to be more stable than the neutral form on the right by a ΔD of 4.6kJ/mol. This differs by nearly an order of magnitude from the experimental value of 30.4 kJ/mol.[27] Several years ago Tortonda et al. obtained similar results in modeling neutral and zwitterionic glycine with an ellipsoidal cavity continuum solvent model, but nonetheless, they succeeded in calculating an aqueous glycine infrared absorbance spectrum that was comparable to that of experiment.[28] Likewise, for the purposes of the present paper, any error in the neutral−zwitterion energy difference does not necessarily have a profound effect on the validity of the computed response properties. All that is needed is a method to calculate a reasonable zwitterionic amino acid structure that is energetically stable. The COSMO model appears to fulfill this requirement.
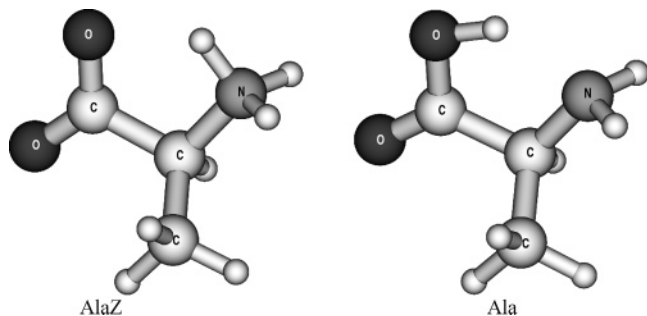
Chiroptical Properties of Small Amino Acids

J. Phys. Chem. A, Vol. 110, No. 11, 2006 **4117**

Figure 2. Optimized structure of alanine zwitterion, and the neutral form it reverts to in the absence of solvent effects.

Solutions of glycine do not exhibit optical rotation. From it may be concluded that either (1) the glycine molecule has a structure with a plane of symmetry or (2) there is a fast exchange between degenerate chiral conformers. The former case was found to be true as the optimized structure of the glycine zwitterion molecule converges to C_s symmetry, illustrated in Figure 1.

As the convergence criteria are tightened, it is apparent that this molecule would converge to a structure that has a mirror plane. Unfortunately, the imprecision inherent in the way that the COSMO model is implemented prevented the use of the tight geometry convergence criteria that one would like to use for the optimization of molecules. With the DFT grid tightened to "m5", the SCF convergence tightened to 10^{-7}, and the COSMO number of geometrical segments per atom (NSPA) number increased to 162, the molecular energy fluctuated by about 10^{-5} hartree by the end of the optimization cycle. Because these geometries inevitably did not converge as well as we intended, the optical rotatory data calculated on the basis of those geometries also deviates from ideal. This is most apparent with glycine, which we know should have a specific rotation of zero. The optimized glycine structure we obtained, which to the human eye does appear to be perfectly symmetric, in fact has a O−C−C−N dihedral of 0.147° instead of zero. This slight asymmetry results in a calculated specific rotation of −2.3 deg· cm³/(g·dm). This number is indicative of approximately how much variance should be expected by computing specific rotations in case small numerical errors in the structure due to a solvation model and the DFT integration grid may occur.

Alanine. Alanine serves as the prototype chiral amino acid. Its relatively small number of electrons lends to rapid calculations of structures and response properties. In addition, the fact that alanine has fewer atoms than the other chiral amino acids means that it will have fewer local minimum structures that need to be investigated. In fact, for the alanine zwitterion we found only one minimum, depicted in Figure 2.

For this zwitterionic structure we calculated a specific rotation of +4.0 deg·cm³/(g·dm), which compares reasonably well with the experimental value of +2.42.[29] Among the common amino acids that are optically active, alanine has the smallest specific rotation in its zwitterionic form. Djerassi noted that this small rotation corresponds to the fact that two of the groups attached to the alanine's chiral carbon, −NH₃⁺ and −CH₃, are isoelec-tronic with one another and that the optical activity of the molecule must result from the charge difference of the N and C nuclei and the perturbation this causes to the electronic structure (otherwise the molecule would have a plane of symmetry).[30]

Thus at this initial stage we appear to have correctly calculated the sign of a very small specific rotation of a single molecule at a single wavelength caused by a small electronic perturbation

of an otherwise symmetrical molecule. Without further informa-tion it would be fair to argue that such an agreement between calculation and experiment for a specific rotation so small in magnitude may merely have been the result of good fortune. However, the comparatively large error margins reported in ref 3 and related studies do not necessarily contradict our results. From ref 3 and previous computational work on optical rotation[13,17,22] it is also clear that a subset of molecules with small optical rotations would have an absolute average error much smaller than the average error of a larger set of molecules that spans a range of optical rotations over several orders of magnitude. To gain more confidence in the computational results for a molecule with a very small optical rotation at 589.3 nm, one could, for instance, compare the optical rotation of a molecule at various wavelengths closer to the excitation energy where the optical rotation is much greater to see if the relative (percent) deviation between theory and experiment is ap-proximately of the same magnitude. Fortunately, some experi-mental ORD plots for various amino acid solutions can be found in the literature, although those experiments were performed at a pH of one where alanine is found predominantly in its cationic (protonated) form.

To facilitate comparison with the experiment, the optimized geometries of the alanine cation were calculated. Protonation of the alanine zwitterion occurs at one of the oxygen atoms of the COO⁻ group. This reduces the site symmetry of the group from C_{2v} to C_s and makes possible a greater number of local minimum conformations for the cation than for the zwitterion. Three such minima were found here. They are depicted in Figure 3.

Structure I is predicted to dominate at room temperature, with a Boltzmann weight of nearly 97%. As can be seen in Table 1, the contribution to the Boltzmann-averaged specific rotation from the two higher energy conformers is negligible. As such, for the rest of this work the only cationic amino acid conforma-tions that will be considered will be those analogous to structure I, the structures with the additional proton attached to the oxygen most distant from the amino group and orientated between the two carboxylate oxygen atoms.

Figure 4 shows that the shape of the ORD for alanine in acidic solution has been faithfully reproduced. A comparison of experimental measured optical rotation to our calculated value at 589.3 nm would show that the calculated value is greater in magnitude than that of experiment; TDDFT is known to overestimate the magnitude of ORD with respect to both experiment and CCSD calculated results.[31] Such a deviation could arise from an overestimation of the magnitude of the rotatory strengths, an underestimation of the electronic excitation energies, or a combination of both.

What may also be concluded from the ORD that may not be determined from single wavelength calculations is that the calculated optical rotations are not merely too high in magnitude, but in fact, the entire calculated ORD curve appears to be red-shifted compared to the experiment. This error also seems to be characteristic of B3LYP in general, as it is known to yield electronic excitation energies that are somewhat lower than experiment.[32] Conversely, the Hartree−Fock method is known to overestimate the electronic excitation energies, which results in a calculated ORD curve that is blue-shifted.[8] The dependence of the calculated lowest excitation energy on the computational method is shown in Table 2. If the only error occurring were the underestimation of the electronic transition energies, chang-ing the functional from DFT to HF should make the magnitude of the calculated specific rotation smaller than that of experi-
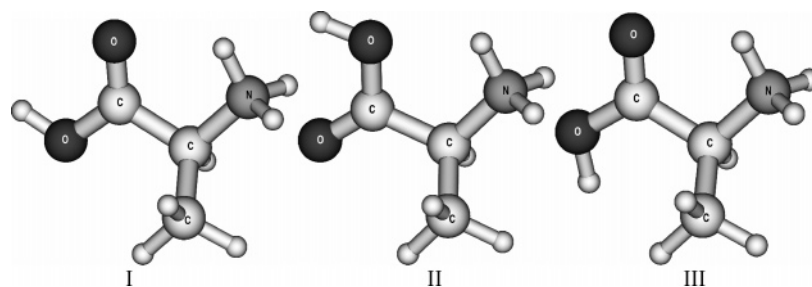
Figure 3. Optimized local minimum structures of the alanine cation.

TABLE 1: Relative Energies and Specific Rotations of Alanine Cation Conformers B3LYP/d-aug-cc-pVDZ//B3LYP/aug-cc-pVDZ, Boltzmann Factors Calculated from ΔD at 293.15 K[a]

conformer	ΔD (kJ/mol)	Boltzmann factor	specific rotation (deg·cm^3/(g·dm))	Boltzmann factor × specific rotation
I	0.0	0.97	34.6	33.4
II	8.3	0.03	−17.3	−0.5
III	12.5	0.01	22.3	0.1
Boltzmann av specific rotation				33.0
experimental specific rotation				13.7

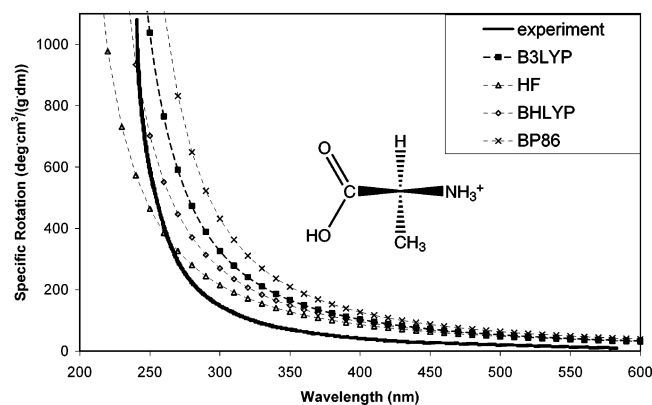[a] Experimental data are from ref 29, pp 203.

Figure 4. Specific rotation of cationic alanine as a function of wavelength. Methods used include Hartree−Fock (HF), Becke "half-and-half" LYP (BHLYP) and Becke−Perdew 86. Experimental data are from ref 30, pp 217.

ment. Because this is not the case for the alanine cation, it is apparent that some other error is present in the calculations that is causing the magnitude of the specific rotation to be overestimated, and changing to an uncorrelated method just partially cancels this error. It is not our aim to "tune" the fraction of Hartree−Fock exchange in the hybrid functional for best agreement with experiment, as the B3LYP functional is already known to generally yield quite accurate linear response properties. We are simply pointing out how a cancellation of errors can sometimes make HF results appear closer to experiment if one only considers the rotation at 589.3 nm instead of modeling the ORD over a larger range of wavelengths. Because the B3LYP functional will be used throughout the rest of this work, one must keep in mind that these predictable errors are likely to be present in the calculations of the more complex amino acids as well.

Proline. Next to alanine and the achiral glycine, proline is arguably the next easiest amino acid to model. Proline is ubiquitously referred to as an "amino acid" due to its biological role though it is sometimes termed an "imino acid", because the nitrogen atom, bound in a five-membered ring, is bonded to one less hydrogen atom than is the case for the other amino

TABLE 2: Calculated Lowest Excitation Wavelength and Specific Rotation for Alanine Geometries at B3LYP/aug-cc-pVDZ, Response Calculations with d-aug-cc-pVDZ Basis

		HF	BHLYP	B3LYP	BP86
cation	excitation wavelength (nm)	177.2	195.0	206.4	213.7
	specific rotation at 589.3 nm	32.6	34.1	34.6	40.9
zwitterion	excitation wavelength (nm)	168.5[a]	192.3	215.6	239.2
	specific rotation at 589.3 nm	−1.2	−2.5	3.9	18.0

[a] The sign of rotatory strength for lowest energy transition does not match the sign calculated with the density functional methods.

acids. The unique ring moiety formed by one nitrogen and four carbon atoms limits the conformational space of proline, thus limiting the number of calculations that must be performed to model its optical rotation.

For the proline zwitterion only two conformations were found. These conformers differ by the direction of puckering of the five-membered pyrrolidine ring and can be identified by whether the γ carbon of the ring is bent toward the carboxylate group (endo) or away from the carboxylate group (exo), depicted in Figure 5. It is not completely clear from the literature whether the endo or the exo configuration is energetically favored at neutral pH.

In 1978 Jankowski et al. performed semiempirical calculations on proline based on the Karplus equation and NMR data and found a ratio of endo to exo conformations of 63:37 at a pH of 7.2.[33] A few years later Haasnoot and co-workers used a similar method to deduce that these proline zwitterions have a 50:50 mole fraction in solution, and thus the ΔG between the two was precisely zero.[34] Stepanian et al. calculated that the in the vapor phase the neutral endo conformer is favored by 2.0 kJ/mol over the exo at the CCSD(T)/6-31++G** level of theory.[35] Pecul and co-workers performed gas-phase B3LYP/aug-cc-pVDZ calculations on proline and found that the lowest energy endo conformer was energetically favored over the exo by 1.63 kJ/mol in the neutral form and 0.84 kJ/mol in the protonated form.[12] Most recently, Cappelli et al. calculated the two conformers to be less than 0.1 kJ/mol apart using the B3LYP/6-31+G(d) method and the IEF-PCM solvent model, with the exo form actually becoming favored over the endo form by around 3.95 kJ/mol when three explicit water molecules are added to the model.[36]

The results of our calculations are collected in Table 3. For the zwitterions, they indicate that the endo conformer may be slightly favored in the zwitterionic form, by a ΔD of 1.6 kJ/mol. With only two conformers that are close in energy, even a small change in energy can alter the ratio of Boltzmann populations considerably. An average computed on the basis of our calculated 66:34 ratio, which closely agrees with Jankowski's semiempirical ratio, results in a specific rotation of −115.4 deg·cm^3/(g·dm) for the zwitterionic form of proline. A 50:50 average suggested by Haasnoot's experimental data gives an average specific rotation of −128.8 deg·cm^3/(g·dm),

Chiroptical Properties of Small Amino Acids

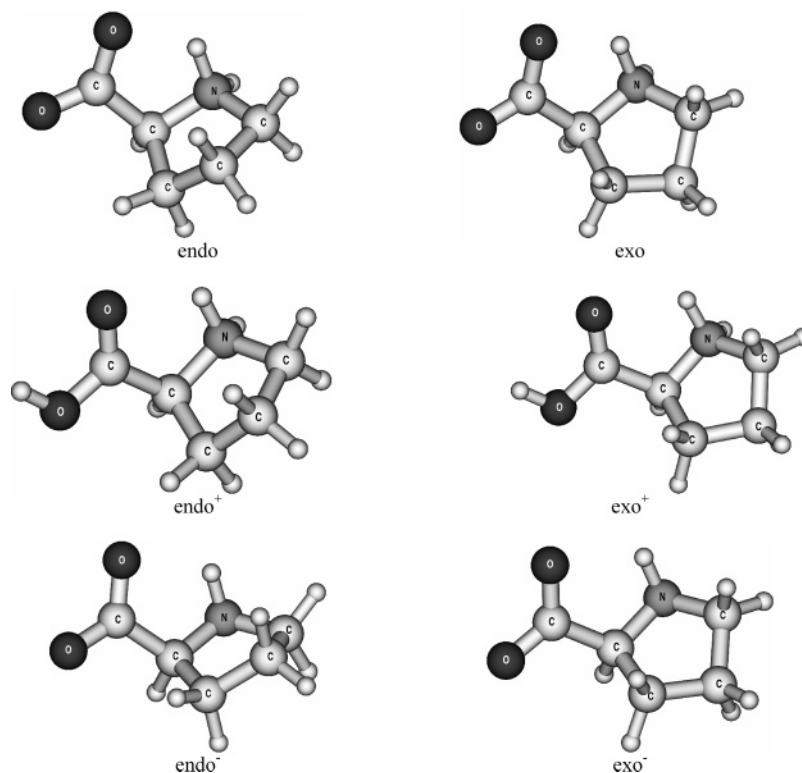*J. Phys. Chem. A, Vol. 110, No. 11, 2006* **4119**

Figure 5. Low lying optimized structures of the proline zwitterion, cation and anion. A pair of higher energy anions with the imino hydrogen on the opposite side of the ring, as well as two pair of higher lying cation conformers with the COOH groups in configurations analogous to alanine structures II and III were also found, but because they did not have significant Boltzmann populations at room temperature, they are not shown here.

TABLE 3: Relative Energies and Specific Rotations of Proline Conformers B3LYP/d-aug-cc-pVDZ//B3LYP/aug-cc-pVDZ, Boltzmann Factors Calculated from ΔD at 293.15 K[a]

conformer	ΔD (kJ/mol)	Boltzmann factor	specific rotation (deg·cm^3/g·dm)	Boltzmann factor × specific rotation
Zwitterions				
endo	0.0	0.66	−86.7	−57.1
exo	1.6	0.34	−171.0	−58.3
Boltzmann av specific rotation				−115.4
experimental specific rotation				−85.0
Cations				
exo$^+$	0.0	0.54	−108.8	−59.1
endo$^+$	0.4	0.46	−46.4	−21.2
Boltzmann av specific rotation				−80.3
experimental specific rotation				−52.6
Anions				
endo$^-$	0.0	0.56	−74.3	−41.8
exo$^-$	0.6	0.44	−243.3	−106.5
Boltzmann av specific rotation				−148.3
experimental specific rotation				−93

[a] Experimental data are from ref 29, pp 7967.

which is actually further from the experimental value of −85.0.[29] Regardless of the Boltzmann populations used here, the product is always a calculated specific rotation that is of the correct sign and somewhat higher in magnitude than the experimentally measured value. This could be expected on the basis of previous experience with optical rotation calculations using the B3LYP functional, including our results for alanine.

Our calculations on the proline cation showed two conformers that are so close in energy that a conclusive assignment of the ground state is not possible, although it appears that the exo form is slightly favored over the endo by 0.4 kJ/mol, which yields a 54:46 exo-to-endo ratio. Experimentally derived data agree
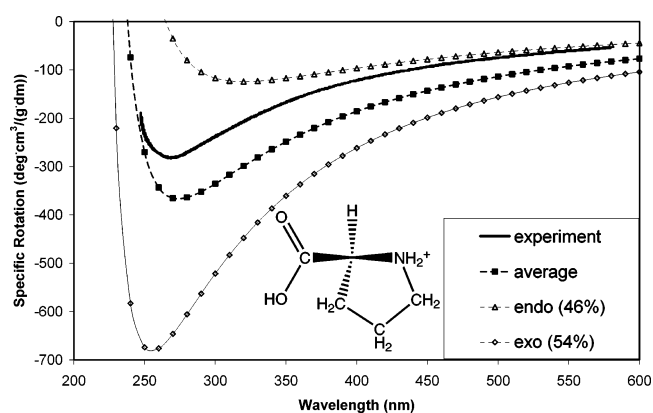

Figure 6. Specific rotation of cationic proline as a function of wavelength Experimental data are from ref 30, pp 222.

that the endo and exo forms are equally populated at low pH.[34,37] Cappelli et al. calculated a slightly higher energy difference that results in a 72:28 exo-to-endo ratio with the IEF-PCM model.[36] Pecul et al. calculated that the lowest energy endo conformation is favored over the exo by 0.84 kJ/mol in gas phase.[12] As was the case with the zwitterionic form, there is disagreement in the literature regarding the exo-to-endo ratio. Again our average specific rotation calculated for the proline cation is of the same sign as the experimental value regardless of the Boltzmann factors used since there is no cancellation of ORs of opposite sign. More insight can be gained by comparing the calculated ORD to experiment.

As can be seen in Figure 6, the ORD of the endo conformer is consistently lower in magnitude than that of experiment, and the ORD from the exo conformer is higher. Our Boltzmann averaged ORD provides a better fit to experiment than the exo or endo conformers alone and better fits the shape of the ORD,
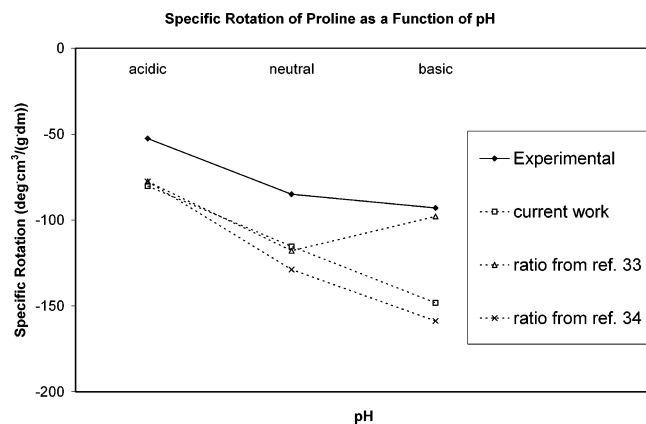
Figure 7. pH effect on optical rotation data for proline. Experimental values from ref 29, pp 7967. Experimentally derived exo/endo ratios used to compute averages are from refs 33 and 34.

including the trough around 270 nm. Generally, we see the same consistent deviation from experiment we saw with alanine: the calculated values are more intense and the calculated ORD is red shifted.

At wavelengths lower than about 250 nm the dependence of the optical rotation on the wavelength becomes very strong, and the numerical agreement with experiment worsens. This is a predictable result of the way in which the ORD is calculated. The computational code used in this work does not take into account the finite lifetimes of the electronic states. As a result, the ORD curves for the endo and exo conformers each will exhibit a singularity as they approach their excitation energies, calculated to be at 207.2 and 208.3 nm, respectively, with positive rotatory strengths.

Experimentally, one would expect the typical bisignate ORD close to 200 nm corresponding to a positive Cotton effect. On the other hand, the trough at 270 nm as well as the long wavelength limit of the optical rotation must be dominated by high lying excitations with negative rotatory strength and are thus expected to be reproduced well by computations without damping[45] such as performed here.

Specific rotation data at high pH are also available in the literature for proline. Consequently, calculations on anionic proline in solution were performed to compare those results with experimental values. Four low lying minimum structures were found differing from one another by exo/endo ring puckering and by whether the imino proton is on the same or opposing side of the ring as the carboxylate group. The two structures with the imino proton cis to the carboxylate group were both calculated to be about 12 kJ/mol lower in energy than the lowest lying trans configuration; therefore only data from these two lowest lying conformers are reported hereafter.

Of these two structures the endo configuration is calculated to be favored over the exo by 0.6 kJ/mol, yielding a 56:44 endo-to-exo ratio at room temperature. Jankowski et al. derived from NMR measurements at a pH of 12.7 that this endo-to-exo ratio should be 86:14.[33] Haasnoot disputed the validity of the trend cited by Jankowski that the endo configuration becomes more favored at higher pH and instead asserts that a 50:50 ratio is equally valid at all pH values.[34] Our energy calculations by themselves do not have the high level of precision that is needed to conclude which experimentally derived ratio is more correct. However, some insight may be gained by comparing the effect these differing proposed exo/endo ratios have on the average specific rotation.

When the optical rotations of the exo and endo forms of proline in various ionization states are averaged on the basis of
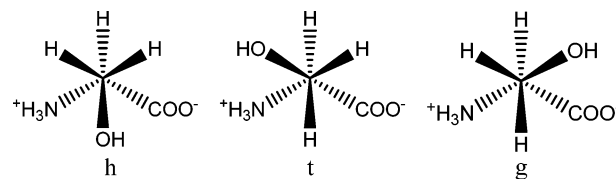

Figure 8. Primary rotamers of serine; page is perpendicular to the $C_\beta-C_\alpha$ bond.

the populations calculated in this work, a trend of increasingly negative specific rotation with increasing pH can be seen (Figure 7). This is in keeping with experiment. If the experimentally derived 50:50 ratio[34] is used to average our data from the respective conformers, this trend still holds true, although a somewhat larger deviation from experiment is seen for the more alkaline solutions. However, if the experimental ratios suggesting that the endo form becomes much more populated at high pH[33] are used to average our optical rotation data, then the specific rotation of proline is calculated to become less negative with increasing pH, and this is inconsistent with experiment.

In the absolute sense, the average specific rotations calculated in this work deviate more for the anionic form than for other forms of proline. Recently, Pecul et al. modeled the specific rotation of anionic proline with TDDFT and found the results unacceptable. Their calculated optical rotation was based almost entirely upon what in this work is termed the exo⁻ conformer, which has an optical rotation that is much higher in magnitude than that of the lowest energy endo⁻ conformer. Unfortunately, ref 12 did not consider data from the endo⁻ conformer. As a result, their calculated specific rotation was several times higher in magnitude than that of experiment. On the basis of this calculated optical rotation, the authors concluded that, "for the anionic form of proline, the DFT values are clearly unreliable because the excitation energies are strongly underestimated, probably because of the DFT self-interaction problem."[12]

The results in this work do not lead to the same conclusion. It is true that as one proceeds from modeling cations to modeling anions, the lowest excitation energies decrease. Thus in a first approximation one might expect the relative errors in the optical rotation calculations to increase along the cation/zwitterion/anion series. However, with a large diffuse basis set and modeling of solvent effects, and on the basis of a conformational search that includes the endo⁻ conformer, our data indicate that the specific rotation of the proline anion can be modeled about as well as that of the cation or zwitterion. The data in Table 3 show that the calculated specific rotations for the proline cation and anion both are of the correct sign and exceed their experimentally measured counterparts by just over 50% in magnitude. The underestimation of the excitation energy seems to affect the modeling of cationic, zwitterionic and anionic proline's specific rotation relatively equally.

Serine. Serine can be considered a derivative of alanine. In terms of the number of atoms and molecular weight it is the smallest of the chiral amino acids after alanine and is thus attractive from a computational point of view. However, in contrast to proline with its atoms bound in a ring, all of the bonds in serine are relatively free to rotate. This gives rise to a number of possible conformations. The primary rotamers, formed by rotation about the $C_\alpha-C_\beta$ bond are depicted in Figure 8.

The naming system is adopted from the works of Marten et al.[38,39] The "t" and "g" refer to situations where the *largest groups* are trans and gauche to one another; this differs from the definitions used in other naming conventions.[40] The "h" designation apparently stems from the word "hindered", which
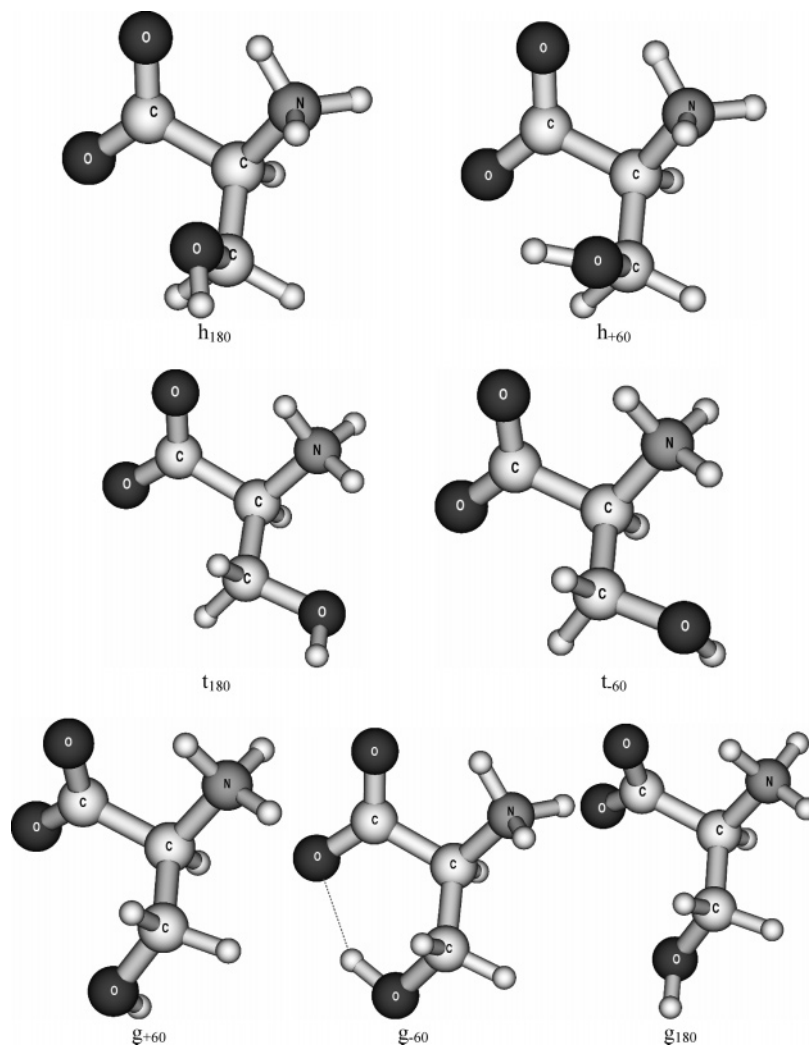
Figure 9. Optimized structures of the serine zwitterion. Structure g_{-60} depicts the strongly hydrogen bonded conformation with an O$-$H distance = 1.87 Å.

was assigned at a time when it was presumed that the configuration with all of the largest functional groups adjacent to one another would be the least favored conformation. Ironically, experiments have since shown that the h configuration is in fact the most energetically favored conformation for serine, followed by the t and the g.[41,42]

For each of the primary rotamers depicted in Figure 7 there can be three possible subrotamers due a rotation of the $-$OH group. This could give rise to up to nine local minima. In actuality, only seven were found, because one each of the h and t rotamer structures did not converge to a minimum due to a steric interaction between the $-$NH$_3^+$ and the $-$OH groups. The optimized structures are depicted in Figure 9. The subrotamers are designated as +60, $-$60 and 180, which refers to the approximate value of the H$-$O$-$C$_\beta$$-C_\alpha$ dihedral angle.

Our calculated Boltzmann factors generally follow the trend that the total mole fraction of the h rotamers exceeds that of the t's, which in turn exceeds that of the g's. However, this agreement is not as good as one would like it to be. By far the biggest source of error has been the intramolecular hydrogen bonding within the g rotamers.

In modeling the average specific rotation of serine, the unusually low energy of one conformer alone was sufficient to disturb our Boltzmann averaging. This structure is depicted as "g$_{-60}$" in Figure 9. It contains an intramolecular hydrogen bond between the $-$OH and the $-$COO$^-$ groups. This so stabilized

this configuration that this structure is predicted to be the overall ground state for serine. It is the energy of this structure that caused the computed Boltzmann population of the g rotamers to be more than twice the value derived from Noszal's experiments, a value in line with other experimentally derived data.[41] Given this consistency and the assertion that such NMR coupling constant derived rotamer populations should error by no more than ±8%,[43] we may conclude that our calculated g rotamer population is indeed significantly higher than it should be, and that it is an overestimation of the extent of intramolecular hydrogen bonding that caused this deviation.

With proline such a deviation between computed and mole fractions of conformers would not have affected the validity of the computed sign of the specific rotation, because all of the relevant conformers had the same specific rotation sign. However, with molecules that can adopt conformations with differing signs of large-magnitude optical rotation changes in calculated conformational populations can alter the sign of the calculated average specific rotation if the resultant value is small. Consequently, errors in the calculated energies can alter the sign of the calculated specific rotation, which is what occurred with our serine zwitterion calculations.

The average specific rotation calculated for zwitterionic serine is +6.3 deg·cm^3/(g·dm), whereas the experimental value from the literature is $-$6.83. The deviation from experiment in the calculated mole fractions (Boltzmann populations) explains this

TABLE 4: Boltzmann Populations of Serine Rotamers B3LYP/d-aug-cc-pVDZ//B3LYP/aug-cc-pVDZ, Calculated from ΔD at 293.15K

		h	T	g
serine zwitterion	computed	0.47	0.27	0.26[a]
	experimental	0.61	0.27	0.11
serine cation	computed	0.79	0.21	<0.01
	experimental	0.76	0.16	0.08

[a] Almost exclusively from the "hydrogen bonded" configuration, g_{-60}. Experimentally derived populations are from Nozsal et al.[42]

TABLE 5: Relative Energies and Specific Rotations of Serine Conformers B3LYP/d-aug-cc-pVDZ//B3LYP/aug-cc-pVDZ, Boltzmann Factors Calculated from ΔD at 293.15 K[a]

conformer	ΔD (kJ/mol)	Boltzmann factor	specific rotation (deg·cm^3/(g·dm))	Boltzmann factor × specific rotation
		Zwitterions		
g_{-60}	0.0	0.26	56.3	14.8
h_{180}	0.2	0.24	10.8	2.6
h_{+60}	0.4	0.23	3.5	0.8
t_{180}	1.5	0.14	−28.3	−3.9
t_{-60}	1.8	0.13	−64.0	−8.0
g_{+60}	18.4	<0.01	144.6	0.0
g_{180}	19.9	<0.01	87.7	0.0
Boltzmann av specific rotation				6.3
experimental specific rotation				−6.83
		Cations		
h_{180}^+	0.0	0.70	33.5	23.5
t_{-60}^+	4.1	0.13	−21.9	−2.8
h_{+60}^+	5.0	0.09	35.0	3.0
t_{180}^+	5.0	0.08	−17.3	−1.5
g_{-60}^+	15.7	<0.01	83.9	0.1
g_{+60}^+	18.3	<0.01	198.2	0.1
g_{180}^+	19.1	<0.01	109.6	0.0
Boltzmann av specific rotation				22.4
experimental specific rotation				14.95

[a] Experimental data are from the ref 29, pp 8605.

error. Table 4 shows that for the serine zwitterion we calculate a population of the g rotamer that is more than twice that calculated from experiment. This inflated g population comes at the expense of the population of the t rotamers, which is therefore calculated to be significantly lower than that of experiment. The g rotamers are calculated to give a large positive optical rotation, +56.3 deg·cm^3/(g·dm). The h rotamers contribute a much smaller positive rotation, +7.2 deg·cm^3/(g·dm). If our computed Boltzmann factors of 0.26 for g and 0.47 for t are used, the resulting contributions to the average, +14.8 deg·cm^3/(g·dm) for the g and +3.4 for the h, added together are greater than the −11.9 deg·cm^3/(g·dm) contribution from the levorotatory t rotamers and cause the average calculated specific rotation to be positive. However, when Noszal's experimentally derived Boltzmann populations were imposed on our results, the g rotamers then contribute only 6.2 deg·cm^3/(g·dm), the contribution from the h rotamers would increase to 4.4 deg·cm^3/(g·dm), and with these added to the −11.9 deg·cm^3/(g·dm) from the t rotamers yields an average specific rotation of −1.3 deg·cm^3/(g·dm), which agrees in sign with the experimentally measured value. These data in Table 5 show that the optical rotation of several of the conformers were an order of magnitude larger than the experimental average. Thus an accurate average is difficult to obtain.

The opposite problem arises when the cations of serine are studied: the method *underestimates* the stability of the g cationic rotamers. However, the correct sign for the optical rotation of the serine cation is still calculated at all frequencies, due to a
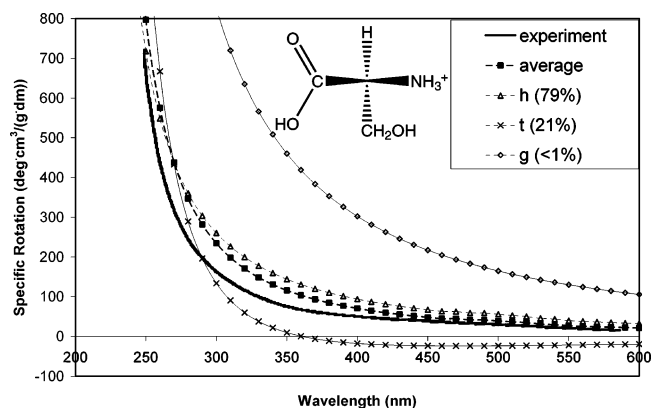

Figure 10. Specific rotation of cationic serine as a function of wavelength Experimental data are from ref 30, pp 217. Plots labeled "h", "t" and "g" are each Boltzmann averages of the optical rotations calculated for the two or three respective subrotamers.

cancellation of errors. As can be seen in Figure 10, if the g population had been as high as it should have been, the Boltzmann average plot would be more in keeping with that of alanine and proline; that is, the calculated ORD would be too large in magnitude and red shifted.

Further insight may be gained by comparing the relative energies of our serine cation conformers with the calculations in the literature. Noguera and co-workers obtained the same relative energies for the primary rotamers as were obtained here in agreement with experiment: h is more stable than t, which is more stable than g.[44] Their calculations did not consider solvent effects, and they found the lowest lying g conformer to be nearly 36.8 kJ/mol higher in energy than the h ground state. Our calculations indicate this difference to be 15.7 kJ/mol. It must be noted that the authors of ref 44 did not include some of the low energy local minimum geometries caused by rotation of the −OH group. However, they reported the structure that we calculate to be the most stable g conformer, which is far too high in energy to be consistent with the population of the g rotamer observed by experiment. From this it may be safely concluded that it is not the solvent model that causes the deviations seen in our relative energies. Instead, those deviations come from approximations that would already affect gas-phase DFT calculations. The COSMO model appears to be partially compensating for such error, though not enough to bring our calculations in line with experiment.

Conclusions

Time-dependent density functional theory response calculations have been shown capable to reproduce the optical rotatory dispersion of cationic, zwitterionic and anionic amino acids in solution with similar margins of error. The method has been used to correctly reproduce the sign of the optical rotation of small amino acid conformers, and to faithfully reproduce the dispersion of the optical rotation of such molecules. A consistent source of error for this method appears to be the underestimation of the electronic excitation energies, which leads to an overestimation of the specific rotation of the molecules at 589 nm and an overall red-shift of their ORD curves.

The usefulness of this method is limited by the ability to predict correct relative energies and the resulting Boltzmann factors for the accessible geometric conformations. The COSMO model seems to work reasonably well for this purpose, but only in cases where intramolecular hydrogen bonding is not a major factor. When intramolecular hydrogen bonding comes into play, DFT coupled with COSMO tends to overestimate the magnitude

Chiroptical Properties of Small Amino Acids

J. Phys. Chem. A, Vol. 110, No. 11, 2006 **4123**

of the bonding in the zwitterions and underestimates it in the cations. If the overall optical rotation is determined by a cancellation of large contributions from individual conformers, the relative errors in the energies can lead to comparatively large deviations from the experimental optical rotation. It is likely that some explicit solvation is necessary to obtain the correct balance between intramolecular and solvent−solute hydrogen bonding.

References and Notes

(1) Polavarapu, P. L. *Chirality* **2002**, *14*, 768.

(2) Kondru, R. K.; Wipf, P.; Beratan, D. N. *Science* **1998**, *282*, 2247.

(3) Stephens, P. J.; McCann, D. M.; Cheeseman, J. R.; Frisch, M. J. *Chirality* **2005**, *17*, S52.

(4) Crawford, T. D. *Theor. Chem. Acc.*, in press.

(5) Mennucci, B.; Tomasi, J.; Cammi, R.; Cheeseman, J. R.; Frisch, M. J.; Devlin, F. J.; Gabriel, S.; Stephens, P. J. *J. Phys. Chem. A* **2002**, *106*, 6102.

(6) Marchesan, D.; Coriani, S.; Forzato, C.; Nitti, P.; Pitacco, G.; Ruud, K. *J. Phys. Chem. A* **2005**, *109*, 1449.

(7) Specht, K. M.; Nam, J.; Ho, D. M.; Berova, N.; Kondru, R. K.; Beratan, D. N.; Wipf, P.; Pascal, R. A.; Kahne, D. *J. Am. Chem. Soc.* **2001**, *123*, 8961.

(8) Giorgio, E.; Viglione, R. G.; Zanasi, R.; Rosini, C. *J. Am. Chem. Soc.* **2004**, *126*, 12968.

(9) Lattanzi, A.; Viglione, R. G.; Scettri, A.; Zanasi, R. *J. Phys. Chem. A* **2004**, *108*, 10749.

(10) da Silva, C. O.; Mennucci, B.; Vreven, T. *J. Org. Chem.* **2004**, *69*, 8161.

(11) Stephens, P. J.; Devlin, F. J.; Cheeseman, J. R.; Frisch, M. J.; Bortolini, O.; Besse, P. *Chirality* **2003**, *15*, S57.

(12) Pecul, M.; Ruud, K.; Rizzo, A.; Helgaker, T. *J. Phys. Chem. A* **2004**, *108*, 4269.

(13) Stephens, P. J.; Devlin, F. J.; Cheeseman, J. R.; Frisch, M. J. *J. Phys. Chem. A* **2001**, *105*, 5356.

(14) Ahlrichs, R.; Bar, M.; Haser, M.; Horn, H.; Kolmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165. Turbomole, Ver. 5.7. Quantum Chemistry Group, University of Karlsruhe. See also http://www.cosmologic.de/turbmole.html.

(15) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(16) PNNL. EMSL Gaussian Basis Set Order Form; Basis sets were obtained from the Extensible Computational Chemistry Environment Basis Set Database, Version 02/25/04, as developed and distributed by the Molecular Science Computing Facility, Environmental and Molecular Sciences Laboratory, which is part of the Pacific Northwest Laboratory, P.O. Box 999, Richland, WA 99352, and funded by the U.S. Department of Energy. The Pacific Northwest Laboratory is a multiprogram laboratory operated by Battelle Memorial Institute for the U.S. Department of Energy under contract DE-AC06-76RLO 1830. Contact Karen Schuchardt for further information.

(17) Grimme, S.; Furche, F.; Ahlrichs, R. *Chem. Phys. Lett.* **2002**, *361*, 321.

(18) Schafer, A.; Klamt, A.; Sattel, D.; Lohrenz, J. C. W.; Eckert, F. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2187.

(19) Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441.

(20) Schaftenaar, G. Molden, 4.4 ed.; Centre for Molecular and Biomolecular Informatics, 2005.

(21) Bour, P.; Kapitan, J.; Baumruk, V. *J. Phys. Chem. A* **2001**, *105*, 6362.

(22) Autschbach, J.; Patchkovskii, S.; Ziegler, T.; van Gisbergen, S. J. A.; Baerends, E. J. *J. Chem. Phys.* **2002**, *117*, 581.

(23) Ruud, K.; Helgaker, T. *Chem. Phys. Lett.* **2002**, *352*, 533.

(24) Lovy, D. WinDIG; Ver. 2.5 http://www.unige.ch/sciences/chifi/cpb/windig/html.

(25) Jensen, J. H.; Gordon, M. S. *J. Am. Chem. Soc.* **1995**, *117*, 8159.

(26) Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999.

(27) Wada, G.; Tamura, E.; Okina, M.; Nakamura, M. *Bull. Chem. Soc. Jpn.* **1982**, *55*, 3064.

(28) Tortonda, F. R.; Pascual-Ahuir, J. L.; Silla, E.; Tunon, I.; Ramirez, F. J. *J. Chem. Phys.* **1998**, *109*, 592.

(29) *The Merck Index*, 12th ed.; Merck & Co., Inc.: Whitehouse Station, NJ, 1996.

(30) Djerassi, C. *Optical Rotatory Dispersion Applications to Organic Chemistry*; McGraw-Hill: New York, 1960.

(31) Crawford, T. D.; Owens, L. S.; Tam, M. C.; Schreiner, P. R.; Koch, H. *J. Am. Chem. Soc.* **2005**, *127*, 1368.

(32) Bauernschmitt, R.; Ahlrichs, R. *Chem. Phys. Lett.* **1996**, *256*, 454.

(33) Jankowski, K.; Soler, F.; Ellenberger, M. *J. Mol. Struct.* **1978**, *48*, 63.

(34) Haasnoot, C. A. G.; Deleeuw, F.; Deleeuw, H. P. M.; Altona, C. *Biopolymers* **1981**, *20*, 1211.

(35) Stepanian, S. G.; Reva, I. D.; Radchenko, E. D.; Adamowicz, L. *J. Phys. Chem. A* **2001**, *105*, 10664.

(36) Cappelli, C.; Monti, S.; Rizzo, A. *Int. J. Quantum Chem.* **2005**, *104*, 744.

(37) Pogliani, L.; Ellenberger, M.; Valat, J. *Org. Magn. Reson.* **1975**, *7*, 61.

(38) Martin, R. B.; Mathur, R. *J. Am. Chem. Soc.* **1965**, *87*, 1065.

(39) Martin, J. P. C. a. R. B. *J. Am. Chem. Soc.* **1972**, *94*, 6141.

(40) Kraszni, M.; Szakacs, Z.; Noszal, B. *Anal. Bioanal. Chem.* **2004**, *378*, 1449.

(41) Fujiwara, S.; Ishizuka, H.; Fudano, S. *Chem. Lett.* **1974**, 1281.

(42) Noszal, B.; Guo, W.; Rabenstein, D. L. *J. Phys. Chem.* **1991**, *95*, 9609.

(43) Feeney, J. *J. Magn. Reson.* **1976**, *21*, 473.

(44) Noguera, M.; Rodriguez-Santiago, L.; Sodupe, M.; Bertran, J. *J. Mol. Struct. (THEOCHEM)* **2001**, *537*, 307.

(45) Autschbach, J.; Jensen, L.; Schatz, G. C.; Tse, Y. C. E.; Krykunov, M. *J. Phys. Chem. A* **2006**, *110*, 2461−2473.