J|A|C|S
COMMUNICATIONS

Published on Web 09/02/2006

Charge-Transfer Transitions in Protein Circular Dichroism Calculations

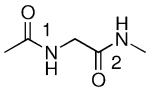
Mark T. Oakley and Jonathan D. Hirst*

School of Chemistry, University of Nottingham, University Park, Nottingham, NG7 2RD, U.K.

Received June 22, 2006; E-mail: Jonathan.hirst@nottingham.ac.uk

The UV absorption spectra of proteins contain many features that can be exploited in studies of protein folding. The far UV spectrum is dominated by an intense band at 190 nm, due to the electronically allowed $\pi_{nb} \rightarrow \pi^*$ transitions in the backbone peptide units. A very weak signal at 222 nm arises from the electrically forbidden $n \rightarrow \pi^*$ transitions. At even higher energy, the vacuum ultraviolet (VUV) region also contains strong absorptions, but their origin is not well established, and charge-transfer transitions are a possible source. Bands at 204 and 189 nm in dipeptides have been assigned to transitions between the terminal carboxylate group and peptide groups,[1] but there have been no previous experimental assignments of interpeptide charge-transfer transitions.

CD spectroscopy is more informative than UV absorption spectroscopy, because of its sensitivity to protein secondary structure. It is widely used to study the conformation of proteins and peptides.[2] Conventional CD spectrometers record at wavelengths down to 178 nm,[3] which is low enough to capture the peptide $n \rightarrow \pi^*$ and $\pi_{nb} \rightarrow \pi^*$ transitions. Synchrotron radiation light sources have a much higher flux than conventional light sources at short wavelengths. Synchrotron radiation circular dichroism (SRCD) spectra reveal new features between 160 and 178 nm.[3−7] The most likely origin of these bands is charge transfer between peptide groups. Calculations[8−11] on *N*-acetylglycine-*N*′-methylamide, **1** (peptide groups are labeled 1 and 2 for future reference), showed that the interpeptide charge-transfer transitions occur between 120 and 175 nm and therefore may be observable in SRCD spectra of proteins.



1

The development of methods to compute the CD spectrum of a protein from its atomic coordinates has received much attention.[12−14] A widely used technique is the matrix method,[15] which treats a protein as a series of separate chromophores that couple with each other electrostatically. These chromophores can be characterized by semiempirical[16] or ab initio[17] calculations. We have used the latter on *N*-methylacetamide (NMA) to model the peptide $n \rightarrow \pi^*$ and $\pi_{nb} \rightarrow \pi^*$ transitions. Matrix method calculations using these parameters on a set of 47 proteins[12] reproduced many of the features of the CD spectra from 190 to 230 nm and performed especially well at 222 nm, where the mean residue ellipticity is highly correlated with the percentage helicity of the protein.

Recently, Gekko and co-workers reported the SRCD spectra of 31 proteins,[6,7] whose atomic coordinates are available in the Protein Data Bank.[18] This is a sufficiently large sample of spectra and structures to permit a meaningful assessment of the accuracy of first-principles CD calculations in the VUV and, through such calculations, to investigate the role of charge transfer.

Calculations on diamide **1** allow us to extend the CD calculations into the VUV. The transition properties of the charge-transfer transitions of **1** were generated from CASSCF/CASPT2 calculations[11] using MOLCAS.[19] For each transition density, the associated electrostatic potential was generated using MOLPRO,[20] and 64 point charges were fitted to reproduce the potential for each transition, with eight charges arranged at the corners of a cube with edges of length 0.1 au around each of the amide C, H, O, and N atoms. These parameters are available as Supporting Information.

Table 1. Energies in nm and Oscillator Strengths of the Charge-Transfer Transitions Used in This Study (from ref 11)

ϕ/°	ψ/°	$n_1 \rightarrow \pi^*_2$	$n_2 \rightarrow \pi^*_1$	$\pi_{nb1} \rightarrow \pi^*_2$	$\pi_{nb2} \rightarrow \pi^*_1$
180	180	155.0 (0.000)	132.6 (0.042)	171.0 (0.040)	131.8 (0.042)
−120	180	153.2 (0.029)	142.1 (0.056)	167.0 (0.015)	138.5 (0.047)
−60	180	156.1 (0.004)	135.9 (0.023)	170.9 (0.051)	136.4 (0.058)
−135	135	154.0 (0.001)	133.5 (0.045)	174.4 (0.168)	122.9 (0.006)
−120	120	155.7 (0.009)	135.2 (0.111)	173.6 (0.202)	126.4 (0.008)
−120	60	166.7 (0.215)	134.9 (0.021)	159.9 (0.056)	135.2 (0.005)
−74	−4	138.2 (0.022)	162.9 (0.091)	153.1 (0.002)	145.9 (0.010)
−62	−41	163.5 (0.121)	159.2 (0.068)	155.3 (0.011)	156.1 (0.042)
48	−57	156.9 (0.123)	162.1 (0.013)	164.4 (0.014)	152.7 (0.015)
−60	−60	170.7 (0.162)	154.5 (0.051)	163.1 (0.020)	145.1 (0.013)

Charge-transfer parameters were generated from the different geometries of **1** reported in our previous work,[11] plus a new one corresponding to Barlow and Thornton's model of the α−helix ($\phi = -62°$, $\psi = -41°$).[21] Initially, all of these geometries were considered, but only those presented in Table 1 proved useful and only they were used in this study. Mirror images of these were used to construct parameters for the left-handed helical and sheet regions of the Ramachandran plot. The fully extended, planar geometry has main-chain dihedral angles (ϕ, ψ) of (180°, 180°). None of the charge-transfer transitions in this geometry is particularly intense. For most of the geometries considered, ϕ and ψ are multiples of 60°. The structure with ($\phi = -135°$, $\psi = 135°$), is representative of a typical β-strand geometry and exhibits an intense $\pi_{nb1} \rightarrow \pi^*_2$ transition at 174.4 nm, that is, excitation of an electron in the nonbonding π orbital of the first peptide group to the antibonding π orbital of the second. The α-helical geometry ($\phi = -62°$, $\psi = -41°$) has an intense excitation at 163.5 nm, which could be the source of the band at 165 nm in the experimental spectrum of polyalanine.[22]

The CD spectra of the 31 proteins were calculated. The charge-transfer chromophores were assigned to the protein by calculating the ϕ and ψ angles along the protein backbone and choosing the chromophore with the closest values. We assess the quality of first-principles CD calculations by the correlation between the experi-

10.1021/ja0644125 CCC: $33.50

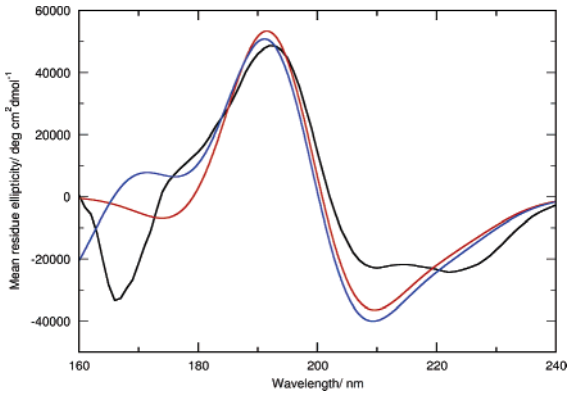


Figure 1. CD spectra of the α-helical protein myoglobin: experimental (black), calculated using only local transitions (red), calculated using local plus charge-transfer transitions (blue).

mental and calculated mean residue ellipticities at various wavelengths. The addition of charge-transfer chromophores to the matrix method makes little difference to the correlation for wavelengths above 190 nm. However, between 170 and 190 nm the agreement between the experimental and calculated spectra improves, most notably at 175 nm, where the Spearman rank correlation coefficient increases from 0.44 to 0.80.

The addition of charge-transfer transitions has the greatest effect on α-helical proteins, such as myoglobin. In the experimental spectrum (Figure 1) there is a positive shoulder at ∼175 nm, an intense negative peak at 165 nm, and a positive peak below 160 nm. In the calculated spectrum, the positive shoulder at 175 nm was reproduced. There is also a broad, intense negative peak at 150 nm. The spectra of typical β-sheet proteins, such as concanavalin A, are affected less by charge transfer. The negative peak at 175 nm is already described by the local $\pi_{nb} \rightarrow \pi^*$ transition. The matrix method does not predict the positive peak at ∼160 nm.

A matrix method calculation was performed on a 20 residue α-helix ($\phi = -62°$, $\psi = -41°$). The transitions in the line spectrum can each be readily identified with a particular excitation. Analysis of the component transitions reveals the following. The $n_1 \rightarrow \pi^*_2$ excitation gives an intense positive band at 164 nm. The $\pi_{nb1} \rightarrow \pi^*_2$ and $n_2 \rightarrow \pi^*_1$ transitions give intense negative bands at 155 and 159 nm, and the $\pi_{nb2} \rightarrow \pi^*_1$ transition gives a weak negative band at 155 nm. After these have been added to the spectrum from the local transitions, the $n_1 \rightarrow \pi^*_2$ band forms a positive shoulder at 169 nm, and the other charge-transfer transitions form a negative band at 153 nm. Inspection of the matrix elements from the Hamiltonian shows that the coupling between local and charge-transfer transitions is small.

The peptide $n' \rightarrow \pi^*$ and $\pi_b \rightarrow \pi^*$ transitions occur at wavelengths shorter than 160 nm, but coupling to these transitions may have an effect at longer wavelengths. Previous calculations including these transitions underestimated the intensity of the band at 190 nm because of a large coupling between the $\pi_b \rightarrow \pi^*$ and $\pi_{nb} \rightarrow \pi^*$.[12] The addition of charge-transfer chromophores to these calculations gave no improvement. We have also calculated CD spectra using Woody and Sreerama's semiempirical peptide parameters[16] and our charge-transfer parameters. This combination produces similar results to those already described. The only noteworthy difference is that, in α-helical proteins, there is a stronger coupling between the $n_1 \rightarrow \pi^*_2$ transition and the local $\pi_{nb} \rightarrow \pi^*$, but a weaker coupling between the $n_1 \rightarrow \pi^*_2$ transition and the local $n \rightarrow \pi^*$.

Detailed knowledge of the electronic structure and charge-transfer transitions in peptides will be an essential part of a more complete and quantitative interpretation of SRCD spectra of proteins. In addition, it may contribute to our understanding of the role of through-bond coupling in electron transfer processes in proteins. The calculations presented here significantly improve the accuracy of calculated protein CD spectra between 170 and 190 nm. They also give the first assignment of an interpeptide charge-transfer band in a CD spectrum. The new calculations perform less well at higher energies and do not predict the positive bands at ∼160 nm in α-helices and β-sheets. It is possible that these bands are caused by nonnearest-neighbor charge-transfer transitions, such as those across hydrogen bonds. Alternatively, these bands could be caused by coupling between charge-transfer transitions and the high-energy local $n' \rightarrow \pi^*$ and $\pi_b \rightarrow \pi^*$ transitions. Future work will explore this further.

Acknowledgment. We thank Professor Kunihiko Gekko for providing the experimental SRCD spectra in electronic form and the Engineering and Physical Science Research Council (EPSRC) for funding (Grant GR/T09224).

Supporting Information Available: Computational details of the matrix method, list of proteins used, transition energies, transition dipole moments, point charges and complete refs 19 and 20. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Hagarman, A.; Measey, T.; Doddasomayajula, R. S.; Dragomir, I.; Eker, F.; Griebenow, K.; Schweitzer-Stenner, R. *J. Phys. Chem. B* **2006**, *110*, 6979.
(2) Sreerama, N.; Woody, R. W. *Methods Enzymol.* **2004**, *383*, 318.
(3) Wallace, B. A. *Nat. Struct. Biol.* **2000**, *7*, 708.
(4) Wallace, B. A.; Janes, R. W. *Curr. Opin. Chem. Biol.* **2001**, *5*, 567.
(5) Miles, A. J.; Wallace, B. A. *Chem. Soc. Rev.* **2006**, *35*, 39.
(6) Matsuo, K.; Yonehara, R.; Gekko, K. *J. Biochem.* **2005**, *138*, 79.
(7) Matsuo, K.; Yonehara, R.; Gekko, K. *J. Biochem.* **2004**, *135*, 405.
(8) Serrano-Andrés, L.; Fülscher, M. P. *J. Am. Chem. Soc.* **1998**, *120*, 10912.
(9) Serrano-Andrés, L.; Fülscher, M. P. *J. Phys. Chem. B* **2001**, *105*, 9323.
(10) Gilbert, A. T. B.; Hirst, J. D. *J. Mol. Struct.* **2004**, *675*, 53.
(11) Oakley, M. T.; Bulheller, B. M.; Hirst, J. D. *Chirality* **2006**, *18*, 340.
(12) Hirst, J. D.; Colella, K.; Gilbert, A. T. B. *J. Phys. Chem. B* **2003**, *107*, 11813.
(13) Woody, R. W.; Koslowski, A. *Biophys. Chem.* **2002**, *101*, 535.
(14) Bode, K. A.; Applequist, J. *J. Am. Chem. Soc.* **1998**, *120*, 10938.
(15) Bayley, P. M.; Nielsen, E. B.; Schellman, J. A. *J. Phys. Chem.* **1969**, *73*, 228.
(16) Woody, R. W.; Sreerama, N. *J. Chem. Phys.* **1999**, *111*, 2844.
(17) Besley, N. A.; Hirst, J. D. *J. Am. Chem. Soc.* **1999**, *121*, 9636.
(18) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235.
(19) Andersson, K. et al.; *MOLCAS*, version 5.2; Lund University, Sweden, 2001.
(20) Amos, R. D. et al.; *Molpro, a package of ab initio programs*, version 2002.1; University of Birmingham: Birmingham, U.K., 2002.
(21) Barlow, D. J.; Thornton, J. M. *J. Mol. Biol.* **1988**, *201*, 601.
(22) Bensing, J. L.; Pysh, E. S. *Macromolecules* **1971**, *4*, 659.

JA0644125