Ab Initio Simulations of the Vibrational Circular Dichroism of Coupled Peptides

Petru Bour and Timothy A. Keiderling*

Contribution from the University of Illinois at Chicago, Department of Chemistry.
Box 4348, Chicago, Illinois 60680

Received October 19, 1992

Abstract: The vibrational circular dichroism (VCD) of peptides in different conformations was simulated using the magnetic field perturbation (MFP) model for the model dipeptide CH$_3$CONH-CH$_2$CONH-CH$_3$. The geometry was optimized in the 4-31G basis set with the torsion angles constrained to mimic the $\alpha$-helical, $\beta$-sheet, 3$\psi$-helical, and polyproline II conformations. An effective mass was used for the appropriate C$_6$ position to simulate the local chirality of L-alanine residues. The normal modes and VCD spectra were calculated at the ab initio quantum mechanical level with the same basis set using analytical derivatives (CADPAC) and the MFP formulations of Stephens and co-workers. These calculations gave simulated VCD spectra that were quite comparable to experimental results with respect to the VCD sign pattern and relative intensities of both the amide I and II modes for polypeptides having these conformations. The MFP results are further shown to be superior to a similar calculation based only on dipole coupling, by comparison to experimental polypeptide spectra.

Introduction

A significant application of vibrational circular dichroism (VCD) for structural studies has developed in the area of peptide and protein conformational analysis. The primary approach to analysis of these spectra, to date, has centered on empirical correlations of VCD spectra for systems of similar secondary structure. Each major secondary structural type yields a distinctive VCD band shape for the amide I and II bands (C=O stretch and C-N-H deformation plus C-N stretch, respectively). These systematic spectral patterns lead to facile predictions of the VCD band shape and a regression analysis based on the resulting components has been developed which yields reasonable predictions of the fractional contribution of the major secondary structure types to the total protein conformation.

Ideally one would like to be able to predict the VCD band shape from a theoretical model based on the peptide geometry.

If such a calculational scheme were shown to be reliable, simulation of spectra based on various probable structures, determined by modeling, for example, could be used for structure determination. Schellman and co-workers and, more recently, Diem and co-workers have attempted to simulate oligopeptide VCD using the simple coupled oscillator or dipole coupling (DC) model. While the DC model appeared to predict the $\alpha$-helical amide I VCD correctly, it failed for the $\beta$-sheet amide I VCD band shape, for the $\alpha$-helical amide II modes, and (as will be shown) for the polyproline II (Pro II) helical amide I VCD.

On the other hand, the magnetic field perturbation (MFP) model as formulated by Stephens and co-workers and implemented at the ab initio level has proven to be quite successful in reproducing the VCD of many small molecules. Similar success has been evident in calculations using other ab initio quantum mechanical approaches to VCD simulation. Part of the key to the success of all of these calculations is the use of ab initio quantum mechanical calculations to determine the quantum mechanical wave functions in the electronic ground state.

initio quantum mechanical force fields to determine the proper relative ordering of transitions. Use of the MFP method has been previously restricted to small molecules due to the size of the basis set needed to obtain the most accurate level of calculation. However, tests in our laboratory with smaller basis sets indicate that the MFP-predicted spectra, particularly for the more intense features, are qualitatively the same as those obtained with the larger basis sets.\(^ {15} \) This observation makes it reasonable to consider ab initio MFP VCD simulation for larger molecules with correspondingly modified goals in terms of precision.

We have recently used the MFP model to delineate the applicability of the coupled oscillator model to strongly coupled, symmetrically equivalent local C–H and C–D stretching modes for a number of small molecules.\(^ {16} \) These calculations demonstrated that although the DC model works well for weakly coupled oscillators, it fails for oscillators nearby groups that are highly polarizable or have delocalized bonding. By contrast, in all cases studied for which experimental data were available, the MFP results were in excellent agreement with the data.

Since the peptide group is an extended π-system and since the peptide groups interact substantially in an oligo- or polypeptide, as evidenced by the well-established conformational sensitivity of their vibrational force fields and frequencies,\(^ {16} \) use of calculations at the ab initio MFP level may be needed for interpretation of peptide VCD. Furthermore, it is expected that such MFP-computed VCD will be a substantial improvement over DC model predictions for these molecules.\(^ {15} \)\(^ {16} \) Due to our computational constraints, an ab initio MFP calculation of the VCD of an extended peptide chain was not realistic, but modeling of an idealized molecule containing two peptide bonds (here termed the “dipeptide”) was possible. Despite the molecular size constraint, the higher energy, amide-centered modes can, in fact, be calculated with some precision and are the sole focus of this report. Since the amide modes are only weakly perturbed by the side chains, these model calculations can be reasonably used to simulate spectra of any sequence and should yield the fundamental amide–amide interactions and their contribution to the VCD band-shape pattern for polypeptides. Previous studies have addressed ab initio conformational energy minimization and force field determinations for dipeptides,\(^ {19} \) but this is the first attempt to simulate their VCD with a high-level theory such as the ab initio MFP.

Even though restricted to the dipeptide level, this representation of near-neighbor amide interaction could eventually be grafted onto a larger VCD calculation that uses a simpler model, such as the DC, to represent longer range amide interactions for which the DC model is applicable.\(^ {16} \) First, however, it is necessary to establish that the MFP model can be successfully used for computation of the characteristic VCD band shapes corresponding to the major polypeptide secondary structure types. That can now be done, and the first such report is the topic of this paper.

### Methods

Our computations were done in two stages. First the molecular geometries of the “glycine-like” model dipeptide, CH$_3$–CONH–CH$_2$–CONH–CH$_3$, was optimized by energy minimization in the 4-31G basis set using the Gaussian 88 program\(^ {20} \) on an IBM 3090 computer. All calculations on the dipeptide were limited to the standard 4-31G basis set by our computer constraints. [Some test calculations on smaller molecules (see below) used alternative, larger basis sets.][10]

### Table I. Geometry Constraints and SCF Energies of the Dipeptide in the Constrained Conformations

<table>
<thead>
<tr>
<th>conformation</th>
<th>ω</th>
<th>φ</th>
<th>ψ</th>
<th>(E(SCF), a.u.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>α-helix</td>
<td>180</td>
<td>-57</td>
<td>-47</td>
<td>-453.1421</td>
</tr>
<tr>
<td>β-sheet</td>
<td>180</td>
<td>-119</td>
<td>113</td>
<td>-453.1475</td>
</tr>
<tr>
<td>310-helix</td>
<td>180</td>
<td>-60</td>
<td>-30</td>
<td>-453.1441</td>
</tr>
<tr>
<td>Pro II</td>
<td>180</td>
<td>-78</td>
<td>-149</td>
<td>-453.1470</td>
</tr>
</tbody>
</table>

The absolute dipole strength changes <14° for these modes is calculated with the three atomic polar and axial tensors \(\rho\) and \(\gamma\) on the terminal amide nitrogen and C=O groups, respectively.

With specific application to this study, a test of rotational strength sensitivity to basis set variation was carried out by selective isotopic substitution of the amide. In these tests, the rotational strengths did vary with basis set, all the computed “intrinsic amide VCD” magnitudes were small. The predicted \(\Delta A_4/\Delta A_6\) values were all \(\leq 10^3\) (or just detectable) for the amide A, I, and II modes.\(^ {22} \) On the primary goal was to compute the conformationally dependent VCD band-shape changes for the amide I and II modes, which generally have relatively intense VCD. On the basis of the results for this single-amide test case,\(^ {22} \) in those cases where the predicted rotational strengths in our calculations for the dipeptide are

\( \Delta A_4/\Delta A_6 \) values obtained were all \(\leq 10^3\) (or just detectable). The absolute values obtained were \(-5\%\) higher than those found experimentally, as is to be expected using SCF-level force fields. Since detailed force field computations are already available for this molecule,\(^ {23} \) extensive discussion is unwarranted here.
Vibrational frequencies were determined. However, the α-sheet without altering the we sought. These results are given for the amide A, I, and I1 modes (each having two components, in and out of phase) in complicate the amide IR and VCD in such a glycine-like is the simplest correction that can improve frequency prediction splitting alone that are commonly used as a basis for DC

Table II. Ab Initio MFP Results for the Highest Amide Modes of the Dipeptide

<table>
<thead>
<tr>
<th>Mode</th>
<th>Glycine-like</th>
<th>Alanine-like</th>
<th>N-Deuterated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\nu$</td>
<td>$\Delta\nu$</td>
<td>$D$</td>
</tr>
<tr>
<td>Ao</td>
<td>3411</td>
<td>-17</td>
<td>38</td>
</tr>
<tr>
<td>Ai</td>
<td>3393</td>
<td>45</td>
<td>-5.8</td>
</tr>
<tr>
<td>Ii</td>
<td>1652</td>
<td>12</td>
<td>763</td>
</tr>
<tr>
<td>Io</td>
<td>1640</td>
<td>190</td>
<td>191</td>
</tr>
<tr>
<td>Ii</td>
<td>1498</td>
<td>-12</td>
<td>610</td>
</tr>
<tr>
<td>Iii</td>
<td>1485</td>
<td>536</td>
<td>-51</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Modes are indicated by their major contribution as o, out of phase, and i, in phase, combinations of the amide A, I, and I1 local oscillators. Column definitions: $\nu$ (cm$^{-1}$), vibrational frequency; $\Delta\nu$ (cm$^{-1}$), splitting between i and o modes; $D$ (10$^{-10}$ D), dipole strength; $R$ (10$^{-10}$ D), rotatory strength.

**Dipeptide Calculations and Results**

In separate geometry optimizations run for each of the dipeptide conformers, all coordinates were varied except the six φ,ψ,ω torsion angles of the main peptide chain, which were fixed to those values shown in Table I. The optimized SCF energies obtained for these four constrained conformations are also included in Table I. For the α-helical and β,δ-helical conformations, 3N–6 positive vibrational frequencies were determined. However, the β-sheet and polyproline II conformations had one and two negative frequencies, respectively.

The force field, atomic polar, and axial tensors and the vibrational spectra were computed with the ab initio program for the "glycine-like" dipeptide (formula as given in previous section). For presentation here, the frequencies were all scaled by a factor of 0.874 to align them better with the experimentally observed frequencies. Such scaling is consistent with the usual experience with SCF force fields and is the simplest correction that can improve frequency prediction without altering the ab initio computed frequency ordering that we sought. These results are given for the amide A, I, and II modes (each having two components, in and out of phase) in columns 2–5 of Table II. The column $\Delta\nu$ indicates the splitting of that amide mode as predicted by the ab initio force field. $\Delta\nu$ is a clear indication of the relative ordering of the vibrational states and can be compared to results obtained from dipolar splitting alone that are commonly used as a basis for DC computations.

Overlap with several C–H deformation modes is expected to complicate the amide IR and VCD in such a glycine-like molecule. Therefore, simulation of a chain of chiral L-alanine residues was accomplished by calculating the VCD of CH$_2$H$_3$H–CONH–CH$_3$H–CONH–CH$_3$H–H. This structure was derived by assigning a mass of 15 to the three hydrogens that would correspond to the side-chain positions in an oligomer of L-amino acid residues and a mass of 30 to the two hydrogens that would correspond to the positions where the main peptide chain would continue in the polymer, as illustrated for the α-helical and β-sheet conformations in Figure 1. [The larger terminal mass was chosen to reduce mixing with the amide modes and to represent the effect of the massive polymeric chain.] Normal mode frequencies as well as dipole and rotational strengths for this alanine-like dipeptide were computed for the amide A, I, and II modes as tabulated in columns 6–9 of Table II. It is important to note that the magnitudes of the rotational strengths for most of the modes in Table II are much larger (by a factor of 4–20) than those found in our test calculations noted above for a single amide. This difference confirms that amide coupling generally leads to much larger VCD intensity than do intrinsic amide chirality effects. Given this, the magnitudes as well as the signs of that we have computed are important because the predicted dipeptide VCD is more reliable for those bands with the largest terminal mass was chosen to reduce mixing with the amide modes and to represent the effect of the massive polymeric chain. Normal mode frequencies as well as dipole and rotational strengths for this alanine-like dipeptide were computed for the amide A, I, and II modes as tabulated in columns 6–9 of Table II. It is important to note that the magnitudes of the rotational strengths for most of the modes in Table II are much larger (by a factor of 4–20) than those found in our test calculations noted above for a single amide. This difference confirms that amide coupling generally leads to much larger VCD intensity than do intrinsic amide chirality effects. Given this, the magnitudes as well as the signs of that we have computed are important because the predicted dipeptide VCD is more reliable for those bands with the largest
The nature of the amide II VCD becomes clear with mass alteration to the alanine-like form, since interference and mixing with other internal coordinates is then reduced and made more realistic. The α-helical amide II VCD is predicted to be strongly negative while the β-sheet amide II VCD is predicted to be a strong negative couplet. The 310-helix is also predicted to have a strong negative amide II, like the α-helix, but in the 310 case the amide II intensity is much larger than the amide I VCD. By contrast, the Pro I helix is predicted to have a weak amide I VCD, much smaller than its amide I VCD.

Overall, the ab initio MFP calculations at the 4-31G level predict the amide I, II region to have a clear pattern of conformationally sensitive sign and intensity patterns, just based on interactions within a dipeptide molecule. The predicted VCD for the amide III region is complex due to mixing of the amide modes with Cα-H deformations. Nonetheless, it can be noted that the amide III VCD signals for the α-helical and 310-helical conformations are predicted to be dominantly positive, while those predicted for the β-sheet and Pro II structures are less clear. In a real peptide molecule, Cα-H deformations arising from the side chains would probably obscure these predicted patterns.

N-deuteration effects were also calculated for the alanine-like dipeptide molecules as listed in Table II, columns 10–13. To simulate the Pro II conformation, a mass of 15 was used for the H(N) to simulate the tertiary amide vibrational characteristics of the proline residue.

Part of the variation observed in the simulated VCD spectra arises from the dependence of the force field on secondary structure. Relative changes in frequency patterns can certainly arise for the isotopic substitutions used to simulate the spectra in Figure 2. Inspection of Table II for the different substitutions shows that, after the "alanine" substitution, the amide A parameters are virtually unchanged from the "glycine" result, which follows from our earlier assertion regarding the small influence of modes far in frequency from the amide modes of interest. On the other hand, the alanine-like amide I frequencies are lower than the glycine-like ones by 1–5 cm⁻¹, and the amide II frequencies are lower by 8–21 cm⁻¹, since those modes are closer in frequency to modes which involve motion of atoms altered by the mass substitution. Amide II intensities (D and R values) are the most sensitive to this substitution, even to the extent of an amide II sign pattern change in the Pro II simulation. Other than for the Pro II structure, which was simulated as two tertiary amides, N-deuteration leaves the amide I and II mode VCD qualitatively the same for the alanine-like computation.

Discussion

Dipeptide Simulation. The results of our MFP calculations have a satisfying stability with regard to the isotopic substitutions made and evidence substantial magnitudes (Table II) in terms of rotational strength as compared to those of the single-amide test calculations. On the basis of our experience with com-
putations of VCD for small molecules, one can have some confidence in the reliability of prediction for bands computed to have large-magnitude \( D \) or \( R \) (absorbance or VCD) values. This is particularly true with regard to basis set effects. However, if these calculations were to exist in isolation, they would do little have large-magnitude \( D \) or \( R \) to the predictions in Figure 2. The calculations exhibit remarkable agreement with those results. L-lysine in \( H_2O \), a 'random-coil' polypeptide that is locally like left-handed helical poly-L-proline II (top to bottom) conformations.

Figure 3. Experimental VCD spectra for (a) albumin in \( H_2O \), a highly \( \alpha \)-helical protein; (b) concanavalin A in \( H_2O \), a highly \( \beta \)-sheet protein; (c) (Alb)\(_2\)Leu(Alb)\(_3\) in CDCl\(_3\), a 3\(_{10}\)-helical oligopeptide; and (d) poly-L-lysine in \( H_2O \), a "random-coil" polypeptide that is locally like left-handed helical poly-L-proline II.

The \( \beta \)-sheet computations are also in good agreement with experimental results. The weak negative amide I VCD reflects what has been seen experimentally in a number of antiparallel \( \beta \)-sheet examples. In addition, the MFP calculations predict a strong amide II negative couplet \( \Delta \lambda \) VCD, again in good agreement with experimental data on models and on proteins. Good models for parallel \( \beta \)-sheets are difficult to find, but protein results indicate that they should at least have negative VCD at the lower frequency end of the amide I range (\( \sim 1630 \text{ cm}^{-1} \)) and negative couplet VCD in the amide II region (Figure 3), which is consistent with these MFP calculations.

One of the most satisfying results was the predicted VCD for the 3\(_{10}\)-helix. Our model peptide work has established that 3\(_{10}\)-helical oligomers yield much weaker VCD of the same sign pattern for the amide I mode and about the same band shape and intensity VCD for the amide A and II modes, as compared to the case of the \( \alpha \)-helix. Our computations reflect this observation very well. In Figure 2, the weak positive couplet nature of the predicted 3\(_{10}\)-helix VCD is not clear, but the Table II values (columns 5, 9, and 13) confirm it. Since the predicted magnitudes are so small for this mode, it is possible that effects other than these near-neighbor interactions will prove to be significant for modeling the VCD of a longer 3\(_{10}\) chain. We have found that other, proposed 3\(_{10}\)-helical geometries give different but still weak amide I VCD while still yielding the same strong amide II VCD. This indicates that these computational results can be taken to be a prediction of weak amide I intensity but should not be viewed as a reliable determination of the specific 3\(_{10}\)-helical amide I VCD band shape.

The Pro II calculations were done with two goals in mind. Experimentally, polyproline II, a left-handed helix of polyalanine bonds, has a strong negative VCD couplet for the amide I mode, which is well matched by the calculated result in Figure 2. In addition, the Pro II conformation is believed to be a good model for the amide I VCD, which is in agreement with results on a-helices in nonaqueous solution, but it is calculated to be biased strongly negative as opposed to the observed positively biased VCD. At this stage, one must regard the amide A theoretical results as being unreliable. Since hydrogen bonding is not included in the model calculation, such an error is understandable but perhaps not truly explained. Finally, the theoretical spectrum in Figure 2 indicates a significant positive VCD in the amide III region, which does agree with experimental results for the \( \alpha \)-helical poly-\( \gamma \)-benzyl-\( \varepsilon \)-glutamate. With experiment: a strong positive VCD couplet \( \Delta \lambda \) for the amide I and a negative VCD for the amide II shifted to the lower energy side of the absorbance maximum. Also in accord with experiment, the negative amide II VCD is here predicted to be maintained upon deuteration (Table II). The predicted amide II VCD couplet is slightly negatively biased, as is usually seen in experiments, but the calculations are not so precise that one should put much stock in that observation. The amide A mode is calculated to have a weak, negative couplet VCD, which is in agreement with results on \( \alpha \)-helices in nonaqueous solution, but it is calculated to be biased strongly negative as opposed to the observed positively biased VCD. At this stage, one must regard the amide A theoretical results as being unreliable. Since hydrogen bonding is not included in the model calculation, such an error is understandable but perhaps not truly explained. Finally, the theoretical spectrum in Figure 2 indicates a significant positive VCD in the amide III region, which does agree with experimental results for the \( \alpha \)-helical poly-\( \gamma \)-benzyl-\( \varepsilon \)-glutamate.
to model the Pro I (right-handed helix of cis peptides) VCD was not as successful, in that while the intensity of the amide I mode was predicted to be low, which is consistent with experiment, the sign pattern was predicted incorrectly.22

Comparison to Dipole Coupling Calculations. To compare the MFP computational method with the previously used simpler and much cheaper (computationally) coupled oscillator approaches to the modeling of VCD spectra, the result of a parallel set of dipole coupling (DC) calculations8-12 is presented in Figure 4. Coupling of only the two amide I and two amide II dipoles was considered. Average frequencies and dipole strengths taken from our single-amide test calculations were used as a basis for the DC calculations. The results of these and other simple model calculations will be discussed separately in detail.22

In summary, as evident from comparison of the spectral simulations based on the DC predictions in Figure 4 with the experimental results in Figure 3, the DC computations fail to reproduce adequately the amide I mode VCD for the β-sheet and Pro II conformations and the amide II VCD for the α-helix, 310-helix, and Pro II helix. They do give good amide II VCD predictions for the β-sheet. However, the amide I VCD predictions for the Pro II helices have the incorrect sign. On this basis alone it is clear that the MFP predictions are superior representations of the oligo- and polypeptide VCD for standard peptide conformations and that the DC predictions are substantially lacking by comparison based on dipeptide computations.

**Conclusion**

The ab initio MFP calculations described here, despite their confinement to just a dipeptide model compound, did better at predicting experimental VCD than have any other models used to date. Given the success of the dipeptide MFP calculations in replicating the observed oligopeptide VCD band shapes for several conformations, one might ask, what is the next step? Our MFP-based calculations only include 1–2 amide–amide interactions, which can be thought of as the intrinsic nearest-neighbor helix term. Their success provides theoretical support to our empirically based proposal, supported by studies of the length dependence of peptide VCD, that VCD has a substantial contribution from short-range effects. The uniformly successful comparison of these predicted VCD spectra, necessarily computed considering only near-neighbor interactions, with experimental data from much longer oligomers and polymers strongly implies that the VCD is dominated by near-neighbor interactions. Extension of these computationally intensive MFP calculations for practical use on polymeric systems may be realized by combining these MFP, pairwise interactions with DC-type computation of longer range interactions. Longer range interactions would presumably lead to weak contributions to the VCD and force field for most structures, but their number may make the total effect significant for extended-chain molecules. This effect will have more importance for those modes which have an MFP-determined, weak near-neighbor contribution to the VCD. We are now pursuing such a hybrid approach to calculation of the VCD for more complex structures utilizing this MFP contribution.

**Acknowledgment.** This work was supported by grants from the National Science Foundation (CHE88-14105) and National Institutes of Health (GM-30147), for which we are most grateful. We wish to thank Dr. R. D. Amos for providing a version of CADPAC compatible with the Titan, Prof. Ron Elber for assistance in implementing it on the UIC facility and Dr. Vladimir Baumruk for experimental data for the figures.