Theoretical UV Circular Dichroism of Aliphatic Cyclic Dipeptides

Kristine L. Carlson, Stephen L. Lowe, Mark R. Hoffmann, and Kathryn A. Thomasson*

Department of Chemistry, University of North Dakota, Grand Forks, North Dakota 58202-9024

Received: July 30, 2004; In Final Form: April 22, 2005

Four cyclic dipeptides (piperazine-2,5-diones), cyclo(L-Pro-Gly), cyclo(L-Pro-L-Leu), cyclo(L-Ala-L-Ala), and cyclo(L-Pro-L-Ala), were modeled from crystal structure data. Conformations resulting from energy minimization using molecular mechanics were compared with traditional ab initio and density functional theory geometric optimizations for each dipeptide. In all computational cases, the gas phase was assumed. The $\pi-\pi^*$ transition feature of the UV circular dichroic (CD) spectra was predicted for each peptide structure via the classical dipole interaction model. The dipole interaction model predicted CD spectra that qualitatively agreed with experiment when MP2 or DFT geometries were used. By coupling MP2 or DFT geometric optimizations with the classical physics method of the dipole interaction model, significantly better CD spectra were calculated than those using geometries obtained by molecular mechanics. Thus, one can couple quantum mechanical geometries with a classical physics model for calculation of circular dichroism.

Introduction

The calculation of electronic spectral phenomena such as UV circular dichroism (CD) is still a major challenge of computational methods. Herein, the classical physics method, the dipole interaction model, is used to calculate CD of a series of aliphatic piperazine-2,5-diones (cyclic dipeptides). The small size of the piperazine-2,5-diones makes them accessible to moderately high-level quantum mechanical calculations through second-order Moller−Plesset perturbation theory (MP2) treatment of electron correlation, and their conformational constraints due to their cyclic nature limits the complexity of their spectra compared to larger molecules. Thus, the effects of molecular geometry on CD spectral calculations via classical physics can be reliably assessed.

Electronic UV circular dichroism is a tool that is still poorly understood theoretically, but is a critical method used to examine secondary structures of proteins. CD measures the difference in absorption of left- and right- circularly polarized light in the absence of a magnetic field. As electrons in the molecule absorb light, the transition of electrons into local excited states induces dipole moments within the molecule. These induced dipole moments interact with one another through both electric and magnetic fields. The amide groups in proteins possess characteristic $\pi-\pi^*$ transitions (180−210 nm, ∼140 nm) and n-π^* transitions (220 nm) whose locations and intensities indicate secondary structural features of the peptide sequence.[1] The piperizine-2,5-diones also exhibit these classical amide transitions, making them excellent models for CD calculations.

The CD spectrum of a molecule is described in terms of the rotational strength, *R*, which is the integrated intensity of the CD band.

$$\boldsymbol{R}(\text{ergs}\cdot\text{cm}^3) = \frac{6909hc}{32\pi^3 N_A}\int\frac{\Delta\epsilon}{\lambda}\,d\lambda = \text{Im}(\vec{\mu}_{oa}\cdot\vec{\text{m}}_{ao}) \qquad (1)$$

The rotational strength (eq 1) arises from interaction of the electric dipole ($\vec{\mu}_{oa}$) and magnetic ($\vec{\text{m}}_{ao}$) dipole and moments of each transition from a ground state to some excited state *a* in some wavelength of applied light, λ. Three parameters describe each CD spectrum: the position of the maximum absorption (ν_{max}), intensity of the absorption ($\Delta\epsilon_{max}$), and the shape of the band. N_A is Avogadro's number; *h* is Planck's constant, and *c* is the speed of light.

There are several methods to predict CD spectra for a molecule with knowledge of the structure. Quantum mechanics allows for direct solution of the dipole and rotational strength, although this is computationally infeasible for large systems. One approach to handle larger molecules requires division of the molecule into a number of separate model chromophores and treating those chromophores quantum mechanically. Coupled with solution of the Schrödinger equation for isolated model chromophores over ground and excited states, this splitting yields the method of Tinoco[2] and the matrix method.[3−5]

Another approach is to calculate CD using classical physics. The dipole interaction model[6−10] is one such classical physics-based method for predicting the CD of peptides and proteins based on the amide chromophore. This model includes all atoms except the amide group as points having nondispersive polarizability, and the amide group as a single point possessing dispersive polarizability. The relationship between the rotational strength, *R*, and the measured difference in absorption of left- and right- circularly polarized light is given by eq 2 for the classical dipole interaction model, assuming a Lorentzian band shape.[8,11−13]

$$\Delta\epsilon = \frac{32\pi^3\bar{\nu}^3 N_A\Gamma}{6909n}\sum_{k=1}^{q}\left\{\frac{R_k}{(\bar{\nu}_k^{\,2}-\bar{\nu}^2)^2+\Gamma^2\bar{\nu}^2}\right\} \qquad (2)$$

In the dipole interaction model, the sum over all dispersive oscillators (light-absorbing units, where there are *q* dispersive oscillators) of the interaction of the rotational strength (R_k) at each wavenumber ($\bar{\nu}$) describes the CD ($\Delta\epsilon$) spectrum. N_A is Avogadro's number, Γ is the half peak bandwidth, *n* is the number of peptide residues, and $\bar{\nu}_k$ is the normal mode

* To whom correspondence should be addressed. E-mail: kthomasson@chem.und.edu. Phone: 701-777-3199. Fax: 701-777-2331.

10.1021/jp046580u CCC: $30.25
Published on Web 06/01/2005

wavenumber. The rotational strength of each segment of the molecule is obtained by dividing the molecule into atoms with isotropic and anisotropic polarizability. The nondispersive oscillators (those with isotropic polarizability) have constant polarizability factors. These atomic polarizabilities were obtained experimentally from fits to molecular polarizabilities of simple organics, beginning with experimental atomic polarizability data determined at the NaD line (589 nm).[11,14] Dispersive oscillators (those whose strength is wavelength-dependent), such as the amide chromophore, have been optimized to reproduce mean polarizabilities and Kerr constants at 589.3 nm for the $\pi-\pi^*$ transition in a variety of simple amides.[15] These original parameters comprise the first classical model to reproduce CD for α-helical structures;[9] however, quantum mechanical treatments of the α-helix did appear earlier.[16,17] More recently, the parameters have been reoptimized to include molecular anisotropies to create three new sets of dipole interaction parameters: (1) general peptide systems (G parameters), (2) α-helical systems (H parameters), and (3) poly-L-Pro-II (J parameters).[10,18,19]

The dipole interaction model has proven successful for a variety of applications, including the prediction of CD spectra for β-sheets,[20] β-turns,[21,22] α-helices,[23,24] β-peptides,[25−27] and is the only method published to obtain the correct $\pi-\pi^*$ spectrum for poly-L-proline II[28,29] and collagen.[30] The model has also proven successful on whole proteins including α-spectrin, tropomyosin,[23] and lactate dehydrogenase.[31] Based on these studies by other groups, we conclude that comparison of theoretical CD for a geometrically optimized structure to the experimental CD reasonably assesses the validity of calculated structures.

Herein the dipole interaction model is tested with geometries predicted for cyclic dipeptides via a variety of molecular modeling techniques including molecular mechanics and quantum mechanics. We address the following questions: (1) What is the sensitivity of the dipole interaction model to small refinements in molecular geometry, and what level of geometric optimization is sufficiently accurate to successfully predict CD spectra of piperazine-2,5-diones? (2) Which of the dipole interaction parameters are best suited for use with cyclic dipeptides? (3) How favorably do the dipole interaction model's predictions compare with experimental values? (4) Does the dipole interaction model recognize poor geometries (i.e., is it a good tool for evaluating molecular geometries)?

Methods

Geometry Optimization. Crystal structures of the cyclic dipeptides were obtained through the Cambridge Structural Database[32] via the ConQuest software.[33] Their CSD codes are: (1) cyclo(L-Ala-D-Ala)[34] (TRDMPP01), (2) cyclo(L-Ala-L-Ala)[34] (LCDMPP01), (3) cyclo(L-Pro-L-Ala)[35] (CLPRAL), (4) cyclo(l-Pro-Gly)[36] (LPROGL01), and (5) cyclo(L-Pro-L-Leu)[37] (PROLEU). These structures were imported into InsightII (Accelyrs, San Diego, CA), where their energy was geometrically minimized using the CVFF force field[38] and a quasi-Newton Raphson algorithm. These calculations were performed on SGI Fuel workstations running Irix 6.5. Geometric optimizations of the structures by energy minimization using quantum mechanical calculations were carried out using Gaussian 98.[39] Pure density functional theory (DFT), hybrid DFT, and traditional ab initio methods were utilized. Specifically restricted Hartree−Fock (RHF), the hybrid DFT Becke3-LYP (B3LYP)[40−42] functional method and two pure DFT functional methods were explored: Becke-VWN (BVWN)[41−43] and Becke−Lee−Yang−Parr (BLYP).[40,41,43] Finally, second-order Moller−Plesset perturbation theory (MP2),[44,45] with frozen core orbitals, was used to investigate the structures of the more flexible peptides. Frequency calculations were carried out at the same level of theory as the geometry optimizations. A series of Pople-style double- and triple-split valence basis sets were used in the calculations at each level: 3-21G,[46−50] 6-31G,[51−56] and 6-311G.[57] Single polarization functions (i.e., d orbitals) were added to the 6-31G[54] and 6-311G[57] basis sets (denoted 6-31G* and 6-311G*), and d and p polarization functions were used with the 6-311G basis set (6-311G**). The 6-31G* calculations used pure d functions (i.e., five functions per set), whereas the triple split calculations used all six Cartesian d functions. The GDIIS algorithm was used with "very tight" geometry optimization convergence criteria. A grid size setting of "ultrafine" (90 radial shells, 590 angular points per shell) was used. For each of these optimizations, SCF convergence was set to 10^{-10}. Solvent effects were not treated in this study because (1) gas-phase calculations have proven insightful in the past with the dipole interaction model;[6,10,18−23,26−29] (2) inclusion of solvent with peptides quantum mechanically has previously proved problematic,[58] and (3) although inclusion of solvent is possible via molecular mechanics, either by a dielectric constant or explicitly at considerably more computational expense, to compare consistently with the quantum mechanics, we chose to treat even the molecular mechanics in the gas phase. Therefore, inclusion of solvent is outside the scope of this paper.

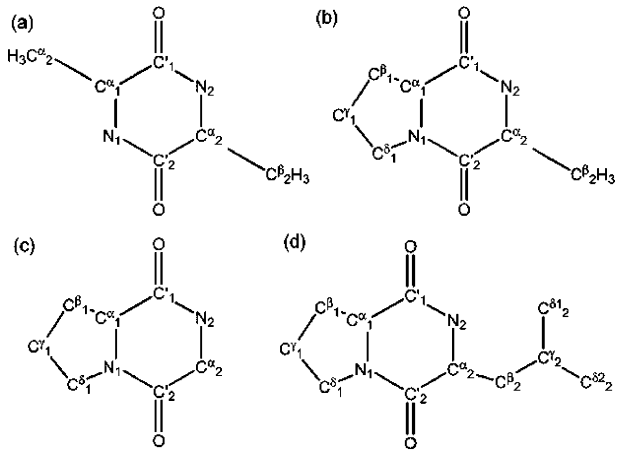

Figure 1. 2D cyclic dipeptide structures: (a) cyclo(L-Ala-L-Ala); (b) cyclo(L-Pro-L-Ala); (c) cyclo(L-Pro-Gly); (d) cyclo(L-Pro-L-Leu).

Structural Comparison. Molecular geometries obtained by both quantum mechanics and molecular mechanics were analyzed with ChemBats3D Pro (CambridgeSoft, Cambridge, MA), and PDB files were generated for each structure using the same software. Values for bond lengths, bond angles, and dihedral angles (Figure 1) were compared for each minimization with crystal structure values (Supporting Information). Standard assignment of dihedral angles are used; for the diketopiperazine ring they are defined as ψ_n (C_k'-N_n-C_n^{α}-C_n'), ϕ_n (N_n-C_n^{α}-C_n'-N_k), ω_n (C_n^{α}-C_n'-N_k-C_k^{α}), where "n" refers to the residue number in the sequence of the dipeptide and "k" is the next residue. The designations of dihedral angles for the proline side chains are $\chi^1{}_n$ (N_n-C_n^{α}-C_n^{β}-C_n^{γ}), $\chi^2{}_n$ (C_n^{α}-C_n^{β}-C_n^{γ}-C_n^{δ}), $\chi^3{}_n$ (C_n^{β}-C_n^{γ}-C_n^{δ}-N_n), $\chi^4{}_n$ (C_n^{γ}-C_n^{δ}-N_n-C_n^{α}), and $\chi^5{}_n$ (C_n^{δ}-N_n-C_n^{α}-C_n^{β}). Dihedral angles of the leucine side chain are similarly referred to as $\chi^1{}_n$ (N_n-C_n^{α}-C_n^{β}-C_n^{γ}), $\chi^2{}_n$ (C_n^{α}-C_n^{β}-C_n^{γ}-$C_n^{\delta 1}$), and $\chi^3{}_n$ (C_n^{α}-C_n^{β}-C_n^{γ}-$C_n^{\delta 2}$). For alanine peptides, the dihedral angle of the side chain refers to the deviation from the plane defined by the amide bond to the next peptide and the alpha carbon of each amino acid, χ^1 (C_n^{β}-C_n^{α}-N_n-C_k'). Pictures of the structures were generated using InsightII (Accelrys, San Diego, CA).

CD Calculation. Cartesian coordinate files generated from the PDB files by InsightII (Accelrys, San Diego, CA) were used to calculate the $\pi-\pi^*$ transition of each optimization's CD spectrum by the dipole interaction model.[13,15,24] This was accomplished through direct use of coordinates for the non-chromophoric portions of each molecule, while the amide (the chromophore) was reduced to a single point and the Eulerian angles between the first chromophore and each successive chromophore were calculated. The original parameters (O),[6] general peptide parameters (G), α-helical systems (H), and poly-L-Pro-II (J parameters)[34] were all used to predict the $\pi-\pi^*$ feature of the CD spectrum for each molecule between 140 and 260 nm with a step size of 0.5 nm. For each of the G-, H-, and J-parameter CD calculations, the location of the amide chromophore was given three possibilities: centered on the N−C′ bond (o), shifted 0.1 Å toward the C′ atom on the N−C′ bond (*x*), and shifted 0.1 Å in the NCO′ plane above the N−C′ bond (*y*). For the original parameter set, only the first location was used because that is what has historically worked best with this model.[6] $\Delta\epsilon$ was calculated every 0.5 cm^{-1} between 150 and 200 nm with bandwidths of 3000, 4000, and 5000 cm^{-1} for each structure. Detailed band analysis (peak location, intensity, and peak ratio) is available in the Supporting Information.

Analysis of CD Calculations. OriginPro 7 (OriginLab Corporation, Northampton, MA) was used to locate the CD spectra peaks and determine half-peak bandwidths, which represent the integrated rotational strength of the combined oscillators. This was accomplished with the Peak-Fitting Module by setting the baseline to $\Delta\epsilon = 0$ and allowing the software to locate peaks automatically. No data preconditioning was used, and all features were fit to Lorentzian bands. A default value of 100 iterations was set for fitting at a 95% confidence value. Published experimental CD spectra were compared with the calculated values for each molecule.

Results

Selection of CD Parameters. Applequist's original parameters gave the best agreement with experiment for every structure calculated for all descriptors of the bands: location of the peak, sign, and half-peak bandwidth (Figure 2). The general parameters (G parameters) predicted an extremely weak band near 180 nm and blue-shifted the band around 205 nm. The poly-L-proline parameters (J parameters) showed the greatest sensitivity to chromophore placement; while the peak locations were typically comparable to those predicted using the original parameters, these parameters sometimes yielded inaccurate signs for the band around 180 nm. The α-helical parameters (H parameters) were also unable to reproduce the experimental CD spectra for any of the cyclic dipeptides tested here. Although the peak locations were relatively accurate, band signs were often incorrect. For all parameters, the ratio of the bandwidth at half-peak for the 205 nm band to the 188 nm band was calculated to have a disproportionately large magnitude compared to the experimental values. The bands, their locations and the ratio of the half-peak bandwidths for all calculations are available in the Supporting Information.

Cyclo(L-Ala-L-Ala). For cyclo(L-Ala-L-Ala), geometry optimization located two minimum energy structures that exhibited either C_2 symmetry or pseudo C_2 symmetry (Figure 3). The first structure (AA-I) exhibited a flat diketopiperazine ring (the central ring formed by the cyclization of the two amino acid backbones) with the methyl "arms" (the side chain methyl groups) raised nearly vertically above the ring. The second structure (AA-II) resembled the crystal structure, possessing a

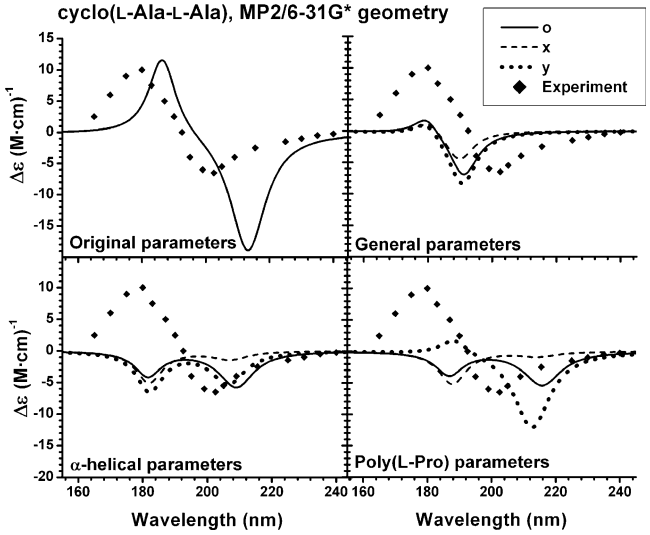

Figure 2. Comparison of different parameter set predictions for the CD spectra of cyclo(L-Ala-L-Ala): the peptide structure was optimized by the MP2 method using a 6-31G* basis set. The "o", "x" and "y" of the legend refers to the position of the pseudoatom for the NC′O chromophore. The "o" positions the pseudoatom halfway between the N and C′ atoms on the N−C′ bond. The "x" indicates a displacement of 0.1 Å toward the C′ atom from the "o" position. The "y" indicates a displacement of 0.1 Å into the N−C′−O plane from the "o" position. Bandwidth for each spectrum is 3000 cm^{-1}. The experimental CD (in water) was obtained from Bowman et al.[61]

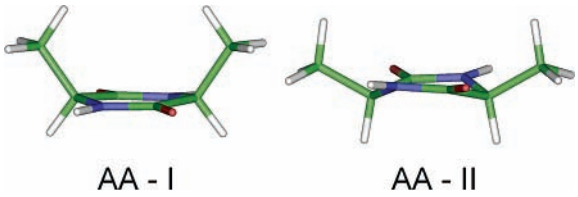

Figure 3. Minimum energy conformations of cyclo(L-Ala-L-Ala). The AA-I structure was obtained by BVWN and RHF optimizations with the larger basis sets from the crystal geometry, and from MP2 optimization starting from the BVWN. The AA-II structure was obtained by minimization of the crystal structure using B3LYP, BLYP, MP2 with all basis sets; and RHF and BVWN with 3-21G and 6-31G* basis sets.

deeply folded diketopiperazine ring while the methyl groups extended outward. Structure AA-I was obtained by BVWN and RHF optimizations with the larger basis sets from the crystal geometry, and from MP2 and CVFF optimization using the BVWN geometry as a starting point. AA-II was obtained by CVFF, B3LYP, BLYP, and MP2 methods starting from the crystal geometry, and by RHF and BVWN with small basis sets. Frequency calculations indicate that both forms are valid energy minima. Analysis of the bond lengths and angles reveals little difference in these properties between the various minimized structures; one key difference, however, is that a majority of the bonds are longer than in the crystal structure; the only bonds that were commonly equal or shorter were the C′−O bonds. The dihedral angles differ greatly between the two conformations (Figure 4). AA-I ϕ values range over 6° to −21°, while ψ varies between −18° and −6°. AA-II has ϕ values between −45° and −26°, and ψ values cover the range 17° to 38°. Furthermore, the MP2 structures have ω torsions that deviate from 0° considerably more than any other method; molecular mechanics and BVWN yield the most planar amide bonds. The potential energies from molecular mechanics and the thermally corrected energies of the various quantum mechanically optimized geometries show a 1.4 kcal/mol dif-

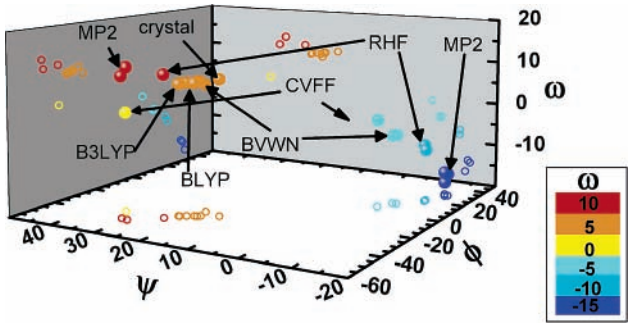

Figure 4. Distribution of ϕ, ψ, and ω angles for cyclo(L-Ala-L-Ala) geometric optimizations. The spheres exist on the three-dimensional location and the open circles represent the two-dimensional shadow on the ω wall. AA-I is characterized by (ϕ,ψ) pairs clustered around (−30, 20). Conformer AA-II is characterized by (ϕ,ψ) pairs around (20, −15).

TABLE 1: Thermal Energies of Cyclic Dipeptides[a]

	cyclo (L-Ala-L-Ala) I	cyclo (L-Ala-L-Ala) II	cyclo (L-Pro-L-Ala)	cyclo (L-Pro-Gly)	cyclo (L-Pro-L-Leu)
CVFF[b]	36.0	37.4	58.6	60.4	54.6
RHF					
3-21G		117.8	142.8	123.2	202.6
6-31G*		118.2	143.4	123.8	203.0
6-311G*	117.9		142.7	123.2	202.0
6-311G**	117.3		142.0	122.6	200.9
B3LYP					
3-21G		109.9	134.2	115.6	190.8
6-31G*		110.7	134.2	115.7	190.4
6-311G*		109.2	133.7	115.3	189.6
6-311G**		110.0	133.2	114.9	188.9
BLYP					
3-21G		107.5	130.3	112.1	185.4
6-31G*		107.5	130.2	112.2	184.9
6-311G*		107.1	129.8	111.8	184.2
6-311G**		106.8	129.4	111.5	183.6
BVWN					
3-21G	108.5		131.6	113.2	187.3
6-31G*	108.5		131.5	113.3	186.8
6-311G*		108.3	131.1	112.9	186.2
6-311G**		108.0	130.7	112.6	185.7
MP2					
3-21G	111.6	111.8	134.9	116.5	191.8
6-31G*	111.4		134.6	117.5	193.7
6-311G*	111.4	111.1	133.5	116.8	192.2
6-311G**	110.2	110.6	133.0	116.4	

[a] All energies are in kcal/mol. [b] Energies from CVFF and the crystal structure as given by InsightII in the minimization output with the CVFF force field.

ference for molecular mechanics and a 0.4 kcal/mol difference for MP2, with the AA-I structure consistently having the lower energy (Table 1). This would correspond to a Boltzmann distribution at 25 °C favoring the AA-I conformation (66:34) using the MP2 energies and (91:9) using the molecular mechanics energies.

The different structures produced observable differences in predicted CD spectra (Figure 5). This is not surprising since the earlier dipole interaction model calculations on this peptide predicted different CD for planar and puckered diketopiperazine ring.[6] Optimization of the crystal structure[35] using MP2 yielded AA-II structures that produced CD with a weak positive band in the region where experiment was negative; the earlier dipole interaction model calculations did not make this observation for any puckered diketopeperazine ring.[6] This difference is probably due to the rigid assignment of bond lengths, bond

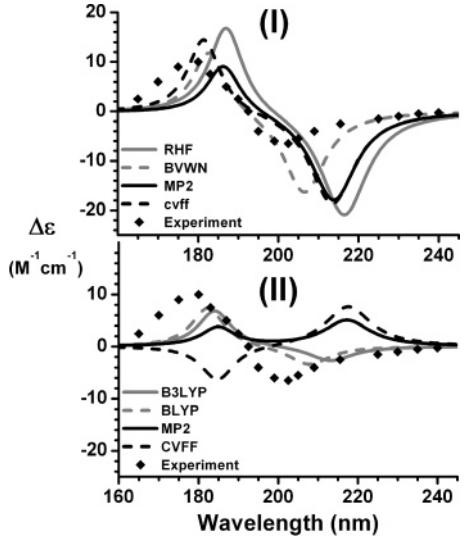

Figure 5. Calculated CD spectra of cyclo(L-Ala-L-Ala). All quantum mechanically optimized structures were optimized with the 6-311G* basis set. (I) AA-I geometries; (II) AA-II geometries. Bandwidth for each spectrum is 3000 cm^{-1}. The experimental CD (in water) was obtained from Bowman et al.[61]

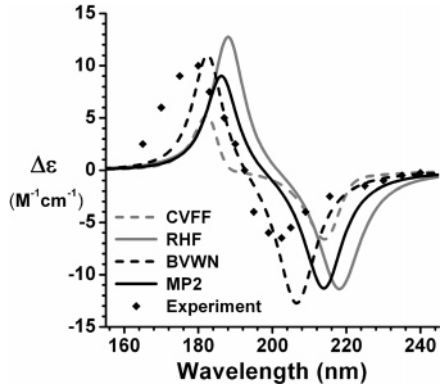

Figure 6. Boltzmann-weighted CD spectra of cyclo(L-Ala-L-Ala). All quantum mechanically optimized structures shown used the 6-311G* basis set. The bandwidth for each spectrum is 3000 cm^{-1}. The experimental CD (in water) was obtained from Bowman et al.[61]

angles, and torsion angles of the original dipole interaction model calculation. Optimizations yielding AA-I conformations matched band signs and predicted stronger bands than AA-II conformations. The difference in the CD predicted for the two kinds of structures may be a reflection of the difference in energies. The CD spectra of MP2 geometrically optimized structures from the BVWN/6-311G* geometry resemble experimental values better as described by band location, sign, and half-peak bandwidth. When spectra were weighted using a Boltzmann weighting scheme utilizing the energy differences from the MP2−6311G** optimizations, the composite spectrum more closely resembled experiment than the spectrum of either conformation alone for the QM optimized structures (Figure 6).

Cyclo(L-Pro-L-Ala). The molecular mechanics minimization of cyclo(L-Pro-L-Ala) and all quantum mechanically calculated structures resemble closely the crystal structure.[59] Bond lengths and angles do not vary much between the structures, but the CVFF minimization deviates most from the crystal structure. The dihedral angles ϕ and ψ of the CVFF (MM) structure deviate significantly from those of all other structures, including the crystal. The proline ϕ_1 falls significantly out of the common range for proline rings,[36] and the crystal structure itself

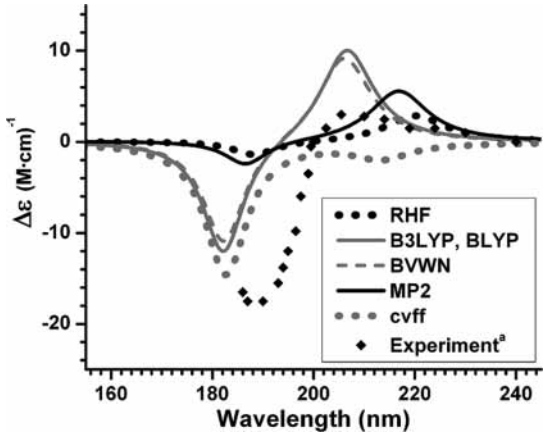

Figure 7. Predicted CD of cyclo(L-Pro-L-Ala). Calculations are for the largest basis set. The CD of the B3LYP structure is coincident with the CD predicted using the BLYP structure. Units of $\Delta\epsilon$ are M^{-1} cm^{-1} and units of wavelength are nm. The bandwidth is 3000 cm^{-1}. [a]Experimental CD obtained from Pancoska et al.[34]

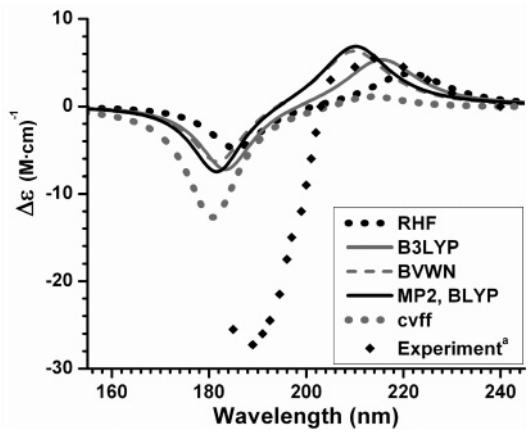

Figure 8. Predicted CD spectra of cyclo(L-Pro-Gly). Units of $\Delta\epsilon$ are M^{-1} cm^{-1} and units of wavelength are nm. The bandwidth for each spectrum is 4000 cm^{-1}. [a]Experimental CD from Pancoska et al.[59]

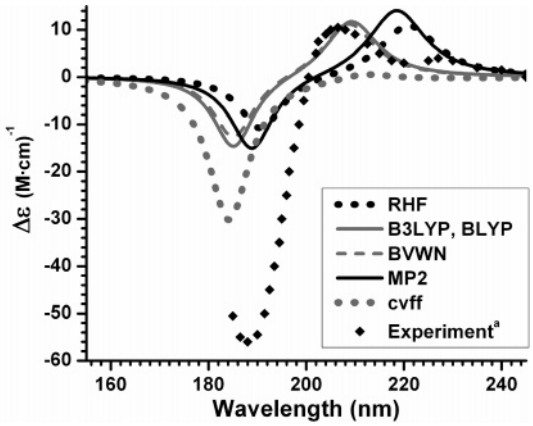

Figure 9. Predicted CD spectra of cyclo(L-Pro-L-Leu). Units of $\Delta\epsilon$ are M^{-1} cm^{-1} and units of wavelength are nm. The bandwidth used was 3000 cm^{-1}. [a]Experimental CD obtained from Pancoska et al.[59]

approaches the limit in the −30° range.[59] The amide dihedral angle (ω), however, is only slightly more planar than that of the crystal structure. This is not terribly surprising since the sp^2 hybridization of the amide carbon is a heavily weighted parameter in the force field, owing to the known nature of peptide bonds to approach planarity when not involved in complex turns.

The CD calculated for cyclo(L-Pro-L-Ala) varies with geometry (Figure 7). The CVFF-minimized geometry predicts two bands, but the band around 210 nm has an incorrect sign. Every QM geometry predicts the existence of both $\pi-\pi^*$ bands with correct signs, although each is wider than the corresponding experimental CD features. Calculated CD spectra for the structures obtained by both pure DFT methods (BVWN and BLYP) resemble experimental CD[59] most closely among the set of geometric optimizations performed here.

Cyclo(L-Pro-Gly). There is very little difference in the structures obtained by the various minimization methods and the crystal[59] structure of cyclo(L-Pro-Gly), with a couple of notable exceptions in dihedral angles. The CVFF predicted dihedral angles ϕ and ψ are significantly different from any others; as with cyclo(L-Pro-L-Ala), the proline ϕ_1 falls significantly out of the common range for proline rings, being at −19°; all other optimizations yielded values between −37° and −40°. The relatively subtle structural differences are reflected in the CD spectra calculated (Figure 8). The experimental band at 215 nm[6] fails to appear in the CD spectra calculated from the CVFF-minimized structure, while all other minimizations have CD spectra that possess this feature. As with cyclo(L-Pro-L-Ala), experimental CD spectra were obtained in 2,2,2-trifluoroethanol.[60] Again, the best agreement of theoretical and experimental CD occurs for the series of pure DFT or MP2 minimized structures.

Cyclo(L-Pro-L-Leu). The various geometrically optimized structures for cyclo(L-Pro-L-Leu) are a homogeneous group with only CVFF predicting a significantly different structure. The dihedral angles obtained by all four quantum techniques are within 1.5° when comparing minimizations using the same basis sets; moreover, basis set effects on geometry are small, even between 3−21G and 6-31G*. The CVFF geometry, however, had ϕ and ψ values that varied as much as 20° from the quantum mechanical and crystal structures. Despite the structural similarity among the quantum mechanical geometries, the dipole moments vary between 1.4 and 2.0 D, and the structures produce different calculated CD spectra (Figure 9). Following the same trend as the other dipeptides, the structures obtained by both pure DFT methods yield calculated CD spectra that match experimental CD[21,22] well. However, in this case, the MP2 calculations yielded results that are in poorer agreement with experiment than the DFT results. The CVFF-minimized structure produced a calculated CD spectra that lacks the $\pi-\pi^*$ band at 205 nm. All RHF geometries improve on this, but redshift this band nearly 20 nm from its experimentally observed location. The calculated CD spectra from the hybrid DFT geometries bring this band closer to the real value, but not as close as the pure DFT results.

Discussion

All optimizations of these peptides support the observations that crystal structures for peptides have a tendency to underestimate isolated molecule bond lengths, particularly the distances between aliphatic carbons. This observation is not new, because for the dipole interaction model to predict the CD spectra for poly(L-Pro) I and II[28] and for cyclic dipeptides,[6] the aliphatic carbon−carbon distances had to be lengthened to 1.54 Å.[61] Earlier molecular mechanics optimizations on cyclo(Gly-L-Pro-Gly)$_2$[58] and cyclo(Pro)$_3$[58] also lengthened bond distances enough so that the dipole interaction model was able to predict reasonable CD without requiring all C−C bonds lengths to be identical. Quantum mechanical optimizations for gas-phase molecules are rarely compared with crystal bond data, but are

compared with data from high accuracy spectroscopic methods for small molecules (e.g., microwave, IR). Quantum mechanical calculations can bracket the experimental values. (Although there are many studies that demonstrate this fact, the reader is referred to ref 62, which represents a systematic modern study.) It is important to remember that crystal structures represent the solid state, whereas CD is measured in solution, and our calculations are in gas phase. Without crystal packing forces, it may very well be natural for the bonds to lengthen slightly in both solution and in the gas phase. All the optimizations done herein indicate that this observation of short bond lengths in crystal structures very well may be real and not just an artifact of the dipole interaction model.

Geometric optimization of cyclo(L-Ala-L-Ala) located two separate minimum energy structures, although not all methods were able to converge on both. When the calculated CD spectra for the two structures were Boltzmann averaged using the MP2−6311G** energies, the composite spectra exhibited a considerable improvement in agreement to experiment. The best agreement of calculated CD of a single conformation with experimental CD[58] occurs for the flattened structure (AA-I) obtained using BVWN and to a lesser extent the MP2-calculated structure. However, the upward-puckered structure (AA-II) of B3LYP and BLYP leads to CD spectra possessing qualitatively good agreement with the experimental values. A quantitative match to experiment may require taking this dynamic behavior into account.

Hooker et al. suggested that cyclo(L-Ala-L-Ala) could potentially exist in three conformations, two of which were similar to those found herein, and the third with the methyl group straight up and orthogonal to the diketopiperazine ring.[63] Hooker et al. indicated that small distortions in geometry play only a secondary role in optical calculations but a major role in conformational energy calculations.[63] This observation may be a result of using the matrix method, which, at the time Hooker et al.[26] did the calculations, did not include the side chain atoms in the optical calculation; i.e., the calculations did not account for the amide−aliphatic group interactions. We observe that the conformational energies are quite close in the MP2 optimizations, but the molecular mechanics force field CVFF overestimates the energy gap between the conformations. Our molecular mechanics energy separation is similar to those that Hooker et al. discussed. However, the high-level calculations performed here suggest that CD spectra do vary significantly with structure. Geometric optimization of all other peptides examined in this study located a single minimum energy conformation, a result consistent with the highly constrained nature of these systems.

(1) What is the sensitivity of the dipole interaction model to small refinements in molecular geometry, and what level of geometry optimization is sufficiently accurate to successfully predict CD spectra of cyclic dipeptides? In general, the DFT methods of structural determination for these molecules are comparable to MP2 in their ability to provide structures that give good agreement with theoretical CD when the dipole interaction model is used to predict their CD spectra, while working with a computational efficiency similar to RHF. In agreement with the cyclo(L-Pro)$_3$ work of Lowe et al.,[21] even a 3-21G basis set was sufficient with any quantum mechanical method to qualitatively predict the CD of each of these peptides. Use of this small basis set with even the simplest ab initio method, RHF, resulted in structures for which the dipole interaction model calculated spectra significantly closer to the experimental spectra than those obtained using the CVFF force field, particularly in respect to the positive band at approximately 205 nm. Using a larger basis set, however, altered the geometries of every molecule enough to see a distinct improvement in the quality of the predicted spectra. The 6-31G* basis set was sufficiently large to accurately describe molecular geometries, as measured by agreement of theoretical CD spectra with experimental values. Expansion to the larger basis set 6-311G*, however, resulted in a slight improvement of CD spectra and a significant improvement in calculated energies. Structures predicted using the 6-311G** basis set were nearly identical to those calculated with 6-311G*, and their CD spectra were indistinguishable; thus, use of the additional polarization functions, which comes at significant computational cost, is unnecessary to sufficiently describe these molecules for prediction of their CD spectra with the dipole interaction model.

(2) Which of the dipole interaction parameters are best suited for use with cyclic dipeptides? All available parameters (original, G, H, and J) were used to predict CD spectra of these molecules. No other parameter set approached the descriptive capability of the original set (Figure 2). Regardless of the peptide studied, all G-parameter obtained spectra were indistinguishable. The H parameter set generally gave two positive $\pi-\pi^*$ bands, and the J parameter descriptive ability varied widely between various structures, making it completely nonpredictive. The success of Applequist's original parameters in predicting CD that closely matches experimental values is somewhat surprising, given past results using the model; for example, cyclo(L-Pro)$_3$[59] and the poly(R)-β-aminobutyric acid in an antiparallel sheet were treated best by the H (helical) parameters.[61] The flexible cyclic hexapeptide, cyclo(Gly-L-Pro-Gly)$_2$, on the other hand, was treated most accurately with the J (poly-L-Pro II parameters).[61] The piperazine-2,5-diones, however, may not be out of line with respect to parameter set success given the rigidity of this group of molecules and that they are not representative of larger typical secondary structures such as helices, sheets, or turns although they are excellent model amides.

(3) How favorably do the dipole interaction model's predictions compare with experimental values? The original parameters of the dipole interaction model produce qualitative agreement with experiment, but they have a tendency to overestimate the ratio of the half-peak bandwidths of the two bands. This may be because the bandwidth is assumed to be identical for each band in the model. This assumption has been used historically with the dipole interaction model to simplify the computation and to prevent users from simply over-fitting the calculation with a collection of Lorentzian bandwidths.

There was significant discrepancy in the 190 nm band intensity for cyclo(L-Pro-L-Ala), cyclo(L-Pro-Gly), and cyclo(L-Pro-L-Leu). The gas-phase structural and CD calculations failed to reproduce the experimental spectral feature of a deep $\pi-\pi^*$ band at 190 nm in hydrogen-bonding solvents such as 2,2,2-trifluoroethanol. The depth of the peak is both a structural and an electronic effect of solvent hydrogen bonding,[59] and without accounting for structural changes to the peptide and the influence of solvent anisotropies on the amide chromophore, this feature cannot be reproduced. It may be possible to do molecular dynamics simulations with molecular mechanics of the peptides in explicit solvent and calculate the CD including the solvent at a variety of different snapshots, but that is beyond the scope of this current exploratory study, which focuses on the sensitivity of the dipole interaction model to various methods of geometric optimization.

(4) Does the dipole interaction model recognize poor geometries (i.e., is it a good tool for evaluating molecular geometries)? The dipole interaction model is capable of spotting

serious structural problems by either failing to produce a CD spectrum that compares well with experiment (e.g., by predicting wrong band signs or with missing bands) or it will not predict a CD spectrum at all. The dipole interaction model relies on mixing of dipole moments through an origin-independent matrix relying on the relative positions of all atoms in the molecule. While small deviations from the correct molecular geometry can "stretch" the elements of the Hermitian mixing matrix, causing theoretical CD with little or no resemblance to experiment, significant deviations obliterate the symmetry of the matrix and result in nonexistence of its inverse.

The CD predicted using poor proline geometries from the CVFF molecular mechanics consistently failed to predict either the existence of the CD between 205 and 214 nm or the band was the wrong sign or extremely weak. The model distinguished between the two conformations of cyclo(L-Ala-L-Ala). The CD spectra predicted from the lower energy cyclic alanine dipeptide resembled experimental values more, whereas the higher energy conformation led to calculated CD having variable resemblance to the experimental CD, with some structures even leading to bands of reversed sign. All optimizations generally increased the aliphatic carbon−carbon bond lengths, suggesting that the slightly longer bond lengths are real. The dipole interaction model failed to predict CD for any unoptimized crystal structure, further indicating the longer bond lengths are real. Thus, the dipole interaction model does have the ability to spot serious structural problems or differences in models of these peptides.

Conclusions

The dipole interaction model qualitatively describes the $\pi-\pi^*$ transition feature of the UV CD spectra of piperazine-2,5-diones (cyclic dipeptides). The additional effort to obtain accurate geometries via ab initio and DFT methods over molecular mechanical optimization resulted in significant improvement of the quality of CD spectra predicted using the classical dipole interaction model. However, optimizing some of the geometries at the higher ab initio levels and with larger basis sets proved to be nontrivial. One can couple quantum mechanical geometries with a classical physics model for calculation electronic UV circular dichroism. Consequently, if a physically accurate molecular structure is obtained, it should be possible to use the classical dipole interaction model to predict the UV CD of biological molecules as well. If ensembles of structures are necessary to reconstruct CD spectra, obtaining good energies to quantify the Boltzmann averaging or a method to follow the dynamics of the molecule (e.g., molecular dynamics) may also be a critical step in faithful description of a composite CD.

Acknowledgment. This work has been supported by the UND Seed Program, NIH ND BRIN grant and the NIH Grant Number P20 RR016741 from the INBRE Program of the National Center for Research Resources, and the ND Computational Chemistry and Biology Network (CCBN). K.C. has been supported by a Department of Education GAANN fellowship.

Supporting Information Available: Additional structural information for the various optimization of the peptides can be found in the following supplemental tables and figures. (1) Bond lengths, bond angles, and dihedral angles for each of the molecules in this study. (2) CD band analysis for each of the peptides in this study, including peak location, intensity, and half-peak bandwidths. (3) Graphs of all calculated CD for each minimized structure using all parameter sets (original, G, H, and J) and a bandwidth of 3000 or 4000 cm^{-1}. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Sreerama, N.; Woody, R. W. Circular Dichroism of Peptides and Proteins. In *Circular Dichroism: Principles and Applications*, 2nd ed.; Berova, N., Nakanishi, K., Woody, R. W., Eds.; John Wiley & Sons: New York, 2000; p 601.
(2) Tinoco, I., Jr. *Adv. Chem. Phys.* **1962**, *4*, 113.
(3) Bayley, P. M.; Nielsen, E. B.; Schellman, J. A. *J. Phys. Chem.* **1969**, *73*, 228.
(4) Woody, A.-Y. M.; Woody, R. W. *Biopolymers (Biospectroscopy)* **2003**, *72*, 500.
(5) Hirst, J. D. *J. Phys. Chem. B* **2003**, *107*, 11813.
(6) Sathyanarayana, B. K.; Applequist, J. *Int. J. Peptide Protein Res.* **1985**, *26*, 518.
(7) Applequist, J.; Sundberg, K. R.; Olson, M. L.; Weiss, L. C. *J. Chem. Phys.* **1979**, *71*, 2330.
(8) Applequist, J.; Sundberg, K. R.; Olson, M. L.; Weiss, L. C. *J. Chem. Phys.* **1979**, *70*, 1240.
(9) Applequist, J. *J. Chem. Phys.* **1979**, *71*, 4332.
(10) Bode, K. A.; Applequist, J. *J. Phys. Chem.* **1996**, *100*, 17825.
(11) Applequist, J.; Carl, J. R.; Fung, K.-K. *J. Am. Chem. Soc.* **1972**, *94*, 2952.
(12) Applequist, J. *Biopolymers* **1981**, *20*, 387.
(13) Applequist, J. *J. Chem. Phys.* **1980**, *73*, 3521.
(14) Applequist, J. *J. Chem. Phys.* **1979**, *71*, 1983.
(15) Applequist, J. *J. Chem. Phys.* **1979**, *71*, 4324.
(16) Woody Robert, W. *J. Chem. Phys.* **1968**, *49*, 4797.
(17) Woody, R. W.; Tinoco, I., Jr. *J. Chem. Phys.* **1967**, *46*, 4927.
(18) Bode, K. A.; Applequist, J. *J. Phys. Chem.* **1996**, *100*, 17820.
(19) Bode, K. A.; Applequist, J. *J. Phys. Chem. A* **1997**, *101*, 9560.
(20) Applequist, J. *Biopolymers* **1982**, *21*, 779.
(21) Huber, A.; Nkabyo, E.; Warnok, R.; Skalsky, A.; Kuzel, M.; Gelling, V. J.; Dillman, T. B.; Ward, M. M.; Guo, R.; Kie-Adams, G.; Vollmer, S.; Ngassa, F. N.; Lowe, S. L.; Ouporov, I. V.; Thomasson, K. A. *J. Undergrad. Chem. Res.* **2003**, *2*, 145.
(22) Sathyanarayana, B. K.; Applequist, J. *Int. J. Peptide Protein Res.* **1986**, *27*, 86.
(23) Bode, K. A.; Applequist, J. *Biopolymers* **1997**, *42*, 855.
(24) Applequist, J. *J. Chem. Phys.* **1979**, *71*, 4332.
(25) Bode, K. A.; Applequist, J. *Macromolecules* **1997**, *30*, 2144.
(26) Applequist, J.; Bode, K. A. *J. Phys. Chem. A* **2000**, *104*, 7129.
(27) Applequist, J. *J. Phys. Chem. A* **2000**, *104*, 7133.
(28) Applequist, J. *Biopolymers* **1981**, *20*, 2311.
(29) Thomasson, K. A.; Applequist, J. *Biopolymers* **1991**, *31*, 529.
(30) Caldwell, J. W.; Applequist, J. *Biopolymers* **1984**, *23*, 1891.
(31) Bode, K. A.; Applequist, J. *J. Am. Chem. Soc.* **1998**, *120*, 10938.
(32) Allen, F. H. A. *Acta Crystallogr.* **2002**, *B58*, 380.
(33) Bruno, I. J.; Cole, J. C.; Edgington, P. R.; Kessler, M.; Macrae, C. F.; McCabe, P.; Pearson, J.; Taylor, R. *Acta Crystallogr.* **2002**, *B58*, 389.
(34) Sletten, E. *J. Am. Chem. Soc.* **1970**, *92*, 172.
(35) Cotrait, M.; Leroy, F. *Cryst. Struct. Commun.* **1979**, *8*, 819.
(36) Von Dreele, R. B. *Acta Crystallogr. B* **1975**, *B31*, 966.
(37) Karle, I. L. *J. Am. Chem. Soc.* **1972**, *94*, 81.
(38) Dauber-Osguthorpe, P.; Roberts, V. A.; Osguthorpe, D. J.; Wolff, J.; Genest, M.; Hagler, A. T. *Proteins: Struct., Funct., Genet.* **1988**, *4*, 31.
(39) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; 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.; Rega, N.; Salvador, P.; Dannenberg, J. J.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; 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.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *GAUSSIAN98*, Revision A.11.4; Gaussian, Inc.: Pittsburgh, PA, 2002.
(40) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.
(41) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
(42) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200.
(43) Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. *Chem. Phys. Lett.* **1989**, *157*, 200.
(44) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618.
(45) Binkley, J. S.; Pople, J. A. *Int. J. Quantum Chem.* **1975**, *9*, 229.
(46) Dobbs, K. D.; Hehre, W. J. *J. Comput. Chem.* **1987**, *8*, 880.
(47) Dobbs, K. D.; Hehre, W. J. *J. Comput. Chem.* **1987**, *8*, 861.
(48) Binkley, J. S.; Pople, J. A.; Hehre, W. J. *J. Am. Chem. Soc.* **1980**, *104*, 939.

(49) Gordon, M. S.; Binkley, J. S.; Pople, J. A.; Pietro, W. J.; Hehre, W. J. *J. Am. Chem. Soc.* **1982**, *104*, 2797.
(50) Pietro, W. J.; Francl, M. M.; Hehre, W. J.; DeFrees, D. J.; Pople, J. A.; Binkley, J. S. *J. Am. Chem. Soc.* **1982**, *104*, 5039.
(51) Binning, R. C., Jr.; Curtiss, L. A. *Int. J. Quantum Chem.* **1991**, *40*, 781.
(52) Ditchfield, R.; Hehre, W. J.; Pople, J. A. *J. Chem. Phys.* **1971**, *54*, 724.
(53) Gordon, M. S. *Chem. Phys. Lett.* **1980**, *76*, 163.
(54) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213.
(55) Hariharan, P. C.; Pople, J. A. *Mol. Phys.* **1974**, *27*, 209.
(56) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257.
(57) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650.
(58) Hooker, T. M., Jr.; Bayley, P. M.; Radding, W.; Schellman, J. A. *Biopolymers* **1974**, *13*, 549.
(59) Pancoska, P.; Fric, I.; Blaha, K. *Collect. Czech.Chem. Commun.* **1979**, *44*, 1296.
(60) Applequist, J. *Biopolymers* **1981**, *20*, 2311.
(61) Bowman, R. L.; Kellerman, M.; Johnson, W. C., Jr. *Biopolymers* **1983**, *22*, 1045.
(62) Cremer, D.; Kraka, E.; He, Y. *J. Mol. Struct.* **2001**, *567*−*568*, 275.
(63) Lowe, S. L.; Pandey, R. R.; Czlapinski, J.; Kie-Adams, G.; Hoffmann, M. R.; Thomasson, K. A.; Pierce, K. S. *J. Peptide Res.* **2003**, *61*, 189.