ChemComm

A New Interpretation of the Structure and Solvent Dependence of the far UV Circular Dichroism Spectrum of Short Oligopeptides

Journal:	*ChemComm*
Manuscript ID	CC-COM-02-2019-001513.R1
Article Type:	Communication

COMMUNICATION

A new interpretation of the structure and solvent dependence of the far UV circular dichroism spectrum of short oligopeptides†

Anshuman Kumar,[a] Reinhard Schweitzer-Stenner*[b] and Bryan M. Wong*[a]

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

UV circular dichroism (UVCD) spectroscopy is a prominent tool for exploring secondary structures of polypeptides and proteins. In the unfolded state of these biomolecules, most of the individual residues primarily sample a conformation called polyproline II. Its CD spectrum contains a negatively biased positive couplet with a pronounced negative maximum below and a weak positive maximum above 200 nm. It is traditionally rationalized in terms of an excitonic coupling mechanism augmented by polarization effects. In this work, we carry out new time-dependent density functional theory calculations on the cationic tripeptide GAG in implicit and explicit water to determine the transitions that give rise to the observed CD signals of polyproline II and β-strand conformations. Our results reveal a plethora of electronic transitions that are governed by configurational interactions between multiple molecular orbital transitions of comparable energy. We also show that reproducing the CD spectra of polyproline II and β-strand conformations requires the explicit consideration of water molecules. The structure dependence of delocalized occupied orbitals contributes to the experimentally-observed invalidation of Flory's isolated pair hypothesis.

Far UVCD spectroscopy is one of the most employed approaches for determining the secondary structure of proteins and peptides in solution.[1, 2] Each secondary structure gives rise to a fingerprint with a characteristic sequence of positive and negative maxima. The structural sensitivity of the UVCD spectra results from the orientational dependence of couplings between electronic transitions in different peptide groups.[3]

Unfolded and disordered proteins, polypeptides, and even very short tripeptides such as glycylalanylglycine (GAG) frequently exhibit UVCD spectra with a pronounced negative maximum at ca. 190 nm. Additionally, poly- and oligopeptides with a substantial alanine or proline content exhibit a weak positive maximum at ca. 214 nm (Figure 1).[4, 5] The conventional understanding of such CD spectra is that they arise from a random coil state in which the backbone of the peptide samples the entire sterically allowed region of the Ramachandran plot.[1] However, this notion was already challenged ca. fifty years ago when Tiffany and Krimm assigned it to a polyproline II (pPII) conformation that exhibits dihedral angles of ϕ=-70° and ψ=140°.[6] Their interpretation was challenged because it was generally believed that short peptides cannot adopt any well-defined conformation in aqueous solution. However, thanks to the work of Woody and colleagues,[7-10] Tiffany and Krimm were eventually proven to be correct. Today, we know that the very pronounced pPII signal in the spectrum of alanine-containing peptides reflects the high propensity of the corresponding amino acid residue for this conformation.[11, 12]

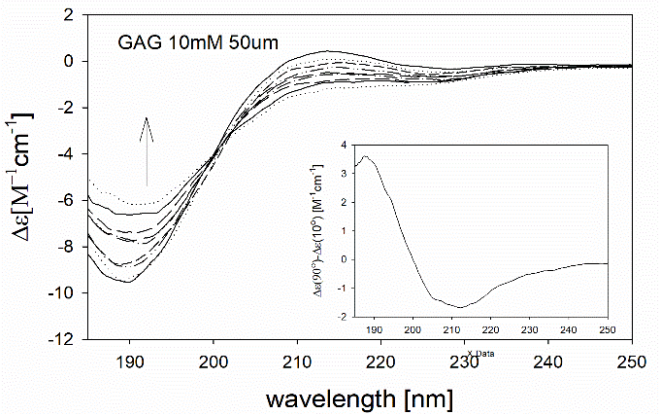

Fig. 1 UV ECD spectrum of cationic GAG in water measured as a function of temperature. The arrow indicates the direction of the changes below 200 nm. The inset shows the difference spectrum calculated by subtracting the spectrum measured at 100°C from the spectrum taken at 90°C. Taken from Ref. 5 and modified.

[a] Department of Chemical & Environmental Engineering, Materials Science & Engineering Program, University of California-Riverside, 900 University Avenue, Riverside, CA 92521, USA. E-mail: bryan.wong@ucr.edu
[b] Department of Chemistry, Drexel University, 3141 Chestnut Street, Philadelphia, PA 19104, USA. E-mail: rs344@drexel.edu

† Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

While the relationship between the strongly negatively-biased positive couplet in Figure 1 and the pPII conformation has been firmly established over time, an explanation of the underlying physics turned out to be challenging. Generally, the UVCD spectra of secondary structures are explained in terms of excitonic coupling models, which describe the rotational strength of electronic transitions in terms of either dipole moments or transition charge densities associated with individual electronic transitions between the HOMO and the lowest unoccupied MOs of individual peptide groups.[3, 7] The canonical UVCD spectra were thus described by considering interactions between the π(HOMO)→π*(LUMO) (NV_1) and n→π* transitions, which were shown to produce considerable rotational strength at the wavelength position of the latter. The n→π* transition is barely visible in the corresponding absorption spectra owing to its rather small electronic transition dipole moment. The rotational strength, however, is written as the dot product of the electric and magnetic transition dipole moment so that it can become detectable even in the absence of a visible contribution to the absorption spectrum.

The above model can only account for symmetric CD signals where the integrated intensities of the negative and positive bands are identical. Apparently, this notion does not apply to the pPII signal in Figure 1. To remedy the situation, several attempts have been made to include polarization effects in the formalism for the rotational strength.[7, 13] Empirically, the polarizability tensor is treated as a ground state property with empirical values for chemical bonds listed in the literature.

We wondered whether a more definite physical assessment of the UVCD spectra of short peptides could be provided by time dependent density functional theory (TDDFT) calculations. Previous studies[14-16] have utilized TDDFT calculations to examine the underlying physics of the CD spectra of model peptides and proteins in their folded state. A similar investigation of unfolded peptides is still outstanding. We recently performed DFT calculations on four different cationic GxG peptides (x = A, V, I, L) in implicit and explicit water (10 H_2O molecules) in order to obtain the energetics of their pPII and β-strand conformations.[17] For the present study we used the optimized geometry of these two conformations in TDDFT calculations at the ωB97X-D/cc-pVTZ level of theory to calculate the UVCD and corresponding absorption spectra of GAG in implicit and explicit water.[17] We specifically chose the ωB97X-D functional for our studies since it contains an asymptotically-correct (range-separated) portion of nonlocal exchange in conjunction with dispersion corrections, which are essential for accurately predicting charge-transfer excitations and hydrogen-bonding interactions,[18] respectively. It is also important to note that prior work by Neto et al.[19] has shown that the optimal range-separated parameter, ω, required to accurately predict excited-states in explicit solvent is 0.25, which is already close to the default value of ω = 0.2 used in the ωB97X-D functional. To test the effects of (1) changing the range-separation value, (2) using different exchange-correlation functionals, (3) utilizing a larger basis set (such as aug-cc-pVTZ), and (4) altering the positions of explicit water molecules, we spot-checked our calculations by comparing the computed ECD spectra obtained separately with each of these effects. As shown in Figs. S1 – S4 in the Supplementary Information, we did not observe any significant change in the ECD spectra within the relevant wavelength range, which further validates our choice of the ωB97X-D/cc-pVTZ level of theory used in our studies. The dihedral angles of all conformations are listed in Table S1, and the corresponding structures are depicted in Figure S5 and S7. Further details of our computational methods are given in the Supplementary Information (SI).

Figure 2 shows the calculated CD spectra and the underlying electronic transitions of the pPII conformation. The corresponding absorption spectra are displayed in Figure S6 and S8 for GAG in the pPII and β- conformation, respectively. The CD spectra exhibit a negatively-biased positive couplet with a positive maximum at 202 nm and a negative maximum at 172 nm, which resemble the shape of a canonical pPII signal. As one would expect, the calculated CD maxima and corresponding absorption bands are at higher energies than the corresponding experimental values, from which they appear blueshifted by 14 and 18 nm. The couplets obtained with implicit and explicit water are very similar. It is slightly redshifted from the position of the corresponding absorption band for implicit water, whereas it coincides with the absorption band for explicit water.

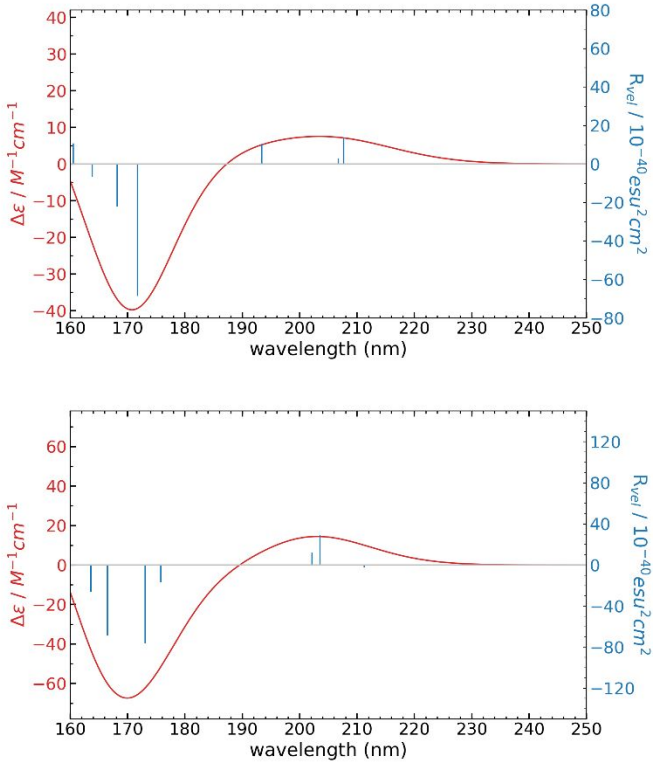

Fig. 2 ECD spectra of GAG in implicit (upper panel) and explicit water (lower panel) calculated for the optimized pPII conformation.

Experimentally, the CD spectrum of the β-strand conformation exhibits a negative couplet with a weak and broad negative maximum at ca. 212 nm and a positive maximum at ca. 190 nm, similar to the difference spectrum shown in Figure 1.[2, 8, 20] The calculation for the optimized β-strand structure of GAG

in implicit water reproduces the experimental spectrum only partially (Figure 3). The spectrum in the region between 220 and 160 nm exhibits local maxima at 202 and 172 nm, respectively. The latter is redshifted relative to the corresponding absorption band (Figure S8). In contrast to this result for implicit water, the CD spectrum of the β-strand conformation in explicit water meets the consistency criterion. It features a positive maximum at 176 nm (slightly redshifted from the absorption band) and a broad negative maximum at 202 nm. The much closer correspondence between the spectrum calculated with explicit water and the experimental spectrum suggest the necessity of performing calculations on peptide-water clusters for a good reproduction of β-strand CD spectra. This result is generally in line with findings of other researchers[16, 21] who found that the inclusion of explicit solvent is essential for calculations aimed at reproducing the CD spectra of folded polypeptides.

It is generally believed that the absorption band at 190 nm results from a π→π* transition between the HOMO and LUMO, and the corresponding CD spectra reflect excitonic couplings between electronic transitions of different peptide groups.[7] Our calculations reveal a more complex and significantly different picture. The band at 176 nm, which corresponds to the 190 nm band in the experimental spectrum, reflects contributions from multiple electronic transitions, each of which is produced by configurational interactions between electron transfers between occupied and unoccupied molecular orbitals. The number and oscillator strength of these transitions depend on the peptide conformation as well as on the water model. In the CD spectrum of the explicit water model, the negative maximum of the pPII signal reflects two major and two minor contributions. The corresponding band in the β-strand spectrum is dominated by a single transition. Since the analysis of the excitations involve complex configuration interactions (CI) composition, we employ natural transition orbitals[22] (NTO) to analyse the nature of these excitations. NTOs provide a compact representation of the orbitals and offer a convenient way to visualize excitations via 'hole' and 'particle' descriptors, respectively. Some of the natural transition orbitals involved in these transitions are in part delocalized over both peptide groups and the alanine residue between them. Others are more localized at one of the peptide groups and/or the C-terminal carboxylate group. (Figure S10). The lowest energy peak is dominated by the second excited state, and visualization of the NTOs implies charge transfer character for this excitation. The hole is localized towards the alanine residue and the particle is localized towards the carboxylate group. The negative dominant peak involves the fifth and sixth excited state, and NTOs depict the redistribution of electron away from the carboxylate group. A list of relevant energies with their rotatory strength for the first five excited states are provided in Tables S2 and S3 for GAG in the pPII and β conformations, respectively.

The influence of water and geometric conformation on the electronic structure of GAG can be inferred from Figures S9-S12 which depict the NTOs that contribute most to the relevant excited states. The influence of explicit water and the backbone conformation can be illustrated by a comparison of corresponding NTOs. For implicit water, the holes of pPII and the β-strand are both mostly localized in the plane of the C-terminal peptide group, while the corresponding particle is localized perpendicular to the plane of the C-terminal peptide group. Hence, the underlying transition clearly has charge-transfer character. For GAG in the pPII conformation with implicit water, the corresponding NTOs contribute around 99% to the transition into the first excited state at around 202 nm. Similarly, the negative maximum at around 170 nm arises from the fourth excited state where the relevant NTOs contribute nearly 93% for this transition. The hole is mostly delocalized across the two peptide bonds while the particle is localized mostly on the N-terminal peptide group. We note that in the negatively-biased positive couplet observed here in the CD spectra, the peaks are dominated by a transition into a single excited state rather than by an admixture of multiple excited-state transitions, as we will see for the case of spectra obtained with explicit water molecules around the GAG in the pPII conformation.

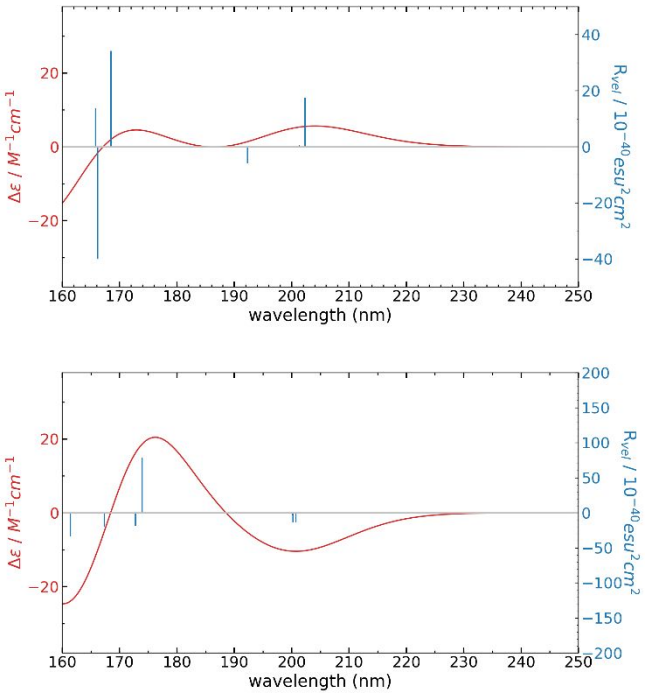

Fig. 3 ECD spectra of GAG in implicit (upper panel) and explicit water (lower panel) calculated for the optimized β-strand conformation.

In explicit water, the NTOs of the β-strand conformation are delocalized over both peptide groups and the C_α-C_β bond of the alanine residue. Both the hole and particle of the pPII conformation are delocalized over the C-terminal peptide and the adjacent carboxylate group for the first and second excited states. For the fifth excitation, which contributes to the negative peak in the CD spectra, the particle is delocalized over both peptide groups while the hole is localized at the alanine residue. In both cases the hole exhibits contributions from water orbitals. The particle, however, involves less water. In the β-strand conformation, water molecules hydrogen-bonded to the peptide groups contribute to the NTO whereas the particle does not exhibit a recognizable contribution from water molecules.

The simulation of the CD spectrum for GAG adopting a β-strand conformation in explicit water yielded a positively-biased negative couplet. The lowest-energy peak arises from transitions into the first and second excited state which are very close in energy. From Figure S12 in the SI we note that the negative peak at around 202 nm arises from transitions into the first and second excited state that have no significant charge transfer. We do not show particle/hole pairs for the second excited-state transition since they are very similar to those of the first excited-state transition. The positive peak at around 178 nm is dominated by transitions into the fourth excited state. The NTOs show a slight charge-transfer character in which the hole is more localized towards the alanine residue and the particle is localized at the peptide bond at C-terminal carboxylate group.

Taken together, our results show that the electronic structure of accessible excited states are far more complicated than generally envisaged. The degree of NTO delocalization is both structure and solvent dependent. Water is involved in the NTOs of both states, though the distribution depends on the conformation. A transition into the particle seems to decouple the peptide from water. These results reveal that the classical excitonic coupling models are unsuitable for describing the electronic structure of polypeptides and that the explicit consideration of hydrogen bonded water (not just the effect of hydrogen bonding) is pivotal for understanding the energetics and electronic structure of peptides. Apparently, one does not have to invoke polarization effects due to very high lying electronic transitions to reproduce the experimental CD spectra of pPII conformations.

The very fact that the UVCD spectra of very short peptides and unfolded polypeptides or proteins look practically indistinguishable strongly suggests that the results obtained for GAG offer some general insight about the electronic structure of unfolded peptides and proteins. UVCD spectra of different GxG peptides exhibit the similar negatively-biased signals, which differs only in terms of the peak intensities of their negative and positive maxima. Some of the peptide spectra do not show the latter. This reflects *different* pPII/β-strand mixtures.[2,5,7] The structure dependence of the delocalization of occupied orbitals invalidate Flory's isolated pair hypothesis, which are in line with experimental results.[23, 24]The huge difference between the corresponding occupied MOs of GAG in implicit and explicit water strongly suggest that solvation energies of different residues are not additive.

One might wonder whether a calculation for only two equilibrium conformations can really account for the observed CD spectra. Since the energy landscape is relatively flat, peptides still sample a significant portion of the Ramachandran plot. For GAG this is particularly true for the glycine residues, while the conformation space of the alanine residues is rather restricted.[4] However, the similarity of UVCD spectra of GxG peptides with rather different positions of pPII and β-strand-like distributions suggest that the spectra are either pPII or β-like as long as the peptide samples the Ramachandran plot about a ψ-value of 100^0. Hence, a consideration of the peptide's dynamics by means of MD/DFT calculations would most likely solely vary the amplitudes of the calculated CD couplet. Any attempt to explicitly account for electronic peptide-water interactions would require that a hydration shell of the size considered in this study should be included in the TDDFT part of the calculations, which would make them to very expensive to perform.

R.S.S. was supported in part by a grant from the National Science Foundation (DMR 1707770). A.K. and B.M.W. acknowledge partial support from a UC Riverside Collaborative Seed Grant and the Office of Naval Research (Grant N00014-18-1-2740).

Conflicts of interest

There are no conflicts to declare.

Notes and references

1. R. W. Woody, *Adv. Biophys. Chem.*, 1992, **2**, 37-79.
2. N. Sreerama and R. W. Woody, *Anal. Biochem.*, 2000, **287**, 252-260.
3. W. Moffit, *J. Chem. Phys.*, 1956, **25**, 467-478.
4. A. Hagarman, T. J. Measey, D. Mathieu, H. Schwalbe and R. Schweitzer-Stenner, *J. Am. Chem. Soc.*, 2010, **132**, 540-551.
5. S. E. Toal, D. J. Verbaro and R. Schweitzer-Stenner, *J. Phys. Chem. B*, 2014, **118**, 1309-1318.
6. M. L. Tiffany and S. Krimm, *Biopolymers*, 1968, **6**, 1767-1770.
7. R. W. Woody, *J. Am. Chem. Soc.*, 2009, **131**, 8234-8245.
8. N. Sreerama and R. W. Woody, *Protein Sci.*, 2003, **12**, 384-388.
9. Z. Shi, R. W. Woody and N. R. Kallenbach, *Adv. Protein Chem.*, 2002, **62**, 163-240.
10. Z. Liu, K. Chen, A. Ng, Z. Shi, R. W. Woody and N. R. Kallenbach, *J. Am. Chem. Soc.*, 2004, **126**, 15141-15150.
11. J. Graf, P. H. Nguyen, G. Stock and H. Schwalbe, *J. Am. Chem. Soc.*, 2007, **129**, 1179-1189.
12. S. Toal, D. Meral, D. Verbaro, B. Urbanc and R. Schweitzer-Stenner, *J. Phys. Chem. B*, 2013, **117**, 3689-3706.
13. W. E. Ronish and S. Krimm, *Biopolymers*, 1974, **13**, 1635-1651.
14. J. Šebek, Z. Kejík and P. Bouř, *J. Phys. Chem. A*, 2006, **110**, 4702-4711.
15. J. Kaminský, J. Kubelka and P. Bouř, *J. Phys. Chem. A*, 2011, **115**, 1734-1742.
16. J. Šebek, B. Gyurcsik, J. Šebestík, Z. Kejík, L. Bednárová and P. Bouř, *J. Phys. Chem. A*, 2007, **111**, 2750-2760.
17. N. V. Ilawe, A. E. Raeber, R. Schweitzer-Stenner, S. E. Toal and B. M. Wong, *Phys. Chem. Chem. Phys.*, 2015, **17**, 24917-24924.
18. N. V. Ilawe, R. Schweitzer-Stenner, D. DiGuiseppi and B. M. Wong, *Phys. Chem. Chem. Phys.*, 2018, **20**, 18158-18168.
19. A. P. V. Neto, D. F. S. Machado, T. O. Lopes, A. J. Camargo and H. C. B. de Oliveira, *J. Phys. Chem. B*, 2018, **122**, 8439-8450.
20. S. Brahms, J. Brahms, G. Spach and A. Brack, *Proc. Natl. Acad. Sci. U S A*, 1977, **74**, 3208-3212.
21. E. S. Pysh, *Biopolymers*, 1974, **13**, 1557-1562.
22. R. L. Martin, *J. Chem. Phys.*, 2003, **118**, 4775-4777.
23. S. E. Toal, N. Kubatova, C. Richter, V. Linhard, H. Schwalbe and R. Schweitzer-Stenner, *Chem. Eur. J.*, 2015, **21**, 5173-5192.
24. R. Schweitzer-Stenner and S. E. Toal, *Biophys. J.*, 2018, **114**, 1046-1057.