Kilmartin, J. V., & Rossi-Bernardi, L. (1971) *Biochem. J. 124*, 31–45.
Kilmartin, J. V., Fogg, J., Luzzana, M., & Rossi-Bernardi, L. (1973) *J. Biol. Chem. 248*, 7039–7043.
Maeda, T., Imai, K., & Tyuma, I. (1972) *Biochemistry 11*, 3685–3689.
Nagai, K., & Kitagawa, T. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 2033–2037.
Ogawa, S., & Shulman, R. G. (1972) *J. Mol. Biol. 70*, 315–336.
Ondrias, M. R., Rousseau, D. L., Kitagawa, T., Ikeda-Saito, M., Inubushi, T., & Yonetani, T. (1982) *J. Biol. Chem. 257*, 8766–8770.
Scott, T. W., Friedman, J. M., Ikeda-Saito, M., & Yonetani, T. (1983) *FEBS Lett. 158*, 68–72.
Shaanan, B. (1983) *J. Mol. Biol. 171*, 31–59.
Shibayama, N., Morimoto, H., & Kitagawa, T. (1986) *J. Mol. Biol. 192*, 331–336.
Shibayama, N., Inubushi, T., Morimoto, H., & Yonetani, T. (1987) *Biochemistry 26*, 2194–2201.
Simolo, K., Stucky, G., Chen, S., Bacley, M., Scholes, C., & Mclendon, G. (1985) *J. Am. Chem. Soc. 107*, 2865–2872.
Takahashi, S., Lin, A. K.-A. C., & Ho, C. (1980) *Biochemistry 19*, 5196–5202.

Theoretical Study of the Contribution of Aromatic Side Chains to the Circular Dichroism of Basic Bovine Pancreatic Trypsin Inhibitor†

Mark C. Manning‡ and Robert W. Woody*
Department of Biochemistry, Colorado State University, Fort Collins, Colorado 80523
Received April 7, 1989

ABSTRACT: Circular dichroism (CD) spectroscopy is frequently employed to determine the secondary structure composition of a protein. However, this assumes that the far-UV region of the spectrum, which is used for these analyses, is due only to contributions from the polypeptide backbone. Basic bovine pancreatic trypsin inhibitor (BPTI) possesses an unusual far-UV CD spectrum, which has made such an analysis difficult. One possible reason for the discrepancy is that other chromophores, such as the aromatic side chains (four tyrosines, four phenylalanines), might be responsible. The CD spectrum of BPTI was calculated by employing a variation of the matrix method. Including only the peptide backbone gave poor agreement between theory and experiment. This was shown to be independent of the quality of the calculation performed. Subsequent inclusion of tyrosine contributions did little to improve the fit. However, further inclusion of the phenylalanine chromophores provided a good fit between the calculated and experimental far-UV spectrum. The important contributions arise from the cluster of aromatic amino acids formed by two tyrosines (Tyr^{21} and Tyr^{23}) and three phenylalanines (Phe^{22}, Phe^{4}, and Phe^{45}). Consideration of both types of side chains and the entire peptide backbone is essential to produce an accurate description of the CD curve. Overall, these results indicate that contributions from aromatic amino acids can significantly perturb the far-UV CD spectrum of a protein, making secondary structure analysis difficult. This is particularly true in systems like BPTI, with low amounts of α-helical structure and clusters of aromatic amino acids.

Recent developments have allowed a greater understanding of the mechanisms of protein folding and the determinants of protein conformation in solution. Basic bovine pancreatic trypsin inhibitor (BPTI)[1] has been a primary system for study in these areas. Its conformation has been extensively examined by nuclear magnetic resonance (NMR) (Snyder et al., 1975; Wagner, 1983; DeMarco et al., 1985; Wüthrich, 1986; Wagner et al., 1987a,b), circular dichroism (Kosen et al., 1981, 1983), and absorption (Kosen et al., 1980) spectroscopy. The role of disulfide bond formation in controlling the folding to the native conformation has been well documented (Creighton, 1984; Marks et al., 1987; States et al., 1987). In addition, BPTI has been the subject of molecular mechanics (Dunfield & Scheraga, 1980; Chou et al., 1985; Billeter et al., 1987) and molecular dynamics (Gelin & Karplus, 1975; McCammon et al., 1979; van Gunsteren et al., 1983) studies. All evidence indicates that BPTI adopts a relatively rigid, compact globular structure in solution similar to the conformation deduced by X-ray crystallography (Deisenhofer & Steigemann, 1975; Bode et al., 1984; Wlodawer et al., 1984, 1987a,b). Despite the wealth of data available on this protein, no detailed theoretical analysis of its circular dichroism (CD) has been published. Secondary structure estimates based upon the far-UV CD spectrum of BPTI have yielded poor results. The large fraction of aromatic residues (4 tyrosines, 4 phenylalanines, and no tryptophans out of 58 residues) suggests that aromatic side chains may make important contributions to both the near- and far-UV CD spectrum of BPTI.

Contributions from aromatic side chains to the CD of proteins and polypeptides have been recognized in both experimental and theoretical studies (Hooker & Schellman, 1970; Chen & Woody, 1971; Strickland, 1974; Strickland & Mercola, 1976; Strassburger et al., 1982). Such effects are

† Financial support for this work was provided by USPHS Grant GM-22994.
* Author to whom correspondence should be addressed.
‡ Present address: Department of Pharmaceutical Chemistry, The University of Kansas, Lawrence, KS 66045.

[1] Abbreviations: BPTI, basic bovine pancreatic trypsin inhibitor; CD, circular dichroism; NMR, nuclear magnetic resonance; UV, ultraviolet.

0006-2960/89/0428-8609$01.50/0

most prominent in systems where aromatic groups are in close proximity. Clusters of aromatic amino acids have been shown to be a common motif in globular proteins (Burley & Petsko, 1985). One such cluster occurs in BPTI (Tyr-21, Phe-22, Tyr-23, Phe-4, Phe-45). These factors, coupled with the detailed structural information available from X-ray crystallography and NMR studies, make BPTI an ideal case for evaluating side-chain contributions to the far-UV CD of proteins.

Materials and Methods

All CD spectra were calculated by employing the matrix method (Bayley et al., 1968) in an origin-independent formulation (Goux & Hooker, 1980), modified to include four transitions on each peptide group (M. C. Manning and R. W. Woody, unpublished results). These included the $n\pi^*$, $\pi_0\pi^*$ (NV_1), $\pi_+\pi^*$ (NV_2), and $n'\pi^*$ excited states. Four transitions were included for each phenylalanine and tyrosine side chain. For phenylalanine, the excited states were L_b (260 nm), L_a (208 nm), B_b (181 nm), and B_a (180 nm). For tyrosine, the excited states were L_a (276 nm), L_b (227 nm), B_b (192 nm), and B_a (191 nm). All wave functions are based upon CNDO/S calculations (Del Bene & Jaffe, 1968; Ellis et al., 1972; Jacques et al., 1981) on model compounds (*N*-methylacetamide for the peptide group, toluene for the phenylalanine side chain, and *p*-methylphenol for the tyrosine side chain). Mulliken ground-state charge distributions (Mulliken, 1955) were employed. Bandwidths were taken to be 12.0 nm for the L_b transitions, 10.5 nm for the L_a transitions of tyrosine, 9.5 nm for the L_a band of phenylalanine, 10.5 nm for the peptide $n\pi^*$ transitions, 10.5 nm for the NV_1 ($\pi_0\pi^*$) transition, 10.4 nm for the B_b and B_a transitions of tyrosine, 9.7 nm for the B_b and B_a transitions of phenylalanine and the $n'\pi^*$ peptide transition, and 6.6 nm for the NV_2 ($\pi_+\pi$) transition. Bandwidths were chosen on the basis of the observed values for model mono- and dipeptides (Nielsen & Schellman, 1967; Strickland, 1974).

The coordinates for the BPTI structure were taken from the Protein Data Bank (form I, 4PTI; form II, 5PTI).

Results and Discussion

Both α- and β-structures are present in BPTI. In particular, there are two short α helices, one near each end of the polypeptide chain. These pack around a long double-stranded β sheet, which runs the length of the molecule. No unusual secondary structural features are present, yet BPTI displays a distinctive far-UV CD with a negative shoulder at 220 nm and a sharp negative band near 200 nm. Its crossover to positive ellipticity is at 195 nm, which is a significantly shorter wavelength than would be expected for such a secondary structure composition. These features have made analysis of the CD difficult. One possible explanation might be that the far-UV CD is not due solely to the polypeptide backbone and its conformation. Other UV-active chromophores are present, such as the aromatic groups of tyrosine and phenylalanine. Their significance is suggested not only by the unusual CD of BPTI but also by the high content of these residues. They are present in nearly twice the numbers expected for a protein this size (Lehninger, 1975). In addition, the interaction of aromatic groups increases when they are in close proximity (Chen & Woody, 1971; Strickland & Mercola, 1976). Such could be the case in BPTI, where a cluster of these groups is formed by Tyr^{21}, Tyr^{23}, Phe^{45}, Phe^{4}, and Phe^{22}. In order to evaluate the impact of aromatic side-chain chromophores on the far-UV CD spectrum of BPTI, detailed calculations must be performed to identify the relative contributions of each part

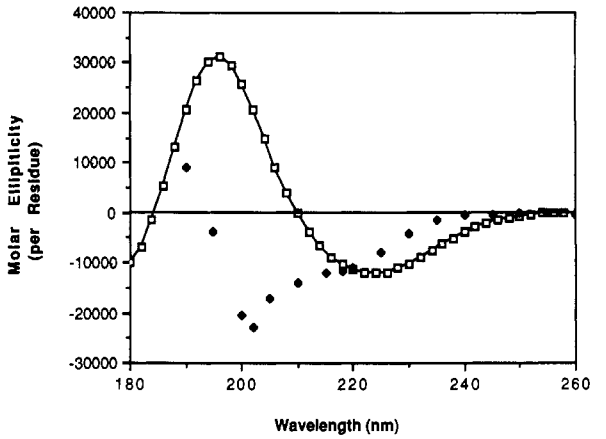

FIGURE 1: Calculated far-UV CD spectrum (□) for the polypeptide chain of BPTI (form II) including only two transitions ($n\pi^*$ and NV_1) per peptide residue. The experimental spectrum (◆) is shown for comparison.

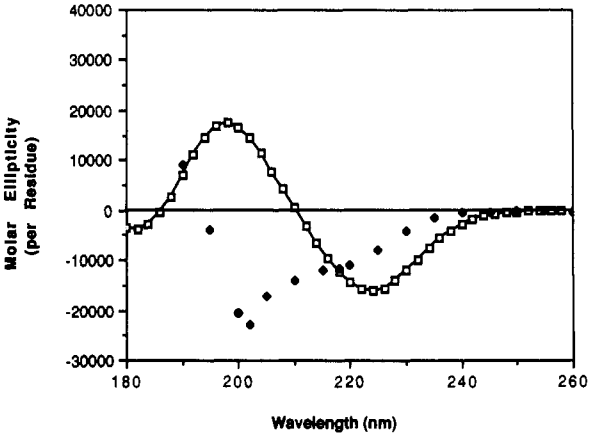

FIGURE 2: Calculated far-UV CD spectrum (□) for the polypeptide chain of BPTI (form II) including four transitions ($n\pi^*$, $n'\pi^*$, NV_1, and NV_2) per peptide residue. All subsequent calculations include all of these transitions. The experimental spectrum (◆) is shown for comparison.

of the molecule, beginning with the polypeptide backbone.

Initially, the CD of BPTI was calculated with consideration only to the polypeptide backbone, in order to ascertain the contributions due to secondary structure. These studies were carried out on the conformation determined for form II (Wlodawer et al., 1984). Including only two transitions per peptide residue ($n\pi^*$ and $\pi_0\pi^*$) yields a far-UV CD curve indicative of a protein containing both α- and β-structures, such as BPTI (Figure 1). This CD spectrum is in poor agreement with that experimentally observed. One possible reason for the discrepancy is an incomplete description of the electronic structure of the peptide backbone. Therefore, two additional transitions ($n'\pi^*$ and $\pi_+\pi^*$) were included for each residue (Figure 2). While changes in intensity are observed, there is little change in band positions between the two theoretical spectra. A second shortcoming which might give rise to the difference in the observed and predicted spectra is that, up to this point, all peptides were assumed to be secondary amides. It has been shown that proline residues, being tertiary amides, have different excited-state energies and ground-state charge distributions (Nielsen & Schellman, 1967). Explicit inclusion of the proline residues leads to further change in the intensity of the calculated far-UV CD curve (Figure 3), yet the general shape and band positions remain incorrect. Therefore, the far-UV CD of BPTI must contain contributions in addition to those of the polypeptide chain.

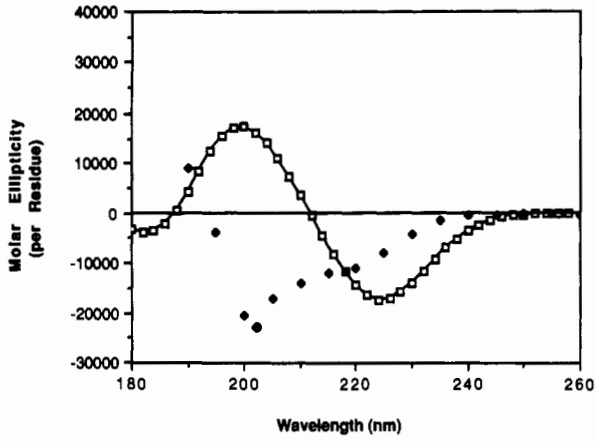

FIGURE 3: Calculated far-UV CD spectrum (□) of the polypeptide chain of BPTI (form II) explicitly including the contributions of the proline residues. All subsequent calculations include this modification. The experimental spectrum (◆) of BPTI is shown for comparison.

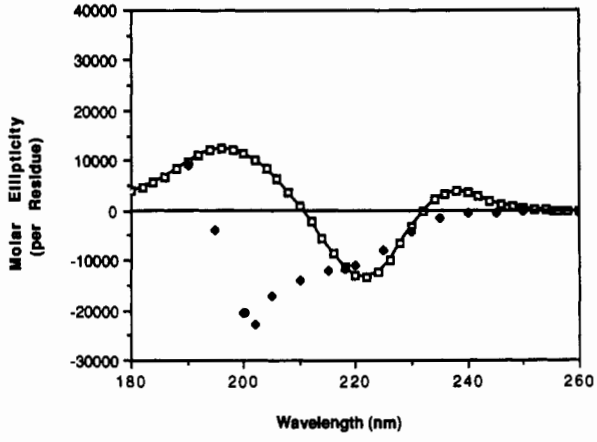

FIGURE 4: Calculated far-UV CD spectrum (□) of BPTI (form II) including the transitions of the polypeptide chain and the tyrosine side chains. The experimental spectrum (◆) of BPTI is shown for comparison.

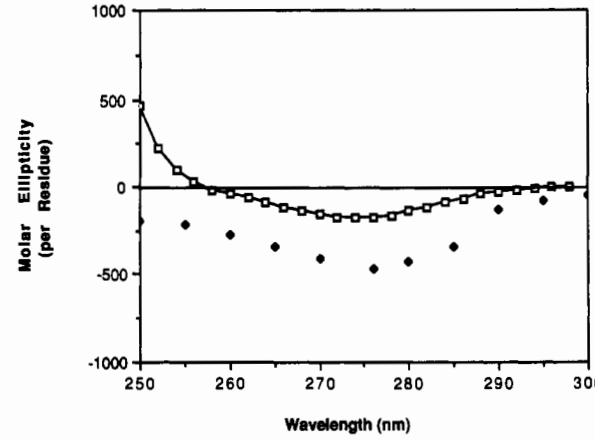

FIGURE 5: Calculated near-UV CD spectrum (□) of BPTI (form II) including the transitions of the polypeptide chain and the tyrosine side chains. The experimental spectrum (◆) of BPTI is shown for comparison.

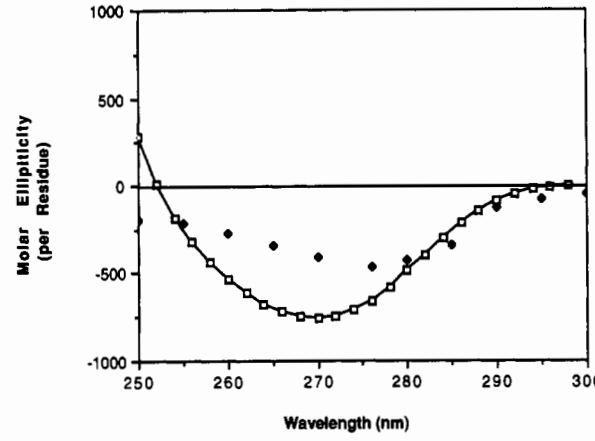

FIGURE 6: Calculated near-UV CD spectrum (□) of BPTI (form II) including the transitions of the polypeptide chain and the tyrosine and phenylalanine side chains. The experimental spectrum (◆) of BPTI is shown for comparison.

Many factors besides conformation of the peptide backbone can affect the far-UV CD spectrum of a protein. Unfortunately, the importance of these factors is still poorly understood. Both tertiary structure and secondary structure deviations can play a significant role in modulating CD spectra (Manning et al., 1988). These factors are implicitly included in these calculations, which are based upon a single well-defined structure. Another possibility is contributions from other UV-active chromophores. Aromatic side chains are the most prominent of these moieties. Neither histidine nor tryptophan is present in BPTI, so only phenylalanine (Phe) and tyrosine (Tyr) rings need be considered. It is generally assumed that the rotational strengths from Tyr-localized transitions will be much larger than for Phe, particularly in the near-UV. Therefore, only Tyr excited states were first included in the calculations. All the previous peptide interactions and modifications (four transitions per peptide group; explicit inclusion of prolines) were retained in these computations. While the shape of the theoretical curve is altered upon inclusion of the Tyr side-chain transitions, the theoretical CD curve still does not agree with the experimental data (Figure 4). Inclusion of tyrosine also produces CD bands in the near-UV. Experimentally, a weak negative band near 275 nm is observed and is attributed primarily to the L_b transition of tyrosine. A weak negative feature is predicted (Figure 5), but with approximately one-third the actual intensity. Nearly all the intensity of this feature is due to the L_b transition of Tyr^{21}. Strong electronic coupling is observed between Tyr^{21} and Tyr^{23}. The polypeptide chain adopts a β-sheet conformation in this region, placing these two side chains on the same side and in close proximity. Rotation of either tyrosine ring by as little as 10° leads to dramatic changes in the sign and intensity of the 275-nm band, indicating strong electronic coupling between these two chromophores (data not shown). No noticeable differences are discerned in the far-UV region upon rotation of any of the tyrosine side chains. As the introduction of only the tyrosine groups did not reconcile the predicted and observed far-UV CD spectra, the phenylalanine side chains were also considered.

Phenylalanine rings have the same four transitions as tyrosine rings, although they are typically less intense and occur at slightly higher energies. For example, the L_b transition of phenylalanine is near 260 nm (vs ~275 nm for tyrosine). Considering the small distance between many of the phenylalanine and tyrosine rings, especially in this aromatic cluster, and the near degeneracy of their electronic transitions, strong interactions may be anticipated. Inclusion of the phenylalanine side chains leads to marked changes in the predicted CD spectra, both in the near- (Figure 6) and in the far-UV (Figure 7). In the near-UV, the sign of the tyrosine L_b band is now predicted to be more intense than observed experimentally. Again, almost all of the intensity arises from the transition on Tyr^{21}. Apparently, interaction with nearby phenylalanine rings is significant and presumably essential for accurate

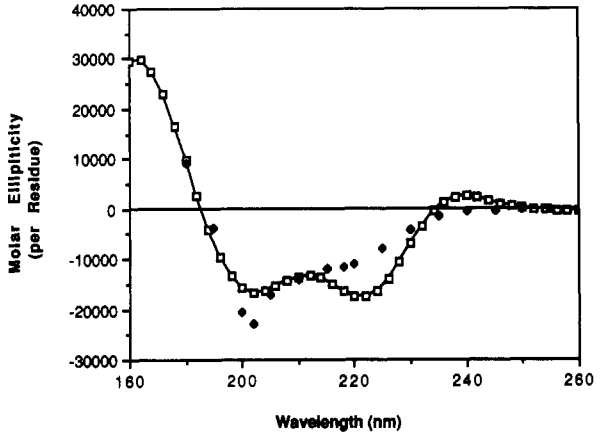

FIGURE 7: Calculated far-UV CD spectrum (□) of BPTI (form II) including the transitions of the polypeptide chain and the tyrosine and phenylalanine side chains. The experimental spectrum (◆) of BPTI is shown for comparison.

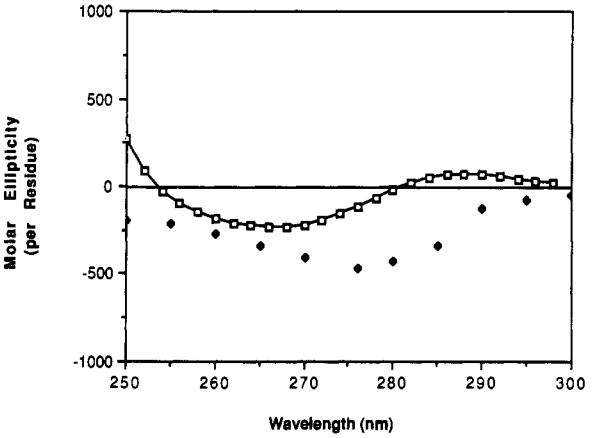

FIGURE 8: Calculated far-UV CD spectrum of form I (□) of BPTI including the transitions of the polypeptide chain and the tyrosine and phenylalanine side chains. The experimental spectrum (◆) of BPTI is shown for comparison.

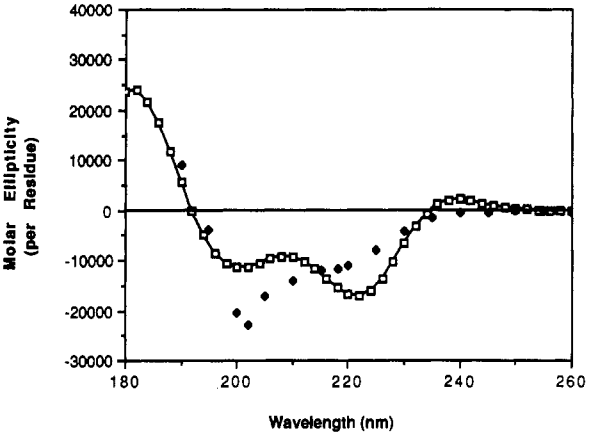

FIGURE 9: Calculated near-UV CD spectrum of form I (□) of BPTI including the transitions of the polypeptide chain and the tyrosine and phenylalanine side chains. The experimental spectrum (◆) of BPTI is shown for comparison.

prediction of rotational strengths. Further analysis indicates that both coupled oscillator and one-electron terms contribute to the rotational strength of the 275-nm band. Thus, the static charge distribution surrounding the tyrosine rings is as important as the electronic coupling between the various chromophores.

Such interaction is necessary to correctly calculate the far-UV CD as well. Consideration of the phenylalanine side chains produces dramatic changes in the computed CD curve. As in the near-UV, it is the attainment of a correct description of the interactions between the different chromophores that is important, rather than the contributions of the phenylalanine moieties per se. Overall, the agreement with experiment is good. The strong negative band near 200 nm is reproduced, as is the shoulder at 220 nm. There is a significant difference between the two CD spectra near 230 nm. A positive feature arising from the L_a band of tyrosine is predicted but not observed experimentally. On the basis of curve fitting of the experimental spectrum, disulfide-based transitions were postulated to produce a strong negative CD signal in this region (Kosen et al., 1981). As disulfides were not included in these calculations, this may be the reason for the discrepancy.

As mentioned previously, BPTI exists in a number of isomeric forms in the solid state. To date, three such isomorphs have been isolated and characterized by X-ray crystallography. Coordinates for the first two (forms I and II) have been deposited in the Protein Data Bank. Dihedral angles may vary between forms by as much as 10–15°, both for backbone and side-chain groups. The calculated spectra to this point have been based upon form II. Predicted spectra for form I in the far- (Figure 8) and near-UV (Figure 9) display larger discrepancies with experiment, suggesting that the solution conformation is nearer to that of form II than form I.

In conclusion, the matrix method for calculating the CD spectra of proteins has been extended to include up to four transitions on each peptide unit and aromatic side chain. This method has been applied to a protein, BPTI, which has been well characterized in terms of its solution and solid-state structure. Consideration of only the polypeptide chain leads to a predicted CD spectrum that is consistent with the secondary structural composition of BPTI. However, this spectrum is markedly different from the one obtained experimentally. Inclusion of the tyrosine side chains leads to little improvement in the far-UV CD. In order to correctly compute the CD spectra of BPTI, both tyrosine and phenylalanine side-chain chromophores must be included along with the polypeptide backbone. This leads to good agreement with experiment in both the near- and far-UV. A detailed analysis of the near-UV CD spectrum produced two important results. First, nearly all of the rotational strength for the 275-nm band arises from the L_b excited state on Tyr^{21}, which is at the center of a cluster of aromatic side chains. There is evidence of strong electronic coupling between each of these rings. Second, both one-electron and coupled oscillator terms contribute significantly to the L_b transition rotational strength. Therefore, theoretical methods considering only coupled oscillator interactions are probably not adequate to accurately predict the near-UV spectrum of proteins. Overall, these studies provide evidence that the aromatic side chains make significant contributions to the far-UV CD spectrum of BPTI. Such contributions account for the unusual CD displayed by BPTI. This work further suggests that aromatic side chains, especially when in close proximity and in proteins with low to moderate α-helix content, may perturb far-UV CD spectra enough to make secondary structure estimations unreliable.

ACKNOWLEDGMENTS

We thank the Colorado State University Computer Center for a grant of computer time.

REFERENCES

Bayley, P. M., Nielsen, E. B., & Schellman, J. A. (1968) *J. Phys. Chem. 73*, 228–243.

Billeter, M., Havel, T. F., & Kuntz, I. D. (1987) *Biopolymers 26*, 777–793.
Bode, W., Walter, J., Huber, R., Wenzel, H. R., & Tschesche, H. (1984) *Eur. J. Biochem. 144*, 185–190.
Burley, S. K., & Petsko, G. A. (1985) *Science 229*, 23–28.
Chen, A. K., & Woody, R. W. (1971) *J. Am. Chem. Soc. 93*, 29–37.
Chou, K.-C., Nemethy, G., Pottle, M. S., & Scheraga, H. A. (1985) *Biochemistry 24*, 7948–7953.
Creighton, T. E. (1984) *Proteins*, W. H. Freeman, New York.
Deisenhofer, J., & Steigemann, W. (1975) *Acta Crystallogr. 31B*, 238–250.
Del Bene, J., & Jaffe, H. H. (1968) *J. Chem. Phys. 48*, 1807–1813.
De Marco, A., Zetta, L., & Kaptein, R. (1985) *Eur. Biophys. J. 11*, 187–193.
Dunfield, L. G., & Scheraga, H. A. (1980) *Macromolecules 13*, 1415–1428.
Ellis, R. L., Kuehnlenz, G., & Jaffe, H. H. (1972) *Theor. Chim. Acta 26*, 131–140.
Gelin, B. R., & Karplus, M. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 2002–2006.
Goux, W. J., & Hooker, T. M., Jr. (1980) *J. Am. Chem. Soc. 102*, 7080–7087.
Hooker, T. M., Jr., & Schellman, J. A. (1970) *Biopolymers 9*, 1319–1348.
Jacques, P., Faure, J., Chalvet, O., & Jaffe, H. H. (1981) *J. Phys. Chem. 85*, 473–479.
Kosen, P. A., Creighton, T. E., & Blout, E. R. (1980) *Biochemistry 19*, 4936–4940.
Kosen, P. A., Creigton, T. E., & Blout, E. R. (1981) *Biochemistry 20*, 5744–5754.
Kosen, P. A., Creigton, T. E., & Blout, E. R. (1983) *Biochemistry 22*, 2433–2440.
Lehninger, A. L. (1975) *Biochemistry. The Molecular Basis of Cell Structure and Function*, 2nd ed., Worth Publishers, New York.
Manning, M. C., Illangasekare, M., & Woody, R. W. (1988) *Biophys. Chem. 31*, 77–86.
Marks, C. B., Naderi, H., Kosen, P. A., Kuntz, I. D., & Anderson, S. (1987) *Science 235*, 1370–1373.
McCammon, J. A., Wolynes, P. G., & Karplus, M. (1979) *Biochemistry 18*, 927–942.
Mulliken, R. S. (1955) *J. Chem. Phys. 23*, 1833–1840.
Nielsen, E. B., & Schellman, J. A. (1967) *J. Phys. Chem. 71*, 2297–2304.
Snyder, G. H., Rowan, R., III, Karplus, S., & Sykes, B. D. (1975) *Biochemistry 14*, 3765–3777.
States, D. J., Creigton, T. E., Dobson, C. M., & Karplus, M. (1987) *J. Mol. Biol. 195*, 731–739.
Strassburger, W., Glatter, U., Wollmer, A., Fleischchauer, J., Mercola, D. A., Blundell, T. L., Glover, I., Pitts, J. E. Tickle, I. J., & Wood, S. P. (1982) *FEBS Lett. 139*, 295–299.
Strickland, E. H. (1974) *CRC Crit. Rev. Biochem. 2*, 113–174.
Strickland, E. H., & Mercola, D. (1976) *Biochemistry 15*, 3875–3884.
van Gunsteren, W. F., Berendsen, H. J. C., Hermans, J., Hol, W. G. J., & Postma, J. P. M. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 4315–4319.
Wagner, G. (1983) *Q. Rev. Biophys. 16*, 1–57.
Wagner, G., Braun, W., Havel, T. F., Schaumann, T., Go, N., & Wüthrich, K. (1987a) *J. Mol. Biol. 196*, 611–639.
Wagner, G., Brühwiler, D., & Wüthrich, K. (1987b) *J. Mol. Biol. 196*, 227–231.
Wlodawer, A., Walter, J., Huber, R., & Sjölin, L. (1984) *J. Mol. Biol. 180*, 301–331.
Wlodawer, A., Deisenhofer, J., & Huber, R. (1987a) *J. Mol. Biol. 193*, 145–156.
Wlodawer, A., Nachman, J., Gilliland, G. L., Gallagher, W., & Woodward, C. (1987b) *J. Mol. Biol. 198*, 469–480.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.