

Flow Linear Dichroism of Some Prototypical Proteins

Benjamin M. Bulheller,[†] Alison Rodger,[‡] Matthew R. Hicks,[‡] Timothy R. Dafforn,[§]
 Louise C. Serpell,^{||} Karen E. Marshall,^{||} Elizabeth H. C. Bromley,[⊥]
 Patrick J. S. King,[⊥] Kevin J. Channon,[⊥] Derek N. Woolfson,^{⊥, #} and
 Jonathan D. Hirst^{*†}

*School of Chemistry, University of Nottingham, University Park, Nottingham NG7 2RD, U.K.,
 Department of Chemistry, University of Warwick, Coventry CV4 7AL, U.K., School of
 Biosciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, U.K., Department of
 Biochemistry, School of Life Sciences, University of Sussex, Falmer BN1 9QG, U.K., School of
 Chemistry, University of Bristol, Bristol BS8 1TS, U.K., and Department of Biochemistry,
 University of Bristol, Bristol BS8 1TD, U.K.*

Received April 3, 2009; E-mail: jonathan.hirst@nottingham.ac.uk

Abstract: Flow linear dichroism (LD) spectroscopy provides information on the orientation of molecules in solution and hence on the relative orientation of parts of molecules. Long molecules such as fibrous proteins can be aligned in Couette flow cells and characterized using LD. We have measured using Couette flow and calculated from first principles the LD of proteins representing prototypical secondary structure classes: a self-assembling fiber and tropomyosin (all- α -helical), FtsZ (an $\alpha\beta$ protein), an amyloid fibril (β -sheet), and collagen [poly(proline)II helices]. The combination of calculation and experiment allows elucidation of the protein orientation in the Couette flow and the orientation of chromophores within the protein fibers.

Introduction

Linear dichroism (LD) spectroscopy, a differential absorption technique closely related to circular dichroism (CD) spectroscopy, is attracting renewed attention.¹ LD experiments require the samples to be aligned in some way, and in the case of polymers, the alignment is typically provided by two methods: electric field^{2,3} and flow. Studies of the dichroism of polymers^{2,4,5} and proteins^{6,7} oriented in a flow have an extensive history; early investigations focused on the use of double refraction of flow to characterize myosin.^{8,9} The focus in this work is on flow LD (henceforth termed LD). LD spectra have provided information on how molecules such as gramicidin in membranes, α_1 -antitrypsin, and the amyloid-related protein A β_{1-42} align in shear flow.¹⁰ LD has been used to detect and characterize fibril formation by amyloid peptides¹¹ and can furnish insight into

how agents dock to enzymes or fibers.^{12–14} The technique has been used to study the polymerization of protein fibers such as G-actin into F-actin,¹⁰ the orientation of bacteriorhodopsin in liposomes,¹⁵ and the binding of myosin to actin¹⁶ as well as to measure the kinetics of insertion of peptides into lipid membranes.^{17,18} However, to exploit fully the information available from LD measurements, one needs to understand the relationship between protein conformation and LD spectra, and ideally, one would like to be able to construct structural models and use them to calculate the electronic structure directly.¹⁰ In this paper, we report on our new calculation methodology for achieving this goal.

LD measurements require the difference in absorption of light polarized parallel ($A_{||}$) and perpendicular (A_{\perp}) with respect to an orientation direction:

$$LD = A_{||} - A_{\perp} \quad (1)$$

The related technique of electronic CD spectroscopy is widely used for secondary structure determination of proteins.¹⁹ In a

[†] University of Nottingham.

[‡] University of Warwick.

[§] University of Birmingham.

^{||} University of Sussex.

[⊥] School of Chemistry, University of Bristol.

[#] Department of Biochemistry, University of Bristol.

- (1) Dafforn, T. R.; Rodger, A. *Curr. Opin. Struct. Biol.* **2004**, *14*, 541–546.
- (2) Troxell, T. C.; Scheraga, H. A. *Macromolecules* **1971**, *4*, 519–527.
- (3) Nordén, B. *Appl. Spectrosc. Rev.* **1978**, *14*, 157–248.
- (4) Wada, A.; Kozawa, S. *J. Polym. Sci., Part A* **1964**, *2*, 853–864.
- (5) Troxell, T. C.; Scheraga, H. A. *Macromolecules* **1971**, *4*, 528–539.
- (6) Cerf, R.; Scheraga, H. A. *Chem. Rev.* **1952**, *51*, 185–261.
- (7) Hofrichter, J.; Eaton, W. A. *Annu. Rev. Biophys. Bioeng.* **1976**, *5*, 511–560.
- (8) von Muralt, A. L.; Edsall, J. T. *J. Biol. Chem.* **1930**, *89*, 351–386.
- (9) Edsall, J. T.; Mehl, J. W. *J. Biol. Chem.* **1940**, *133*, 409–429.
- (10) Dafforn, T. R.; Rajendra, J.; Halsall, D. J.; Serpell, L. C.; Rodger, A. *Biophys. J.* **2004**, *86*, 404–410.
- (11) Adachi, R.; Yamaguchi, K.; Yagi, H.; Sakurai, K.; Naiki, H.; Goto, Y. *J. Biol. Chem.* **2007**, *282*, 8978–8983.

- (12) Morimatsu, K.; Takahashi, M. *Anal. Biochem.* **2006**, *358*, 192–198.
- (13) Esbjörner, E. K.; Caesar, C. E. B.; Albinsson, B.; Lincoln, P.; Nordén, B. *Biochem. Biophys. Res. Commun.* **2007**, *361*, 645–650.
- (14) Small, E.; Marrington, R.; Rodger, A.; Scott, D. J.; Sloan, K.; Roper, D.; Dafforn, T. R.; Addinall, S. G. *J. Mol. Biol.* **2007**, *369*, 210–221.
- (15) Rajendra, J.; Damianoglou, A.; Hicks, M.; Booth, P. J.; Rodger, P. M.; Rodger, A. *Chem. Phys.* **2006**, *326*, 210–220.
- (16) Rodger, A.; Marrington, R.; Geeves, M. A.; Hicks, M.; de Alwis, L.; Halsall, D. J.; Dafforn, T. R. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3161–3171.
- (17) Hicks, M. R.; Damianoglou, A.; Rodger, A.; Dafforn, T. R. *J. Mol. Biol.* **2008**, *383*, 358–366.
- (18) Ennaceur, S. M.; Hicks, M. R.; Pridmore, C. J.; Dafforn, T. R.; Rodger, A.; Sanderson, J. M. *Biophys. J.* **2009**, *96*, 1399–1407.
- (19) Berova, N.; Nakanishi, K.; Woody, R. W. *Circular Dichroism: Principles and Applications*, 2nd ed.; Wiley-VCH: New York, 2000.

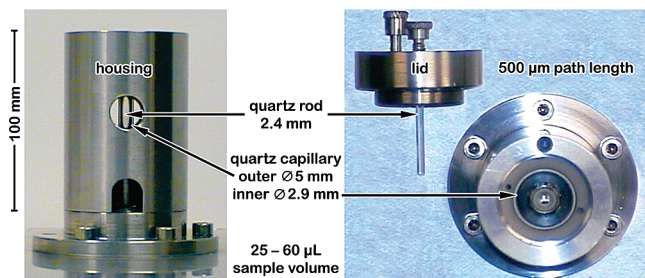


Figure 1. Photos of a microvolume capillary LD cell with labels indicating how the components are assembled.

CD spectrometer, a sample is irradiated with left- and right-circularly polarized light, and because of the chirality of the structural motifs, the two types of polarized light are absorbed to different extents. The differential absorption is wavelength-dependent and allows one to distinguish between different protein folds and estimate the helical content.²⁰ Because of this and other advantages, CD spectrometers are widely available, and much effort has been directed toward understanding the theoretical basis of the technique.²¹

Since circularly polarized light consists of two linearly polarized beams traveling with optical retardation, any CD spectrometer can be used to measure LD. Linearly polarized light may be produced by adding a quarter-wave plate or increasing the voltage across the photoelastic modulator.²² The latter option has been used to obtain the data reported in this work. Despite the close relationship of the techniques, the measurement of LD is more challenging than that of CD. The latter can be measured for a solution in a cuvette, but in order to record an LD spectrum, the sample has to be oriented.²² Currently, the most successful method for aligning fibrous and membrane proteins in solution has proved to be a Couette flow cell, in which the sample is forced to flow through the annular gap formed by a cylindrical cuvette with a quartz rod in the middle (Figure 1) while either the cylinder or the quartz rod rotates.¹⁶ The light is incident on the cell and passes through the annular gap (and hence through the sample), the rod, and then the gap (and sample) on the other side of the cell. Improved cell designs have reduced the amount of sample required to as little as 30 μL , making the technique valuable even on a biochemical volume scale.²³

Globular proteins are usually soluble in aqueous solution and can be crystallized. This facilitates the use of X-ray crystallography to aid in the determination of the structure and orientation. However, fibrous proteins such as FtsZ¹⁴ and tubulin^{16,24} often do not crystallize. Alignment in Couette flow relies on high aspect ratios of the molecules, which is precisely what DNA and fibrous proteins have. LD is thus a valuable tool for obtaining information about the orientation of such molecules, especially if structural information is available about the monomer units (the DNA bases or the units of the protein fiber). For most protein fibers, fiber formation allows an LD measurement, even though the monomers are too small to be

aligned individually, because of the regular order of the monomers in a fiber. Thus, LD can be used to follow fiber formation, as only the fibers give a signal. LD has also been used in the past in the context of electric dichroism to probe the structure and electrical properties of polymers.^{2,5,25}

The interpretation of LD spectra is challenging, since overlapping transitions lead to the cancellation of bands and complicate the analysis of experimental spectra beyond what is possible by qualitative inspection. This is especially true if the protein contains a mixture of different secondary structural elements. Calculations allow one to study both single and multiple transitions as a function of geometry and orientation, and the combination of experiment and calculation can lead to mechanistic and structural insights that cannot be derived from either in isolation. Previously,²⁶ transition polarizations and oscillator strengths of ideal (i.e., perfectly regular) α -helical and β -sheet model peptides were calculated from first principles. In this work, we show the utility of the approach through its application to several proteins exemplifying prototypical secondary structural classes. The study has been made possible by the adaptation of the matrix method, which is used for calculating CD spectra, to the calculation of LD spectra. This new application was implemented in the web interface DichroCalc,²⁷ which now offers both CD and LD calculations, facilitating comparison between experimental spectra and spectra calculated from a (perhaps hypothetical model or an experimentally determined) protein structure. For each protein of interest, one considers the LD calculated for different proposed structures and uses the experiment/calculation comparison to select the experimental fiber geometry. For the protein fibers used in this work, atomic-level structures were available for the monomeric units and in some cases for dimers or higher oligomers, which facilitated the process. If no information on how the fiber units combine to make a polymer is available, then a large series of calculations is required.

Methods

A successful methodology for computing the CD spectrum of a protein from its structure is the (exciton) matrix method.^{28–30} The atomic Cartesian coordinates required for the calculation can be downloaded, for example, from the RCSB Protein Data Bank (PDB).³¹ The matrix method uses monopole charges, which can be obtained from quantum-mechanical ab initio calculations, to represent the electrostatic potential of the electronic transitions of the chromophores within the protein.³² The Hamiltonian matrix needed to calculate the wave function is constructed from the Coulombic interactions between those charges. Such calculations have been reviewed in detail elsewhere.³³ For LD, the calculation is analogous, but we extract the absorbance spectra for the parallel and perpendicular orientations of the protein, which are calculated from the electric transition dipole moments. The z axis is defined as the orientation axis (i.e., the parallel direction); an average over

(20) Kelly, S. M.; Jess, T. J.; Price, N. C. *Biochim. Biophys. Acta* **2005**, *1751*, 119–139.

(21) Schellman, J. A. *Chem. Rev.* **1975**, *75*, 323–331.

(22) Rodger, A.; Nordén, B. *Circular Dichroism and Linear Dichroism*, 1st ed.; Oxford University Press: Oxford, U.K., 1997.

(23) Marrington, R.; Dafforn, T. R.; Halsall, D. J.; MacDonald, J. I.; Hicks, M.; Rodger, A. *Analyst* **2005**, *130*, 1608–1616.

(24) Marrington, R.; Seymour, M.; Rodger, A. *Chirality* **2006**, *18*, 680–690.

(25) Troxell, T. C.; Scheraga, H. A. *Biochem. Biophys. Res. Commun.* **1969**, *35*, 913–119.

(26) Rodger, A.; et al. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4051–4057.

(27) Bulheller, B. M.; Hirst, J. D. *Bioinformatics* **2009**, *25*, 539–540.

(28) Bayley, P. M.; Nielsen, E. B.; Schellman, J. A. *J. Phys. Chem.* **1969**, *73*, 228–243.

(29) Woody, R. W.; Tinoco, I. *J. Chem. Phys.* **1967**, *46*, 4927–4945.

(30) Woody, R. W. *J. Chem. Phys.* **1968**, *49*, 4797–4806.

(31) 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*, 235–242.

(32) Besley, N. A.; Hirst, J. D. *J. Am. Chem. Soc.* **1999**, *121*, 9636–9644.

(33) Bulheller, B. M.; Rodger, A.; Hirst, J. D. *Phys. Chem. Chem. Phys.* **2007**, *9*, 2020–2035.

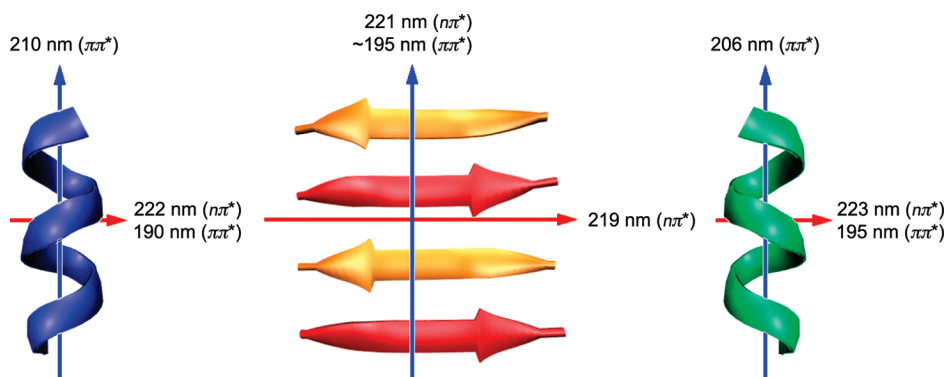


Figure 2. Polarizations of transitions in (left) an α -helix, (middle) a β -sheet, and (right) a P_{II} -helix.

the absorption (A) in the x and y directions accounts for the perpendicular absorption:

$$A_{\parallel} = A_z \quad A_{\perp} = \frac{1}{2}(A_x + A_y) \quad (2)$$

The LD spectrum is calculated by combining eqs 1 and 2. Through the use of standard PDB files to give the geometries of the monomer units, the calculations can be performed via the web interface DichroCalc, which is available at <http://comp.chem.nottingham.ac.uk/dichrocalc>.³³ The main result required from the LD calculations of protein fibers is usually the relative orientation of the monomeric proteins within the fiber. DichroCalc calculates the LD spectrum corresponding to all possible orientations of the fiber components, so the output is more extensive than for CD.

Equation 1 indicates that transitions polarized parallel to the fiber axis lead to a positive band in the LD spectrum, whereas transitions with a perpendicular polarization lead to a negative band. One can therefore anticipate the qualitative features of the LD spectra of regular secondary structure elements by considering the allowed combinations of coupled local $\pi \rightarrow \pi^*$ transitions, which are polarized along the carbonyl bonds, to determine the net polarization for these transitions. For an α -helix, it is known from coupled-oscillator analysis that the ~ 195 nm amide $\pi \rightarrow \pi^*$ transitions are split into two components, one polarized parallel to the helix axis at ~ 210 nm and another polarized perpendicular at ~ 190 nm (Figure 2). An analogous consideration of $n \rightarrow \pi^*$ transitions has been developed through calculations on peptides with structures built into ideal, regular helical, and strand conformations. The $n \rightarrow \pi^*$ transition at ~ 222 nm is polarized perpendicular to the helix axis.¹⁰ For a perfectly parallel helix (such as in a coiled coil, e.g., tropomyosin), one would therefore expect negative bands at ~ 190 and ~ 222 nm in the LD spectrum. These regions are dominated by transitions polarized perpendicular to the helix axis, resulting in a negative LD signal according to eq 1. Conversely, the parallel component of the $\pi \rightarrow \pi^*$ transition would result in a positive band at ~ 210 nm (though in reality it is usually dominated by the two bands on either side of it and thus appears as a small minimum). In β -sheets, the $\pi \rightarrow \pi^*$ transition is polarized parallel to the long axis of a single strand (perpendicular to the sheet) as a result of the lack of the parallel component at ~ 210 nm.²⁶ In actuality, $\sim 50\%$ of residues in proteins are not in regular secondary structure motifs, and proteins often comprise a complex combination of secondary structures. Therefore, we have adopted a more detailed computational approach that deals with the experimentally diverse range of protein conformations.

Table 1. Electronic Transitions Included in the Calculations of the LD Spectra and Their Corresponding Transition Wavelengths and Electric Transition Dipole Moments

transition	wavelength (nm)	electric transition dipole moments (x, y, z) ^a (D)
Backbone		
$n \rightarrow \pi^*$	220	(0.0000, 0.0000, 0.2260)
$\pi \rightarrow \pi^*$	193	(2.0946, 2.0200, 0.0000)
Tryptophan		
1B_a	196	(3.6291, 2.1261, 0.0000)
1B_b	204	(3.6321, -2.4661, 0.0000)
1L_a	263	(-1.0939, 1.2255, 0.0000)
1L_b	283	(0.7198, 0.6912, 0.0000)
Phenylalanine		
1B_b	190	(-0.0009, 5.7320, 0.0000)
1B_a	191	(-5.7153, -0.0025, 0.0000)
1L_a	209	(-0.0018, 0.0000, 0.0000)
1L_b	263	(0.0000, 0.0171, 0.0000)
Tyrosine		
1B_b	193	(-0.0220, -5.3279, 0.0000)
1B_a	196	(5.7956, 0.0015, 0.0000)
1L_a	216	(1.3859, -0.1582, 0.0000)
1L_b	274	(-0.0563, -0.6972, 0.0000)

^a Defined with respect to coordinate systems provided in the Supporting Information of refs 32 and 34.

The electronic excitations of the chromophores are described by parameters derived from ab initio calculations^{32,35} that have been tested on several sets of proteins.^{33,36} Chromophore parameters are available for the peptide³² and aromatic side chain transitions.³⁵ The calculations consider two local peptide backbone transitions, the $n \rightarrow \pi^*$ transition (at 220 nm) and the $\pi \rightarrow \pi^*$ transition (at 193 nm), and electronic transitions associated with the side chains of tryptophan, phenylalanine, and tyrosine. For convenience, we summarize the transitions and their associated properties in Table 1. Each transition density is represented by 20 or more charges fitted to reproduce the ab initio electrostatic potential. Full details of these charges for each excitation (ground state to excited state and between different excited states) are available in the Supporting Information of earlier publications^{32,34} as well as online from the DichroCalc web interface.²⁷ The interaction terms in the exciton Hamiltonian matrix, corresponding to the mixing of electronic states, are calculated from the Coulombic interactions between these sets of charges. The unitary transformation that diagonalizes the Hamiltonian matrix is used to transform the transition dipole moments of the individual chromophores. The resultant electric transition dipole moments are used to compute the absorption and LD spectra. Inclusion of higher-energy local peptide transitions and transitions involving charge transfer between covalently neighboring peptide groups³⁴ made little quantitative difference, and the calculations reported here do not include them.

The $n \rightarrow \pi^*$ band is evident in several of the experimental absorption and LD spectra. The calculations do not reproduce this band, despite the reasonable agreement observed between experimental and calculated CD spectra. The discrepancy in the case of

(34) Bulheller, B. M.; Miles, A. J.; Wallace, B. A.; Hirst, J. D. *J. Phys. Chem. B* **2008**, *112*, 1866–1874.

(35) Rogers, D. M.; Hirst, J. D. *J. Phys. Chem. A* **2003**, *107*, 11191–11200.

(36) Hirst, J. D.; Colella, K.; Gilbert, A. T. B. *J. Phys. Chem. B* **2003**, *107*, 11813–11819.

the LD spectra indicates that the calculated transitions at ~ 222 nm do not have a sufficiently large electric transition dipole moment, which in turn suggests perhaps that the mixing between the $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ states is significantly underestimated for the conformations in question. However, investigative calculations using other parameter sets for which the coupling between these states is stronger³⁷ did not improve matters (data not shown).

The calculated far-UV spectra consider only the peptide chromophores, while the near-UV calculations, where experimental data are available, take aromatic side chains into account. The intensity observed experimentally depends on the orientation factor (S) of the sample. If the sample is not oriented at all, then $S = 0$ and no LD spectrum is observed. If all of the proteins are perfectly oriented (as is the case, for example, in a crystal), $S = 1$ and the intensity is greatest. Adachi et al.¹¹ suggested that $S \approx 0.1$ for the amyloid fibrils that they studied. However, the orientation factor is difficult to estimate a priori, and we do not have a solid basis for doing so in the current study. The degree of orientation is related to the velocity gradient between the walls of the rotating Couette cylinders. The angular velocity has to be chosen carefully to maintain a laminar flow (or Couette flow). Turbulence occurs at too high a speed, leading to increased scattering and loss of light.²³ Since the velocity is therefore limited and the alignment of the sample depends on it, perfect orientation of the molecules cannot be achieved; thus, the LD signal is diminished. In addition, samples subjected to shear flow are more prone to form bubbles than stationary samples in a cuvette, increasing the experimental challenge. In contrast, the calculations were carried out on a single defined orientation of a protein, for which $S = 1$ by definition. In the absence of a rigorously defined scaling factor for the calculations, the experimental LD spectra were scaled to the calculated intensity for the purpose of comparison.

Couette flow experiments were performed using a Jasco J815 spectropolarimeter modified for LD measurements (Jasco UK, Great Dunmow, U.K.). Spectra were measured over the appropriate wavelength range using a bandwidth of 2 nm, a data pitch of 0.2 nm, a scanning speed of 100 nm min⁻¹, and a response time of 1 s. Where the signal was small relative to the instrument noise, four or eight spectra were averaged to improve the signal-to-noise ratio. Measurements were carried out at room temperature (21 °C), and baseline spectra of the buffer were subtracted from the sample spectra. The rotation speed of the samples, which gives rise to the velocity gradient responsible for the alignment of the samples, was ~ 3000 rpm. The microvolume Couette cell used had a path length of 0.5 mm and was built in-house; equivalent models are commercially available (Kromatek, Great Dunmow, U.K.). All sample volumes were 60 μ L. Tropomyosin was obtained from Sigma-Aldrich, and FtsZ was supplied by Xi Cheng and R. Pacheco Gomez (University of Warwick). The collagen was guinea pig type-I, a heterotrimer comprising two $\alpha 1$ chains and one $\alpha 2$ chain, supplied by R. Visse (Imperial College London).

In the following, we report the first calculations of protein LD spectra for a number of possible orientations of a given fiber and compare them to data from Couette flow experiments. The combination of calculation and experiment can be used to provide useful information on the assembly of the protein into a fibrous form. In the examples chosen for this work, the answers are essentially known (as in coiled coils) or we have strong indications as to the fiber structure from monomer or small-oligomer crystal structures (FtsZ and an amyloid-like fibril), since the main aim of the work is to show that LD can be used in this way. Comparison with theoretical results can provide insight into protein orientation that is not easily deducible from an analysis of experimental spectral data alone.

Results

For each protein, the calculated LD spectrum for the monomers assembled into fibers in accord with literature data

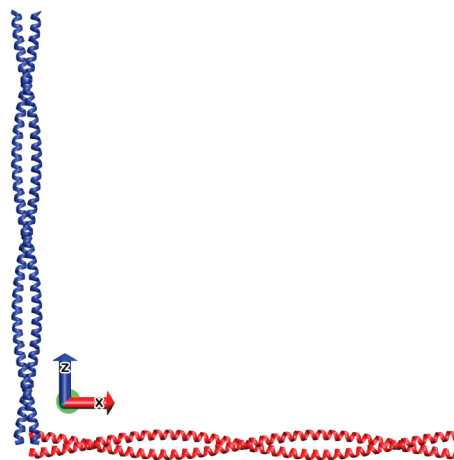


Figure 3. Orientations of the SAF used for the calculations. The colors indicate parallel (blue) and perpendicular (red) orientations.

and aligned perfectly parallel to the flow direction is given. For comparison, the LD of a fiber assembled to be equivalent to the parallel fiber but oriented perpendicular to the flow direction is also shown. The z axis (vertical, blue) represents the parallel orientation with respect to the fiber long axis in the Couette flow; x (red) and y equivalently represent the perpendicular orientation. The colors of the structures indicate parallel (blue, aligned with z) and perpendicular (red, aligned with x) orientations and are used accordingly for calculated spectra throughout the study. Since the x and y coordinates are taken into account as a rotational average (eq 2), alignments with these two axes are equivalent. All of the 3D plots were created using VMD.³⁸ Next to most of the plots showing LD spectra for the parallel and perpendicular orientations of the protein, we show the calculated absorbance spectra that were used to generate the LD spectra. In each case, the top panel of absorbance plots contains the spectra for the parallel orientation, and the results for perpendicular alignment are shown in the bottom panel. For each orientation, two calculated absorbance spectra are required: the absorption in the parallel direction (A_{\parallel} , along the z axis), and the absorption in the perpendicular direction (A_{\perp} , the rotational average according to eq 2). Additionally, the full absorbance spectrum for each orientation is given for comparison with experiment.

Self-Assembling Fibers. Self-assembling fibers (SAFs) have well-understood behavior in the Couette flow. They are designed peptides that assemble into regular fibers and were developed to investigate the fibrillogenesis of proteins.³⁹ Two synthetic peptides, each 28 residues long, assemble into a regular coiled coil (Figure 3, which shows both the parallel and perpendicular orientations). The coiled coil is staggered, leaving “sticky ends” that allow the coils to assemble longitudinally into extended fibers. Lateral bundling of the coiled coils into fibers creates stiff rods tens of micrometers in length. The rigidity of the fibers ensures minimal deformation in shear flow. This, combined with the known orientation of coiled coils in fibers,⁴⁰ which is parallel to the long axis of the thickened fibers, makes SAFs a useful

(38) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33–38.

(39) Pandya, M. J.; Spooner, G. M.; Sunde, M.; Thorpe, J. R.; Rodger, A.; Woolfson, D. N. *Biochemistry* **2000**, *39*, 8728–8734.

(40) Papapostolou, D.; Smith, A. M.; Atkins, E. D. T.; Oliver, S. J.; Ryadnov, M. G.; Serpell, L. C.; Woolfson, D. N. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 10853–10858.

(37) Woody, R. W.; Sreerama, N. *J. Chem. Phys.* **1999**, *111*, 2844–2845.

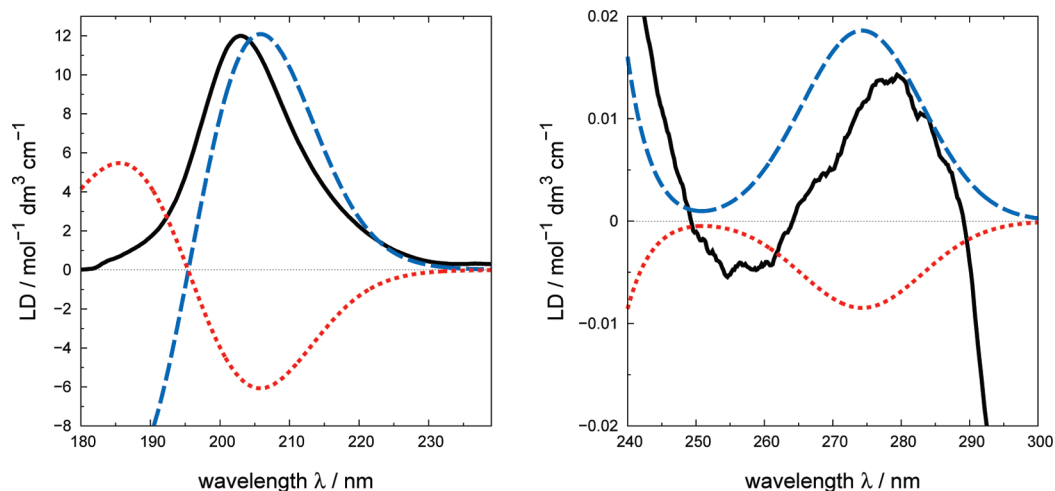


Figure 4. Experimental LD spectrum (solid line) of a SAF aligned in a Couette flow cell and calculated spectra of parallel (dashed) and perpendicular (dotted) orientations: (left) far-UV considering the peptide chromophores; (right) near-UV considering backbone and aromatic side chains. The experimental spectra have been scaled to the calculated intensities for comparison.

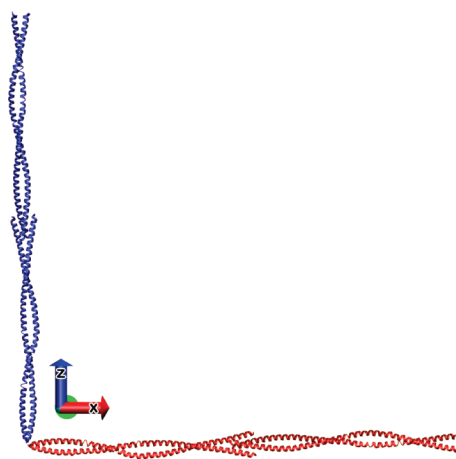


Figure 5. Orientations of tropomyosin used for the calculations. The colors indicate parallel (blue) and perpendicular (red) orientations.

model case. If the α -helices are aligned along the z axis (i.e., parallel to the flow direction), the calculations result in a positive maximum around 205 nm (Figure 4, left), while a perpendicular orientation leads to a minimum at this wavelength. Comparison with experiment indicates that the α -helices of the SAF are indeed oriented parallel to the fiber axis. The positive band around 280 nm (Figure 4, right) is caused by the side chains of one tyrosine residue per peptide monomer. The calculations considering aromatic side chains are sensitive to the regular repeating structure of the SAF peptides, and the orientation predicted from the near-UV data is consistent with the results in the far-UV. The agreement of theory and experiment for this well-understood model case is promising, and we turn now to natural proteins representing the most important secondary structure classes.

Tropomyosin. Tropomyosin (PDB code 2D3E) is a pure α -helical protein consisting of four chains that form two coiled coils (Figure 5).⁴¹ It has a helical content of 92% according to the DSSP algorithm.⁴² Filaments are formed by interactions of overlapping ends of the coils, thereby extending the helix

longitudinally.⁴³ The left-handed superhelix created by the consecutive coiled coils is 38–40 nm long with a diameter of less than 2 nm.⁴⁴ The rod-shaped constitution makes it a good candidate for LD experiments. The near-UV and $n \rightarrow \pi^*$ region of the far-UV LD spectrum of tropomyosin has been characterized previously,³⁹ and LD has also proven useful in the investigation of tropomyosin binding to actin.¹⁶ While SAFs form thick, stiff rods, tropomyosin is a natural polypeptide with greater flexibility, and the rods of tropomyosin are orders of magnitude thinner than SAF fibers. For calculations, conformational flexibility potentially poses an additional challenge. Calculations were carried out with tropomyosin aligned with the z axis (perfectly parallel to the orientation direction) or the x -axis (perpendicular to the orientation direction) (Figure 6). The position of the band caused by the $\pi \rightarrow \pi^*$ transition is reproduced well and shows a positive sign for the protein oriented parallel to the orientation axis. This is in agreement with the experimental spectrum and the expected orientation of tropomyosin in the Couette flow. The experimental intensity at 222 nm is larger for tropomyosin than for the SAF peptide, although the secondary structures of the two are very similar. Interestingly, the calculation does not mimic the intensity enhancement.

FtsZ. FtsZ (PDB code 1FSZ)⁴⁵ is an $\alpha\beta$ -protein and the bacterial homologue of tubulin.⁴⁶ The formation of fibrous FtsZ is a key step in cytokinesis. FtsZ contains two mainly parallel β -sheets surrounded by α -helices. The secondary structure content, estimated using the DSSP algorithm,⁴² is 44% helical and 25% β -sheet. The protein comprises one chain with guanosine diphosphate bound in a polar pocket. This study focuses on the orientation of the FtsZ protein itself, so the nucleotide was neglected in the calculations. FtsZ forms fibers, and the orientation of the monomers within the long axis of the fiber is not unambiguously resolved, though a literature dimer structure suggests a probable fiber geometry as discussed

(43) Phillips, G. N., Jr.; Fillers, J. P.; Cohen, C. *J. Mol. Biol.* **1986**, *192*, 111–131.

(44) Bode, K. A.; Applequist, J. *Biopolymers* **1997**, *42*, 855–860.

(45) Leung, A. K. W.; Lucile White, E.; Ross, L. J.; Reynolds, R. C.; DeVito, J. A.; Borhani, D. W. *J. Mol. Biol.* **2004**, *342*, 953–970.

(46) Nogales, E.; Downing, K. H.; Amos, L. A.; Löwe, J. *Nat. Struct. Biol.* **1998**, *5*, 451–458.

(41) Nitanai, Y.; Minakata, S.; Maeda, K.; Oda, N.; Maéda, Y. *Adv. Exp. Med. Biol.* **2007**, *592*, 137–151.

(42) Kabsch, W.; Sander, C. *Biopolymers* **1983**, *22*, 2577–2637.

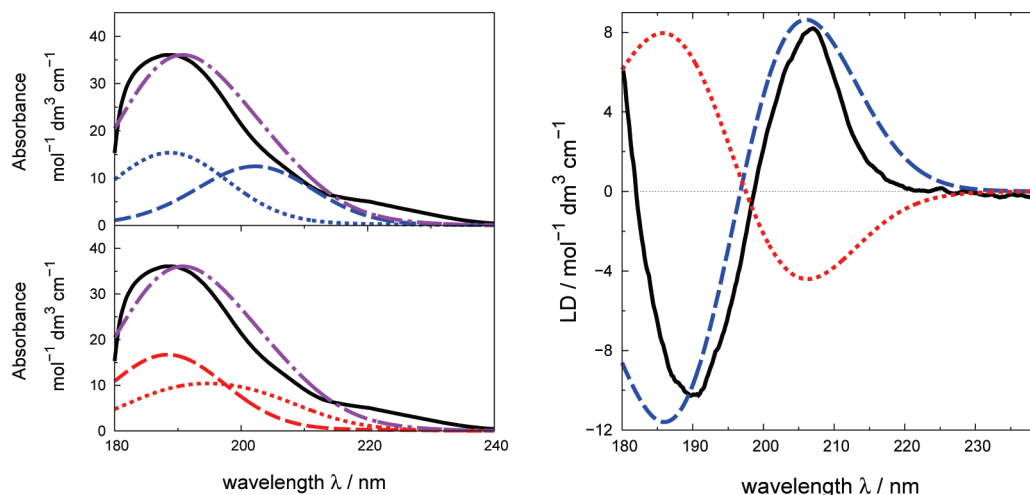


Figure 6. (left) Experimental absorbance spectrum of tropomyosin (solid line) and calculated spectra for parallel (top, blue) and perpendicular (bottom, red) orientations: parallel absorption (dashed), perpendicular absorption (dotted), and full absorption (dot-dashed). (right) Experimental LD spectrum (solid line) of tropomyosin aligned in a Couette flow cell and calculated spectra for parallel (dashed) and perpendicular (dotted) orientations. The experimental spectrum has been scaled to the calculated intensity for comparison.

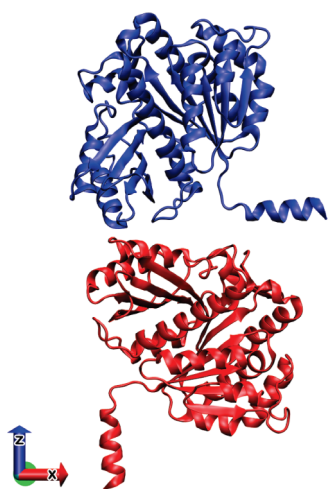


Figure 7. Orientation of FtsZ used for the calculations. The colors indicate parallel (blue) and perpendicular (red) orientations.

below.⁴⁷ The mechanism of polymerization has been studied using LD,^{14,48} and such studies would benefit from a direct comparison between structural models and LD spectra. In contrast to SAF and tropomyosin, both of which have high aspect ratios, FtsZ is globular, so the definition of parallel and perpendicular orientations on the basis of the monomer structure alone is ambiguous. Therefore, the protein was aligned with the z and x axes, respectively, according to the supposed orientation in the fiber determined from X-ray diffraction measurements on a dimer by Löwe and Amos⁴⁷ (Figure 7). In this work, we considered FtsZ, in the first instance, without the guidance provided by the dimer structure (which might not be relevant for the polymer). Therefore, a range of calculations with the monomers oriented in all possible ways was performed. A summary of the calculations for all of the analyzed orientations is available in the Supporting Information. The orientation suggested by Löwe and Amos proved not to be unique in producing the correct spectrum, since the intensity could not

be taken into account. However, in a truly unknown case, the vast majority of orientations eliminated by the LD experiment and calculations would enable a constrained molecular modeling approach to be undertaken to establish the most likely fiber structure. The top structure (blue) in Figure 7 depicts the Löwe–Amos orientation within the fiber. The monomer is therefore oriented parallel in the Couette flow. The calculated LD spectra show a positive maximum at 206 nm for the parallel orientation and a negative band for the structure aligned perpendicular to the orientation axis (Figure 8). The positive band observed experimentally in this region agrees with the calculation of the parallel structure and therefore confirms that the orientation of FtsZ fibers is indeed in accord with that extrapolated from the experimental dimer structure.

Amyloid Fiber. In contrast to helical structures, the theoretical origins of the electronic structure of β -sheet structures are less well understood, even at a qualitative level. For example, Nesloney and Kelly⁴⁹ remarked in 1996 that agreement on a standard β -sheet CD spectrum has yet to be attained and that the textbook spectrum, with a maximum at 195 nm and a minimum at 218 nm, is usually that of poly-L-Lys in an aggregated β -sheet. Previous theoretical studies of β -sheet CD have reported significant sensitivity to the precise structure and regularity of the protein chain and also to the location of the transition dipole moment on the peptide chromophore.⁵⁰ More recent literature⁵¹ suggests that exciton coupling theory predicts a negative band at 175 nm and a positive band 195 nm. However, these theoretical studies for β -sheets are certainly less definitive than for helical structures.

Amyloid fibers have been linked to conditions such as Parkinson's and Alzheimer's diseases and are therefore a focus of current research.^{52–54} A common feature in amyloid-like fibrils is the cross- β pattern, which is characterized by pairs of

(47) Löwe, J.; Amos, L. A. *EMBO J.* **1999**, *18*, 2364–2371.

(48) Marrington, R.; Small, E.; Rodger, A.; Dafforn, T. R.; Addinall, S. G. *J. Biol. Chem.* **2004**, *279*, 48821–48829.

(49) Nesloney, C. L.; Kelly, J. W. *Bioorg. Med. Chem.* **1996**, *4*, 739–766.

(50) Woody, R. W. *Tetrahedron: Asymmetry* **1993**, *4*, 529–544.

(51) Sreerama, N.; Woody, R. W. *Methods Enzymol.* **2004**, *383*, 318–351.

(52) Westermark, P. *FEBS J.* **2005**, *272*, 5942–5949.

(53) Westermark, P.; Benson, M. D.; Buxbaum, J. N.; Cohen, A. S.; Frangione, B.; Ikeda, S.-I.; Masters, C. L.; Merlini, G.; Saraiva, M. J.; Sipe, J. D. *Amyloid* **2005**, *12*, 1–4.

(54) Trojanowski, J. Q.; Lee, V. M.-Y. *Neurotoxicology* **2002**, *23*, 457–460.

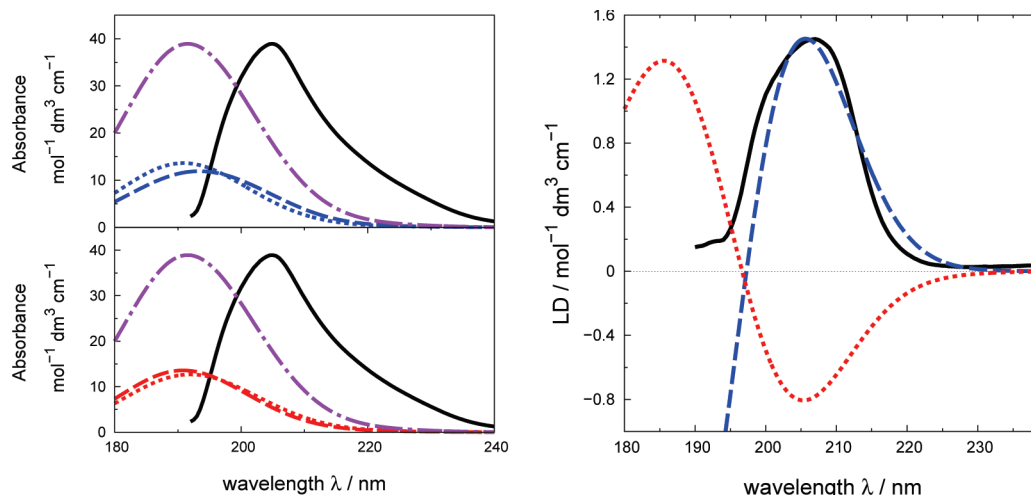


Figure 8. (left) Experimental absorbance spectrum of FtsZ (solid line) and calculated spectra for parallel (top, blue) and perpendicular orientations (bottom, red): parallel absorption (dashed), perpendicular absorption (dotted), and full absorption (dot-dashed). (right) Experimental LD spectrum (solid line) of FtsZ aligned in a Couette flow cell and calculated spectra for parallel (dashed) and perpendicular (dotted) orientations. The experimental spectrum has been scaled to the calculated intensity for comparison.

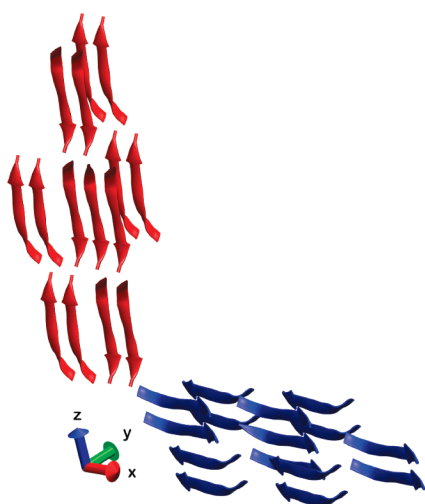


Figure 9. Orientation of the amyloid fibril used for the calculations. The colors indicate parallel (blue) and perpendicular (red) orientations of the β -sheets, which are perpendicular to the β -strands.

β -sheets that are parallel to the long axis of the fibril. The fibril-forming segment GNNQQNY has been studied using X-ray diffraction analysis, and the structure and orientation of the fibrous crystals in the fibrils is known (PDB code 1YJP).^{55,56} A model containing 15 β -strands in parallel sheets was created on the basis of the experimental X-ray diffraction results. The orientation of the parallel segment (Figure 9, blue) was thus defined to be consistent with its orientation in the X-ray data for fibrils: the β -sheets are oriented parallel and the single strands perpendicular to the orientation axis (in contrast to the long axis of the fragment in the PDB structure, which coincides with the axis of the strand). The sheet extends parallel to the shown strands of each of the fragments, so the short axis of the fragment is the long axis of the sheet. The experimental

spectrum shows an intense band at 198 nm that is due to the $\pi \rightarrow \pi^*$ transition polarized parallel to the sheet axis (Figure 2). The calculated band of the parallel sheet is ~ 10 nm blue-shifted (Figure 10). However, the $\pi \rightarrow \pi^*$ band of the perpendicular sheet orientation (Figure 10) is negative and disagrees with experiment.

Every strand in the amyloid fiber has the same structure, and the seven residues in each strand correspond to six peptide bonds. Calculations on individual strands or pairs of strands from the amyloid structure gave similar absorption and LD spectra. Thus, from our calculations, the electronic structure of β -sheets appears to be dominated by nearest-neighbor interactions along individual strands. The largest coupling between electronic excitations is the coupling between $\pi \rightarrow \pi^*$ transitions located on neighboring peptide groups within a single strand (Table 2). This interaction is on the order of $850 \pm 100 \text{ cm}^{-1}$ and gives rise to a splitting of ~ 5 nm between the parallel and perpendicular components in the calculated absorption spectrum. If the splitting were significantly larger, it would lead to a couplet in the LD spectrum. Calculations using a semiempirical parameter set³⁷ where the orientation of the $\pi \rightarrow \pi^*$ electric transition dipole moment is different led to nearest-neighbor $\pi\pi^*-\pi\pi^*$ interactions almost twice as large as in our calculations, resulting in a correspondingly larger exciton splitting. A couplet is not observed in the experimental LD spectrum, indicating that our calculated interaction is at least qualitatively correct. Other interactions between excitations within the strand are noticeably smaller (Table 2). However, the coupling between the $\pi \rightarrow \pi^*$ transition on one peptide group and the $n \rightarrow \pi^*$ transition on the following peptide group is significant. Interactions less than 50 cm^{-1} have little quantitative effect on the computed spectra. Equivalent atoms in the two strands forming the parallel sheet are separated by 4.86 \AA , which is larger than the distance between neighboring peptide groups within a strand. Thus, the interactions between excitations on different strands (Table 3) are weaker and also less regular.

Collagen. Collagen (PDB code 1CLG)⁵⁷ is rich in glycine, proline, and hydroxyproline, usually in the sequence GPO. The GPO content leads to the formation of extended left-handed

(55) Nelson, R.; Sawaya, M. R.; Balbirnie, M.; Madsen, A. Ø.; Riekkel, C.; Grothe, R.; Eisenberg, D. *Nature* **2005**, *435*, 773–778.

(56) Sawaya, M. R.; Sambashivan, S.; Nelson, R.; Ivanova, M. I.; Sievers, S. A.; Apostol, M. I.; Thompson, M. J.; Balbirnie, M.; Wiltzius, J. J. W.; McFarlane, H. T.; Madsen, A. Ø.; Riekkel, C.; Eisenberg, D. *Nature* **2007**, *447*, 453–457.

(57) Chen, J. M.; Kung, C. E.; Fearheller, S. H.; Brown, E. M. *J. Protein Chem.* **1991**, *10*, 535–552.

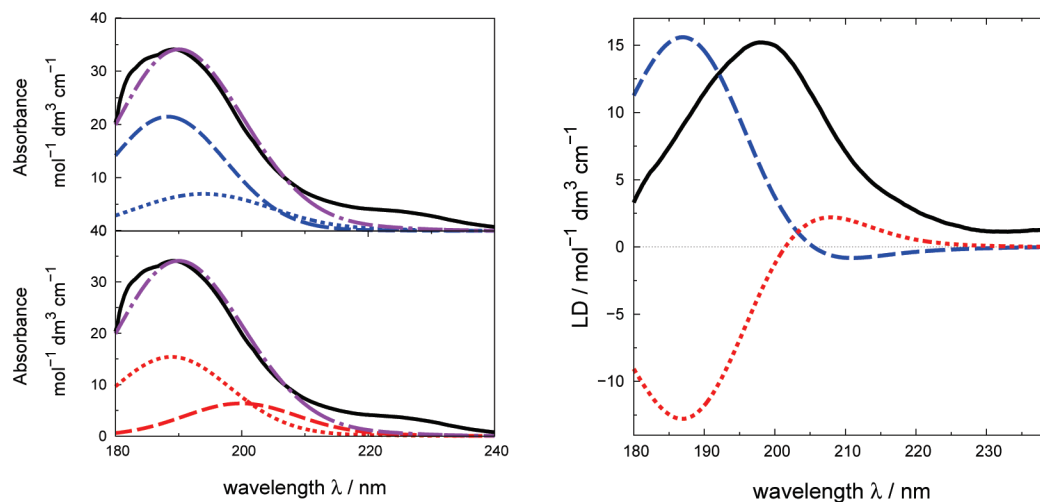


Figure 10. (left) Experimental absorbance spectrum of an amyloid fibril (solid line) and calculated spectra for parallel (top, blue) and perpendicular orientations (bottom, red): parallel absorption (dashed), perpendicular absorption (dotted), and full absorption (dot-dashed). (right) Experimental LD spectrum (solid line) of the amyloid fibril aligned in a Couette flow cell and calculated spectra for parallel (dashed) and perpendicular (dotted) orientations. The experimental spectrum has been scaled to the calculated intensity for comparison.

Table 2. Interaction Energies between Electronic Excitations within a Single Strand

peptide i	$r_{i,i+1}$ (Å)	interaction energy (cm ⁻¹)				
		$\pi\pi_i^* - n\pi_{i+1}^*$	$\pi\pi_i^* - \pi\pi_{i+1}^*$	$n\pi_i^* - n\pi_{i+1}^*$	$n\pi_i^* - \pi\pi_i^*$	$\pi\pi_i^* - \pi\pi_{i+2}^*$
1	3.00	92	-846	26	76	11
2	3.52	336	-934	81	67	59
3	3.63	409	-762	46	-26	16
4	3.44	357	-966	73	67	84
5	3.41	613	-930	114	-15	n/a

Table 3. Interaction Energies between $\pi \rightarrow \pi^*$ Excitations Located on Adjacent Strands within a Parallel β -Sheet

peptide groups ij	interaction energy (cm ⁻¹)			
	$\pi\pi_i^* - \pi\pi_{j-1}^*$	$\pi\pi_i^* - \pi\pi_j^*$	$\pi\pi_i^* - \pi\pi_{j+1}^*$	$\pi\pi_i^* - \pi\pi_{j+2}^*$
1	n/a	340	-29	-29
2	-107	-450	159	67
3	91	-374	79	-97
4	32	-271	79	70
5	98	-406	131	n/a
6	36	-455	n/a	n/a

helices of the poly(proline)-II (P_{II}) type. The helices are stabilized by steric interactions of the proline side chain groups, while the small glycine residues facilitate a very narrow helix diameter. Three P_{II} helices supercoil to form a triple helix with a length of ~ 9 nm (Figure 11). The P_{II} structure is still a challenge for matrix-method CD calculations, and the characteristic negative maximum around 200 nm has not been reproduced.^{33,58} Mandel and Holzwarth⁵⁹ measured the LD of samples of calfskin collagen and poly-L-proline that had been oriented by flow through a set of parallel capillary tubes. They deconvolved the absorption spectrum of collagen into Gaussian bands, resulting in a parallel band at 200 nm with 40% of the $\pi \rightarrow \pi^*$ intensity and a perpendicular band at 189 nm.

The LD spectrum of collagen in Couette flow has recently been reported,¹⁰ and a positive band around 201 nm suggests from estimates of transition polarizations that collagen is probably oriented with its long axis parallel within the fiber.

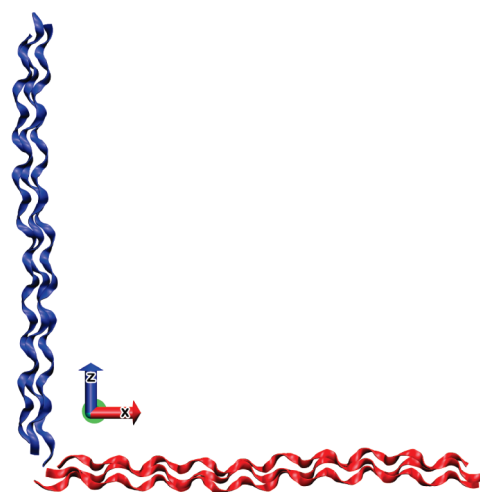


Figure 11. Orientation of collagen used for the calculations. The colors indicate parallel (blue) and perpendicular (red) orientations.

Our calculations were based on the PDB structure 1CLG, a synthetic sequence that is 33% glycine and 66% proline, similar in composition to the experimental sample. The dihedral backbone angles show a small spread around the typical P_{II} values of ($\phi = -75^\circ$, $\psi = 160^\circ$). Calculations on a single helix, in contrast to the triple-helical structure, yielded similar LD spectra, indicating that the interhelical interactions are limited. The calculated LD spectrum (Figure 12) for the parallel orientation of collagen agrees well with experiment. Since the rod-shaped structure is similar to the structures of SAF and tropomyosin, the orientation of collagen can be assumed to be parallel in the fiber, and the calculations support this. We calculate that the parallel-polarized component of the $\pi \rightarrow \pi^*$ transition lies at higher wavelength (201 nm) than the perpendicular-polarized component (190 nm) but is slightly less intense. The agreement between the experimental spectrum and our calculated absorption spectrum is good. These components are 5 nm blue-shifted relative to previous suggestions¹⁰ of 195 and 206 nm but give a similar splitting. Despite this blue shift, the calculated LD band at 206 nm is actually red-shifted by 5 nm compared with the experimental spectrum, suggesting that

(58) Sreerama, N.; Woody, R. W. *Protein Sci.* **2003**, *12*, 384–388.

(59) Mandel, R.; Holzwarth, G. *Biopolymers* **1973**, *12*, 655–674.

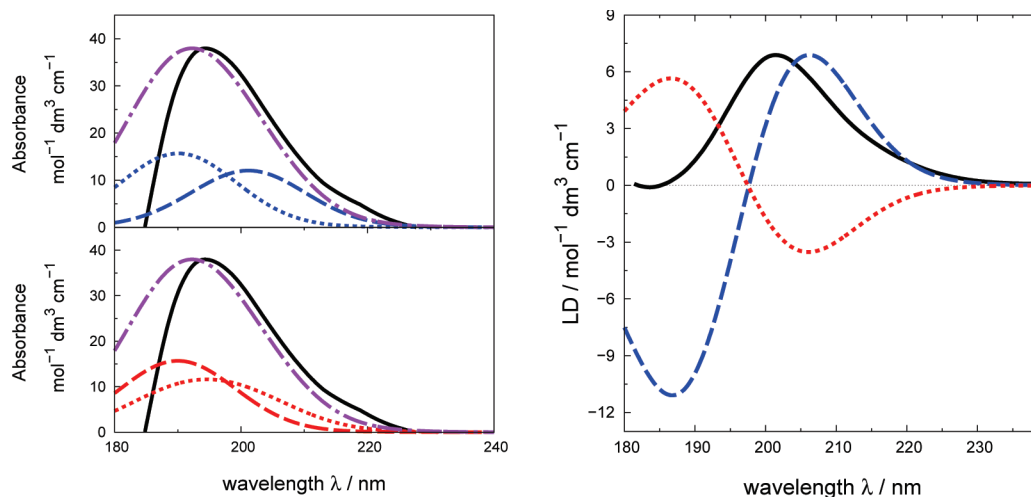


Figure 12. (left) Experimental absorbance spectrum of collagen (solid line) and calculated spectra for parallel (top, blue) and perpendicular orientations (bottom, red): parallel absorption (dashed), perpendicular absorption (dotted), and full absorption (dot-dashed). (right) Experimental LD spectrum (solid line) of collagen aligned in a Couette flow cell and calculated spectra for parallel (dashed) and perpendicular (dotted) orientations. The experimental spectrum has been scaled to the calculated intensity for comparison.

the perpendicular-polarized component of the $\pi \rightarrow \pi^*$ transition is actually more intense than we calculated.

The negative band located at 187 nm in the calculated LD spectrum has no experimental counterpart. Inclusion of charge-transfer transitions in the deep-UV³⁴ did not qualitatively change the calculated LD spectrum. Woody^{60,61} incorporated the coupling of higher-energy transitions into the calculation of the CD spectrum of the P_{II} conformation using a polarizability term and suggested that this may account for the nonconservative nature of the CD in the 180–250 nm region. The additional negative band in our calculated LD spectrum may be another manifestation of this effect. LD depends only on the electric transition dipole moments, while CD is equally dependent on the magnetic transition dipole moments. The promising results of the LD calculations in reproducing the experimental features of a P_{II} structure suggest that the challenge for CD is related to the magnetic transition dipole moments.

Discussion and Conclusion

The aim of this work was to determine whether it is possible to calculate the LD spectra of fibrous proteins and gain information on the orientation of the component units. In contrast to a CD spectrum, an LD spectrum depends only on the electric transition dipole moments. Therefore, there is no sensitivity to the magnetic transition dipole moments, which can be challenging to calculate accurately. This aspect manifests itself in the successful calculation of the LD spectrum of collagen, a structure for which the calculation of the CD spectrum remains problematic. The matrix-method LD calculations produced the correct result for five proteins with known orientations in the fiber. For each of the fibrous proteins studied in this work, there was a clear indication from the literature about how to define the (parallel) orientation axis, from which the definition of perpendicular orientation follows. This was either given by a high aspect ratio or, in the cases of FtsZ and the amyloid-like fibril, experimental data. However, more generally such data may not be available. It is therefore desirable to be able to perform calculations for any given protein

orientation between the perfectly parallel and perpendicular alignments. Once the matrix method methodology for LD had been established, this was the main challenge faced in this work. The LD of an ideal helix is invariant with respect to a rotation about the helix long axis, which can be assumed to be the orientation axis. More complicated folding motifs require consideration of the rotation about the molecular axis for each orientation. This analysis is crucial if LD calculations are to be used for elucidation of unknown orientations instead of confirmation of a previously determined orientation. The functionality is available on DichroCalc²⁷ and allows one to calculate the LD spectrum for every possible orientation of the monomer unit within the fiber. An example report of such an analysis is provided in the Supporting Information.

The *orientation search* defines the z axis as the parallel direction or orientation axis. The LD is invariant with respect to rotation about this axis because of the rotational average over the perpendicular axes (eq 2). As the initial position for an orientation search, the long (internal) axis of the protein is aligned with z . In order to cover all possible orientations, the protein is tilted about one of the perpendicular axes at angles between 0 and 90° with respect to the orientation axis. At each tilt angle, the molecule is rotated about its internal long molecular axis. For symmetry reasons, orientations with angles greater than 90° with respect to the orientation axis do not have to be considered. An LD spectrum is calculated for every orientation, and one can then identify the orientation that gives a calculated LD spectrum in closest agreement with experiment. For each orientation, only a single monomer is considered, i.e. different orientations of monomers within the fiber are not yet taken into account. However, such considerations may be implemented in the future.

For each of the proteins, in the calculated LD spectrum the couplet around 200 nm is reproduced with a good match of the band position. For fibers with high helical content, the theoretical LD spectra agree very well with the expected orientations of SAF, tropomyosin, and collagen, whose long axes are oriented parallel to the flow direction. For FtsZ, which contains α -helices and β -sheets, the results presented in Figure 8 support the orientation determined from X-ray diffraction analyses of a dimer. For the β -sheet example, the amyloid-like fibril, the

(60) Woody, R. W. *Monatsh. Chem.* **2005**, *136*, 347–366.

(61) Woody, R. W. *J. Am. Chem. Soc.* **2009**, *131*, 8234–8245.

calculations agree with the experimental evidence that the sheets are oriented parallel to the orientation axis. Deviations from a perfectly parallel orientation would be apparent in the intensity of the bands, which decreases until the magic angle of 54.7° is reached, at which point no LD signal would be observed. Since the experimental and calculated intensities cannot be quantitatively compared (as the orientation parameter is unknown in the experiments), it can only be concluded that the inclination from the orientation axis is smaller than 54.7° , but this does rule out a perpendicular alignment in the flow direction.

In conclusion, matrix-method calculations provide insight into protein orientation in shear flow for all major secondary structure classes. Online software for performing these calculations has been made available to the scientific community. Improvements to the experimental technique or determination of a defined scaling factor that takes into account the orientation factor in

an experiment could allow a quantitative comparison of the experimental and calculated intensity, which would facilitate a yet more accurate estimate of the orientation of the protein in solution.

Acknowledgment. We thank the Engineering and Physical Sciences Research Council (Grant GR/T09224) for funding and the University of Nottingham for access to High Performance Computing facilities. Xi Cheng and R. Pacheco Gomez (University of Warwick) are thanked for the FtsZ measurements. The collagen was supplied by R. Visse (Imperial College London).

Supporting Information Available: Orientation search report for FtsZ and complete ref 26. This material is available free of charge via the Internet at <http://pubs.acs.org>.

JA902662E