Tyrosine, Phenylalanine, and Disulfide Contributions to the Circular Dichroism of Proteins: Circular Dichroism Spectra of Wild-Type and Mutant Bovine Pancreatic Trypsin Inhibitor†

Narasimha Sreerama,‡ Mark C. Manning,§ Michael E. Powers,§ Jian-Xin Zhang,|| David P. Goldenberg,|| and Robert W. Woody*,‡

Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, Colorado 80523, Department of Pharmaceutical Sciences, School of Pharmacy, Campus Box C238, University of Colorado Health Sciences Center, Denver, Colorado 80262, and Department of Biology, University of Utah, Salt Lake City, Utah 84112

Received March 4, 1999; Revised Manuscript Received June 9, 1999

ABSTRACT: Improved descriptions of the lowest energy excited states of tyrosine and phenylalanine side chains have been developed in order to extend the capabilities of calculating the circular dichroism (CD) spectra of proteins. Four transitions (L_b, L_a, B_b, and B_a) for each of the side-chain chromophores were considered, and the transition monopole charges were obtained from a CNDO/S calculation on models representing the individual groups. Monopole charges at midpoints of the bonds, corresponding to the maximum transition charge densities in the L_b band, and monopole charges representing the vibronic coupling with the B transitions for the L_a transition were also included. The aromatic transitions were combined with the peptide transitions ($n\pi^*$, $\pi_0\pi^*$ $n'\pi^*$, and $\pi_+\pi^*$) and disulfide transitions ($n_1\sigma^*$ and $n_4\sigma^*$) in the framework of the origin-independent matrix method to compute the CD spectra of different crystal forms and Y → L and F → L mutants of bovine pancreatic trypsin inhibitor (BPTI). The structures of the mutants were obtained by replacing the appropriate tyrosine or phenylalanine residue by leucine in the wild-type crystal structure. The CD calculations were performed on the energy-minimized structures. The CD spectrum calculated for the form II crystal structure of BPTI showed the best agreement with experiment. In the far UV, the calculated and experimental CD spectra agree to various extents for the wild-type and mutant BPTI. Among the mutants, the calculated CD spectra of Y4L, Y10L, Y23L, and F45L showed reasonable agreement with experiment, while those of Y21L and F22L, the two residues interacting with most aromatic groups, showed poor agreement. In the near UV, the negative bands predicted for the wild-type and mutant BPTI have much less intensity than observed experimentally.

Circular dichroism (CD)[1] spectroscopy is a valuable tool in the investigation of protein and nucleic acid structure due to its sensitivity to their secondary structures. The utility of CD in the secondary structure estimation of proteins and in monitoring folding and unfolding transitions in proteins, as well as secondary structure of nucleic acids, is well documented (*1*). The existing methods for analyzing protein CD spectra assume that the polypeptide backbone dominates the far-UV CD spectra, which is by and large valid (*2*). However, chromophores other than the amides that make up the polypeptide backbone can make significant contributions, and of these the side chains of the aromatic residues phenylalanine, tyrosine, and tryptophan are particularly important (*3*). Analyses of protein CD spectra dominated by, or having substantial contributions from, the aromatic residues could be in error.

The ability to accurately calculate the CD spectra from first principles would be of great utility in the structural characterization of biomolecules. The matrix method (*4*), which treats a biomolecule as a collection of chromophores, enables one to calculate the CD spectra of proteins and nucleic acids. In these calculations, each chromophore is considered as a single unit that does not overlap or exchange electrons with other chromophores. Each chromophore has a characteristic set of electronic transitions, the properties of which are obtained from experiment, where available, and from molecular orbital calculations on models of individual units. The interaction of these transitions leads to a series of delocalized transitions that are then used to calculate the CD spectra. Application of this method to the polypeptide backbone has been effective in describing the CD spectra of certain secondary structures of proteins (*5*). The calculations of CD spectra of whole proteins carried out by Hirst (*6*) and by

† Supported by NIH Research Grant GM22994.

* Corresponding author: E-mail rww@lamar.colostate.edu; Fax (970) 491-0494.

‡ Colorado State University.

§ University of Colorado Health Sciences Center.

|| University of Utah.

[1] Abbreviations: CD, circular dichroism; BPTI, bovine pancreatic trypsin inhibitor; IPTG, isopropyl thiogalactoside; PDB, Protein Data Bank; BM, bond monopoles; VM, vibronic monopoles; DBM, Debye−Bohr magneton, a unit of rotational strength equal to 0.9273×10^{-38} cgs unit; D, Debye, a unit of dipole moment equal to 10^{-18} cgs unit; $\Delta\epsilon$, difference in the molar absorption coefficients of left and right circularly polarized light; rms, root mean square.

10.1021/bi990516z CCC: $18.00
Published on Web 07/21/1999

Woody and Sreerama (*7*) using the matrix method, and by Bode and Applequist (*8*) using the dipole interaction model, have been moderately successful.

Inclusion of side-chain chromophores should be necessary in order to achieve a high degree of accuracy in the calculation of protein CD spectra. In particular, consideration of aromatic groups is of primary importance. Estimating the aromatic side-chain contributions to far-UV protein CD from experiment is a difficult problem due to their nature. In proteins only about 10% of the residues are aromatic, and their environment is particularly important in determining their CD contributions. Strong coupling of aromatic transitions can occur if two or more aromatic rings are in close proximity, thus intensifying the aromatic CD. Such interactions are more important than the number of aromatic residues in a protein in determining the aromatic CD. The near-UV CD of proteins, where peptide contributions are insignificant, is dominated by aromatic contributions, and each type of aromatic residue gives rise to CD bands at characteristic wavelengths. It is possible to detect small local conformational changes near aromatic side chains that are manifested in changes in near-UV CD, while the far-UV CD remains unaltered within experimental error. Replacing an aromatic residue by another aromatic or nonaromatic residue in proteins may also give rise to a difference in the near-UV CD that can be detected. Aromatic residues can also make significant contributions in the far-UV CD of proteins, for example, the positive CD band observed in a number of proteins at 230 nm (*9, 10*), where most prevalent polypeptide conformations give negative CD. This, and the cases where the aromatic contributions do not lead to obviously anomalous CD, complicate the secondary structural analysis.

In this paper, we present the development of a computational model for incorporating the tyrosine and phenylalanine side chains in the existing framework of CD calculations. We have also considered the disulfide transitions in our calculations. We have considered four transitions for each of the aromatic chromophores, denoted L_b, L_a, B_b, and B_a in Platt's notation (*11*). The computational model was used to calculate the CD spectra of bovine pancreatic trypsin inhibitor (BPTI), a small protein with 58 residues, four phenylalanines, four tyrosines, and three disulfide bonds. A previous study on the importance of aromatic CD contributions in BPTI has been published (*12*). The calculated CD spectra of the wild-type and mutant BPTI are compared with experiment.

MATERIALS AND METHODS

Calculation of Rotational Strengths. We utilize the matrix method (*4*) as modified by Goux and Hooker (*13*) for calculating the rotational strengths. The protein molecule is considered to be a set of nonoverlapping chromophores, and each chromophore is treated as an independent unit with distinct and well-characterized electronic transitions. The energies of these transitions form the diagonal elements of the interaction energy matrix, **V,** and the interactions between different transitions form the off-diagonal elements, representing mixing of excited states within different groups or within a single group under the influence of the time-average field of the rest of the molecule. The properties of the transitions necessary to calculate the off-diagonal elements of **V**, such as transition dipole moments, ground-state charges, and transition monopole charges, are obtained from spectroscopic studies where available or from molecular orbital calculations on models representing individual groups. For disulfide transitions, where the transitions are dependent on the torsion angle of the disulfide bond, we use the analytical expressions developed by Woody (*14*). Diagonalization of the matrix **V** yields the eigenvalues and eigenvectors for the composite transitions of the molecule, the protein. The eigenvalues are the energies of the composite transitions, and the eigenvectors give the mixing coefficients describing contributions of the excited states of individual groups to the delocalized excited states of the composite system. The eigenvectors are then used to calculate the rotational strengths corresponding to each excited state of the composite molecule, in the dipole-velocity formalism:

$$R_{oi} = \frac{e^2\hbar^3}{2mcE}(\psi_0|\nabla|\psi_i) \times (\psi_0|\mathbf{r} \times \nabla|\psi_i)$$

where E is the energy of the transition from the ground state, represented by wave function ψ_0, to the excited state, represented by ψ_i; m is the mass of an electron; e is the charge on the electron; c is the speed of light; and $\hbar = h/2\pi$, h being Planck's constant.

The transition charge densities are represented by a set of point charges, determined from the CNDO/S wave functions (*15*), which reproduce the transition dipole moment for the $\pi\pi^*$ transitions or the electric quadrupole moment for the $n\pi^*$ transitions. The elements of the matrix **V** representing interactions between transitions are then evaluated by using Coulomb's law:

$$V_{ioa;\,job} = \sum_s \sum_t \frac{q_{isoa} q_{jtob}}{R_{is,jt}}$$

where q_{isoa} is the monopole s associated with the transition o $\rightarrow$ a on chromophore i, and $R_{is,jt}$ is the distance between monopoles *is* and *jt*.

Calculation of CD Spectra. The CD spectra were calculated from the rotational strengths and the excited-state energies of the composite transitions obtained from the matrix formalism described above. Gaussian band shapes were assumed for all transitions. The relationship between the mean residue ellipticity and the rotational strength for the kth transition is given by

$$\Delta\epsilon_k = \frac{2.278\, R_k \lambda_k}{\Delta_k}$$

where $\Delta\epsilon_k$ is the molar CD (M^{-1} cm^{-1}) per residue due to transition k, R_k is the rotational strength in DBM, λ_k is the wavelength of the transition, and Δ_k is the half-bandwidth. The half-bandwidth is related to the band position by the empirical relation (*16*):

$$\Delta_k = \kappa\, \lambda_k^{1.5}$$

where the value of 0.0045 for the constant κ produces amide absorption spectra that are similar to those measured by Nielsen and Schellman (*17*).

Preparation of BPTI Mutants. Mutant forms of BPTI were produced in *Escherichia coli* by use of the expression system

described by Goldenberg (*18*). The proteins were produced as fusion proteins with the leader peptide of the *E. coli* OmpA protein and exported to the periplasmic space, where the leader sequence was removed and the polypeptides folded to yield native protein. Mutations in the BPTI sequence were constructed by the method of Kunkel (*19*).

For the preparation of all proteins, *E. coli* HB101 bacteria carrying the appropriate plasmid were grown at 30 °C with 12−14 L of fresh air/min passing through the culture, as described by Zhang and Goldenberg (*20*). Four hours after inoculation, expression was induced by the addition of IPTG (0.2 mM). After an additional 18 h, the medium was cooled to 10−15 °C and concentrated 5-fold by use of a Kontes hollow fiber system. The concentrated bacteria were then centrifuged to yield a wet paste, which was stored at −70 °C until needed.

Purification of the BPTI mutants was performed as previously described (*20*). The concentrations of the purified BPTI mutants were determined spectrophotometrically, assuming an extinction coefficient at 280 nm of 5400 M^{-1} cm^{-1} for wild type and phenylalanine mutants and 4050 M^{-1} cm^{-1} for the tyrosine mutants.

Experimental CD Spectra. The CD spectra were collected at the University of Colorado Health Sciences Center on an Aviv model 62 DS spectropolarimeter equipped with a thermoelectric temperature control unit regulated to 0.1 °C (Lakewood, NJ). Samples were typically prepared at concentrations of 1 and 0.1 mg/mL for collection of near-UV and far-UV spectra, respectively, and loaded into strain-free quartz cells. The CD spectra were collected in 10 mM phosphate buffer at pH 7.0 and the sample temperature was held at 25 °C. Data were collected at either 0.25 or 0.5 nm intervals using averaging times of 4−8 s.

Structures of Wild-Type and Mutant BPTI. The coordinates of the crystal structures of the different crystal forms of BPTI (PDB codes 4pti, 5pti, and 6pti) were obtained from the Protein Data Bank (*21*). Positions of the chromophores that were needed to compute the interaction energy matrix elements were obtained by energy minimization of the crystal structures. Even though the structures used were refined at high resolution, some energy minimization was deemed to be desirable for two reasons. First, hydrogen atoms, not included explicitly in the X-ray diffraction refinements, were introduced in our calculations, with standard bond lengths and bond angles. Second, the mutant structure was generated by replacing the Phe or Tyr side chain by that of Leu. Energy minimization is desirable in structures thus generated to permit the wild-type structure to relax. It also seemed desirable to apply the same minimization protocol and force field to the original wild-type structure before calculation of the CD spectrum. The protocol followed for minimization involved 100 steps of steepest descent, followed by 25 steps of conjugate gradient minimization, which was followed by another 100 steepest descent steps and would allow small rearrangements of atomic positions. All minimizations were carried out with AMBER potentials and the InsightII 95.0/Discover 2.97 molecular modeling software (Biosym Technologies). Missing residues (disordered in the crystal structure; residues 57 and 58 in 6pti) were not replaced. The structure of each mutant protein was obtained by replacing the appropriate tyrosine (or phenylalanine) residue from the wild-type structure by leucine, without otherwise altering the protein structure, and energy-minimizing the resulting structure.

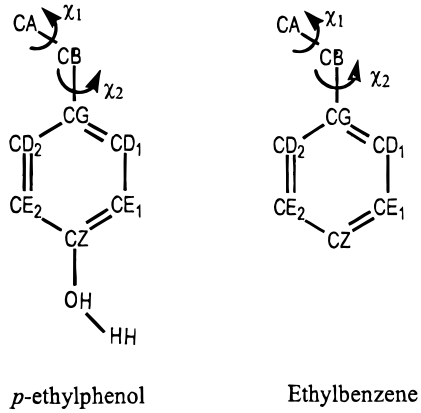

FIGURE 1: Atom labels for the model compounds for tyrosine and phenylalanine, respectively. (Nonpolar hydrogen atoms are not shown.)

RESULTS

Description of Model Compounds and Transitions. The most important step in this calculation is the evaluation of the interaction energy matrix **V**. The transition density is represented as a set of point charges that reproduces the transition dipole moment for the $\pi\pi^*$ transitions or the electric quadrupole moment for the $n\pi^*$ transitions. The transition parameters of the chromophores were obtained from semiempirical molecular orbital calculations on model compounds. The peptide group was modeled by the compound *N*-methylacetamide and the transition properties were calculated by the CNDO/S method (*15*) and deorthogonalized wave functions (*22, 23*). Four transitions per peptide group were included in the calculations. They are the $n\pi^*$ at 220 nm, $\pi_0\pi^*$ (NV_1) at 190 nm, $n'\pi^*$ at 180 nm, and $\pi_+\pi^*$ (NV_2) at 140 nm. Details of the model and the transition monopoles are described by Manning and Woody (*5*).

The aromatic side chains of tyrosine and phenylalanine were modeled by *p*-ethylphenol and ethylbenzene, respectively, and are shown schematically in Figure 1. These aromatic chromophores are derivatives of benzene and the origin of their absorption spectra is similar to that of benzene. In benzenoid chromophores, the absorption spectrum above 180 nm is characterized by three absorption bands. These three bands are due to the four lowest energy transitions involving the excited states analogous to the B_{2u}, B_{1u}, and E_{1u} states of benzene, which are denoted by L_b, L_a, B_b, and B_a (*11*). The transition monopoles and the ground-state charges of these chromophores were calculated by deorthogonalized CNDO/S wave functions of the model compounds, coordinates of which were obtained by standard bond lengths and bond angles. The dihedral angle χ_2 (Figure 1) was taken to be 90°, which corresponds to the most favorable value for aromatic amino acid side chains (*24, 25*). The ground-state charges on nonpolar hydrogens were added to the atom to which the hydrogen is bonded. To maintain charge neutrality, the surplus charge was added to the C_β atom. The ground-state charges were multiplied by a scaling factor to match the experimental dipole moment and were placed at the atomic centers. The monopole changes are given in Tables 1 and 2.

Table 1: Ground-State Charges and Transition Monopole Charges for Phenylalanine[a]

atom	ground-state charges	transitions			
		L_b	L_a	B_b	B_a
CB	0.1240		−0.0207		−0.0516
CG	−0.0193		−0.2632		−0.6580
CD1	−0.0200	0.0858	−0.2131	−0.7263	−0.5328
CD2	−0.0200	−0.0858	−0.2131	0.7263	−0.5328
CE1	−0.0193	0.0114	0.2416	−0.6271	0.6039
CE2	−0.0193	−0.0114	0.2416	0.6271	0.6039
CZ	−0.0261		0.2269		0.5674

[a] All charges are in 10^{-10} esu. The ground-state charges are placed at atomic centers. The transition monopole charges are placed above and below the plane of the ring (the total monopole charge at a given atom is twice the value given in the table; see text for details). The ground-state dipole moment is 0.37 D. The magnitudes of the transition dipole moments for L_b, L_a, B_b, and B_a transitions are 0.4, 2.7, 6.7, and 6.9 D, respectively. The transitions were located at 257 nm (L_a), 206 nm (L_a), 188 nm (B_b), and 188 nm (B_a).

Table 2: Ground-State Charges and Transition Monopoles for Tyrosine[a]

atom	ground-state charges	transitions			
		L_b	L_a	B_b	B_a
CB	0.1563		−0.0209		−0.0476
CG	−0.0059		−0.1891		−0.4297
CD1	−0.0061	0.1441	−0.2763	−0.6632	−0.6281
CD2	−0.0061	−0.1441	−0.2763	0.6632	−0.6281
CE1	0.0400	0.1181	0.2858	−0.6722	0.6495
CE2	0.0400	−0.1181	0.2858	0.6722	0.6495
CZ	0.5981		0.1784		0.4054
OH	−1.7532		0.0126		0.0291
HH	0.9369				

[a] All charges are in 10^{-10} esu. The ground-state charges are placed at atomic centers. The transition monopole charges are placed above and below the plane of the ring (the total monopole charge at a given atom is twice the value given in the table; see text for details). The ground-state dipole moment is 1.8 D. The magnitudes of the transition dipole moments for L_b, L_a, B_b, and B_a transitions are 1.2, 2.8, 6.5, and 6.3 D, respectively. The transitions were located at 278 nm (L_a), 229 nm (L_a), 193 nm (B_b), and 192 nm (B_a).

The transitions L_b, L_a, B_b, and B_a are due to the aromatic ring and are of the $\pi-\pi^*$ type. The transition monopoles are, therefore, mainly located on the atoms of the aromatic ring. Even though the CNDO/S method predicts small transition monopoles on the other atoms, they are negligible in comparison with those on the ring atoms. The transition monopoles were adjusted to give the proper polarization for the transition dipole moment; for L_b and B_b transitions, the polarization is along the short axis of the molecule, and for L_a and B_a transitions, the polarization is along the long axis. The transition monopoles were halved and placed above and below the aromatic ring, mimicking the $\pi-\pi^*$ transition. The distances used were 0.9583 and 0.6861 Å for carbon and oxygen, respectively (*26*). The transition monopole charges (in 10^{-10} esu) for the L and B transitions of phenylalanine and tyrosine side chains are given in Tables 1 and 2. The wavelengths and the electric transition dipole moments are also given in these tables.

The L_b and L_a transitions in benzene are electrically forbidden, while in substituted benzenes they are weakly allowed. The L_b and L_a excited states of benzene have three units of orbital angular momentum corresponding to three nodal planes in the perimeter model (*11*). For the L_b transition the nodal planes, corresponding to zero transition density, pass through the atomic centers, and the antinodal planes, which correspond to maximum transition density, pass through the bond centers. The L_a transition has nodal planes at the bond centers and antinodal planes at the atomic centers. At the atomic centers, the L_b transition in benzene has zero transition charge density and in the substituted benzenes small transition density. However, the transition densities at the bond centers could be large as they lie along the antinodes. Therefore, to properly model the L_b transition, we introduced transition monopoles at the midpoints of bonds of the aromatic ring. We will henceforth refer to these transition monopoles as bond monopoles (BM). The BM were computed from the deorthogonalized CNDO/S wave functions as

$$BM_{ab} = \sum_k A_k BM^k_{ab} = \sum_k A_k \{\sum S_{ij} \sqrt{2}\,(C_{0i}\,C_{1j} + C_{1i}\,C_{0j})\}$$

Here, BM_{ab} is the charge of the transition monopole at the center of the bond between atoms a and b; A_k is the coefficient of configuration k; i and j are the orbitals belonging to the atoms a and b, respectively; S_{ij} is the overlap integral between orbitals i and j; C is a molecular orbital coefficient; and 0 and 1 refer to the ground and excited states, respectively. The BM for the tyrosine and phenylalanine side chains are given in Table 3. The BM alternate in sign around the circumference of the benzene ring, leading to zero transition dipole moment due to BM. They were divided by 2 and placed above and below the plane of the aromatic ring; charges of equal magnitude and like sign were placed 0.9583 Å above and below the midpoint of the bond. For the L_a transition the bond monopole charges are small in comparison to the transition monopole charges at the atomic centers, as they lie near the nodal planes, and were not considered.

Table 3: Bond Monopole Charges for Tyrosine and Phenylalanine L_b Transition[a]

bond	phenylalanine	tyrosine
CG−CD1	0.7061	0.7150
CD1−CE1	−0.7061	−0.7150
CE1−CZ	0.7234	0.7206
CZ−CE2	−0.7234	−0.7206
CE2−CD2	0.7227	0.7008
CD2−CG	−0.7227	−0.7008

[a] The monopole charges were placed 0.9583 Å above and below the plane of the aromatic ring. The total monopole charge at a given atom is twice the value given in the table; see text for details.

Among the four aromatic transitions considered, only B_b and B_a are fully allowed. The partially allowed L transitions derive a part of their intensity from the B transitions through vibronic mixing. In benzene, the L_b and L_a bands in benzene are forbidden but still have oscillator strengths of 0.0014 and 0.094 (*27*), respectively, which are derived from the vibronic mixing with the B transitions. The vibronic coupling is quite significant in the L_a band and we have attempted to model this contribution by a set of vibronic monopoles (VM). The VM charges were calculated by assuming that the two orthogonally polarized B components make equal contributions. The transition dipole moments due to the parallel and perpendicular vibronic components of benzene will be equal in magnitude and lie along the two axes defining the plane

Table 4: Vibronic Monopole Charges for Phenylalanine L_a Transition[a]

atom	parallel component	perpendicular component
CG		0.0849
CD1	−0.1453	0.1697
CD2	0.1453	0.1697
CE1	−0.1453	−0.1697
CE2	−0.1453	−0.1696
CZ		−0.0849

[a] Monopole charges were placed 0.9583 Å above and below the plane of the aromatic ring (the total monopole charge at a given atom is twice the value given in the table). The vibronic monopole charges for the tyrosine side chain were obtained by scaling those for phenylalanine by a factor (0.5989) determined by the ratio of the energy difference between the L_a and the B transitions of benzene and phenol. The parallel component is along the long axis (CZ−CG) and the perpendicular component is along the short axis (midpoint of CD1 and CE1 to midpoint of CD2 and CE2) of the aromatic ring.

of the aromatic ring. The magnitude of the vibronic transition moment for each component is calculated as (*28*)

$$|\mu| = 0.46114\sqrt{f\lambda}$$

where f is the vibronic oscillator strength divided equally between the two components and λ is the wavelength of the L_a transition. Vibronic components in ethylbenzene, the model for the phenylalanine side chain, were taken to be the same as for benzene, while those for phenol (tyrosine side chain) were scaled by a factor (0.5989) determined by the ratio of the energy difference between the L_a and the B transitions of benzene and phenol. The VM for the tyrosine and phenylalanine side chains are given in Table 4. These were also placed 0.9583 Å above and below the plane of the aromatic ring.

Disulfide transitions also contribute to the near-UV CD in proteins. In the near UV, the disulfide group has two $n\sigma^*$ transitions that give broad absorption bands between 250 and 260 nm. The disulfide transition parameters, however, depend on the χ_{SS} angle and, because of the large bandwidth, disulfide CD bands can extend beyond 300 nm. BPTI has three disulfide bonds, and we have included their contributions in our calculations. The energies of the transitions were obtained from the empirical expressions of Kurapkat et al. (*29*) that relate these energies to χ_{SS}, based upon data for model disulfides. The transition monopole charges were obtained from the analytical expressions of Woody (*14*).

BPTI: Different Crystal Forms. BPTI is a small compact globular protein with 58 amino acid residues and has been extensively studied. It has four tyrosine and four phenylalanine residues, and many of these side chains are in close proximity. Such aromatic clusters are important since they lead to mixing of optical transitions, thus enhancing the CD signal. The CD spectrum of BPTI is quite unusual. It has a sharp negative band at 200 nm and a shoulder at 220 nm. It also has a significant negative band at 278 nm. BPTI has been crystallographically characterized in three different crystal forms (*30*−*32*), and structures corresponding to the different crystal forms are available in the Protein Data Bank (*21*). Comparison of the calculated CD spectra (based on the crystal structures) to CD spectra measured in solution should indicate which structure is most similar to that found in solution. Comparison of the three structures indicates that

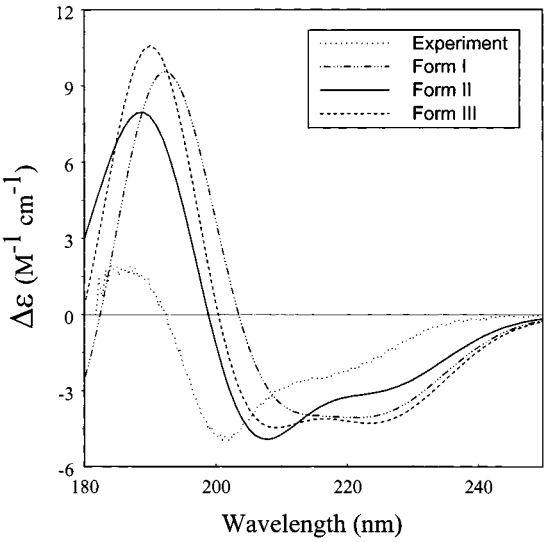

FIGURE 2: Far-UV CD spectra of BPTI: Comparison of the calculated spectra for different crystal forms with the experiment.

the backbone of the protein in the three forms is quite similar. There are, however, differences in the conformations of the side chains located at the surface that are attributed to the effects of crystal packing and intermolecular interactions. For one buried residue, Y21, a large displacement was observed between forms II and III (*32*). In these three crystal forms, the patterns of intermolecular hydrogen bonds are virtually identical, the temperature factors correlate well, and the secondary structures remain unchanged. The dynamic nature of the structure of BPTI is indicated by the differences between the three crystal structures.

The X-ray structure indicates that the protein does not have any unusual structural features and, according to the DSSP assignments (*34*), has approximately 20% α-helix, 24% β-sheet, 23% turns, and 33% other secondary structures. Kosen et al. (*35*) performed an analysis of the BPTI CD spectrum with results that did not agree with the X-ray structure. The methods for analyzing protein CD spectra have since been improved (*36*). The secondary structure analysis of the BPTI CD spectrum by SELCON2 (*37*) gives a relatively poor solution: 21% α-helix, 21% β-sheet, 28% turns, and 41% other; sum of fractions 1.12. Even though the fractions of α-helix and β-sheet determined by CD agree with X-ray values reasonably well, the CD-estimated sum of fractions is 1.12. For a good analysis of protein CD spectrum, the sum of fractions obtained by SELCON2 should be between 0.95 and 1.05, corresponding to a 5% error. The poor analyses of the BPTI CD spectrum could be due to interference from aromatic contributions. A previous calculation of the CD spectrum of BPTI demonstrated the importance of aromatic contributions in obtaining reasonable agreement with experiment (*12*).

The calculated far-UV CD spectra of the three forms of BPTI, corresponding to the energy-minimized X-ray structures, are shown in Figure 2 along with the experimental CD spectrum. All three forms are predicted to have a strong positive band at 190 nm while the experimental spectrum has a small positive band. Theory predicts a negative band around 200 nm with a shoulder for form II, double minima at 200 and 222 nm for form III, and a broad negative band from 200 to 220 nm for form I. The best agreement with

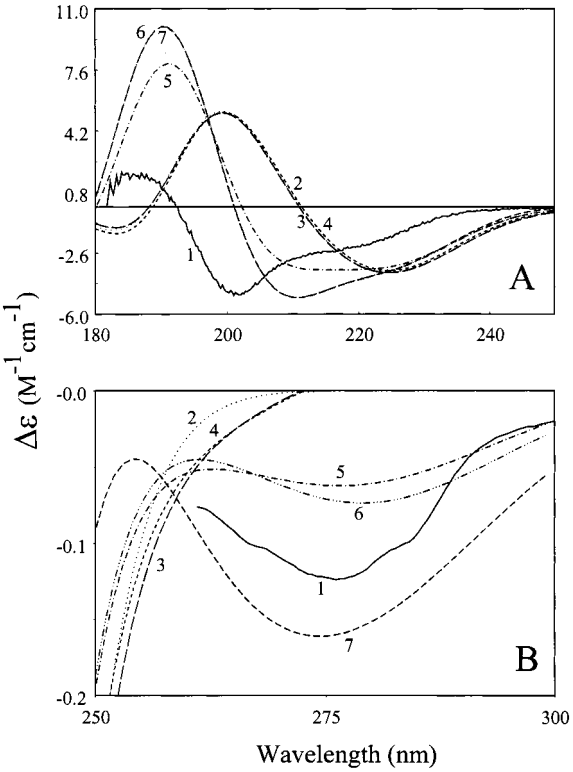

FIGURE 3: Effect of different aromatic transitions on the calculated CD spectra of BPTI: Curve 1 (—), experimental; curve 2 (•••), calculated with peptide transitions only; curve 3 (— — —), calculated with peptide and L_b transitions; curve 4 (---), calculated with peptide, L_b, and L_a transitions; curve 5 (— • —), calculated with peptide, L_b, L_a, and B transitions; curve 6 (— ••• —), calculated with peptide, L_b, L_a, and B transitions, B transition parameters from CNDO/S; curve 7 (— — —), calculated with peptide, L_b, L_a, and B transitions, B transition parameters from CNDO/S, and disulfide transitions. (A) Comparison of calculated and experimental far-UV CD spectra. (B) Comparison of calculated and experimental near-UV CD spectra.

experiment was obtained for form II, which is a joint neutron and X-ray refinement structure and the form for which the highest resolution has been attained. Comparison of the calculated and experimental CD spectra indicates that the crystal structure of form II is most similar to that found in solution. We used this structure in generating the mutant structures and examining the effects of different transitions on the BPTI CD.

We have also computed the CD spectra of these three forms of BPTI by considering only the peptide transitions, which gives an estimate of the main-chain contributions to the CD spectrum. The CD spectra due to the peptide chromophores are very similar for the three forms (data not shown). This is not surprising, given the fact that the main-chain conformations in the three forms of BPTI are very similar. The rms differences in the main-chain atoms between the three forms are 0.4−0.5 Å. The minor differences in the side-chain orientations between the three forms lead to the differences between the predicted CD spectra of the three forms.

Contributions from Different Transitions. To examine the effects of different aromatic transitions on the calculated CD spectrum of BPTI, we included the transitions in a systematic manner. Figure 3 compares the results of the calculation with experimental data (curve 1, —) in the far UV (Figure 3A) and in the near UV (Figure 3B). First, the CD spectrum of BPTI (form II) was calculated considering only the peptide transitions, shown in curve 2 (•••). The L_b transitions of tyrosine and phenylalanine were then introduced into the calculation and the resulting CD spectrum is shown in curve 3 (— — —). Introduction of L_a, including vibronic monopoles, gives the CD spectrum shown in curve 4 (---). Curve 5 (— • —) shows the results after the B transitions were introduced into the calculation. The transition monopole charges used in the calculation of CD spectra shown in curves 2−5 were scaled to match the experimental transition dipole moments. Such a correction factor is necessary because the CNDO/S parametrization, with limited configuration interaction and using the dipole length operator to evaluate oscillator strengths (*38*), overestimates the intensities of transitions. Use of a scaling factor is certainly justified for the L_b and L_a transitions, for which we have reliable experimental transition dipole moments. However, the transition dipole moments of the B transitions are less reliable due to uncertainties in the experimental intensities and the presence of two unresolved components, B_b and B_a. We have calculated the BPTI CD spectra without scaling the B transition monopole charges and the results are shown in curve 6 (— ••• —). The CD spectrum calculated with both aromatic and disulfide transitions is shown in curve 7 (— — —).

The transition monopoles given in Table 1 correspond to the calculation of curve 6, with the L_b and L_a transition monopole charges scaled to match the experimental data, and the CNDO/S monopoles for B transitions are presented. The same set was used in calculating the spectra presented in Figure 2 as well. The CD spectrum, calculated with only peptide contributions (curve 2), does not have any CD bands in the near UV since there are no peptide transitions in this region. Incorporation of the L_b and L_a transitions in the calculations leads to slight changes in the near-UV CD but does not change the CD spectrum in the far UV, as indicated by the overlapping of curves 2−4 in Figure 3A. The B transitions and their interactions have the largest impact on the calculated CD spectrum, in both the near- and far-UV regions. The calculated negative band in the near UV is smaller than the experimental intensity ($-0.11\Delta\epsilon$). The far-UV spectrum resembles the experimental spectrum with a negative band at 205 nm and a shoulder at 225 nm. Incorporation of disulfide transitions does not change the far-UV CD, as reflected in the overlapping of curves 6 and 7 in the far UV (Figure 3A), but it does increases the amplitude of the negative band in the near UV (Figure 3B). There is also a slight blue shift of the position of the negative band. Of the calculated CD spectral curves, curves 6 and 7 show the best agreement with the experiment.

BPTI: Mutants. By use of the transition parameters previously described for the amide and disulfide transitions, and described above for the tyrosine and phenylalanine side chains, with the structures of mutants obtained from the wild-type crystal structure, the CD spectra of the Y → L and F → L mutant BPTI were calculated. The calculated spectra are compared with experiment in the near UV in Figure 4 and in the far UV in Figure 5.

Theory predicts a negative CD band for all mutants in the near UV. The amplitudes of the predicted bands are larger than the experimentally observed bands for Y35L, Y23L, F22L, and F45L and are comparable to experiment for the

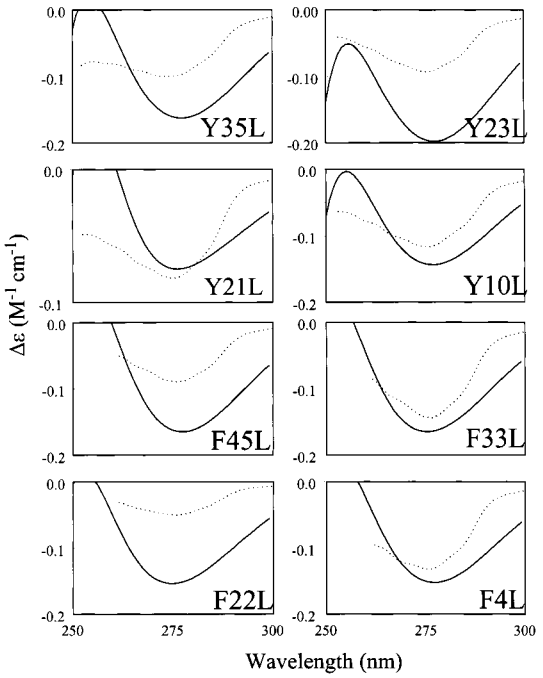

FIGURE 4: Comparison of the calculated (—) and experimental (···) near-UV CD spectra of BPTI mutants.

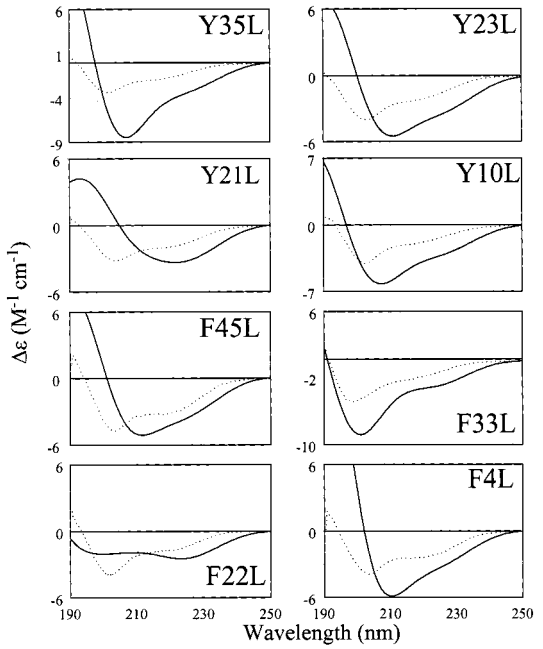

FIGURE 5: Comparison of calculated (—) and experimental (···) far-UV CD spectra of BPTI mutants.

other four mutants. According to the experimental data, F22L has the smallest CD intensity in the near UV, F33L has the largest, and other mutants have intermediate intensities. Our calculations predict similar intensities for all mutants. In our calculations, the near-UV CD is dominated by disulfide contributions as was seen in Figure 3B. When we ignored the disulfide transitions we obtained much smaller amplitudes in the near-UV CD for all mutants, with the smallest amplitude for F22L and the largest for F33L. Thus our calculations of the aromatic contributions reproduce the observed trend in near-UV intensity.

The far-UV CD is more interesting than the near-UV CD due to the importance of aromatic contributions in BPTI. The calculated and experimental CD spectra of mutant BPTI in the far-UV region are shown in Figure 5. The experimental CD spectra of all mutants show the characteristic negative band around 200 nm, which was predicted for the majority of the mutants. The best agreement between theory and experiment was obtained for Y10L, F45L, and Y23L, where the general shape and the amplitudes of the CD bands of the predicted spectra were comparable to the experimental data. For F33L and Y35L the shapes of the CD spectra were predicted well but the predicted amplitudes were about twice as large as experiment. The predicted spectra for Y21L and F22L showed the greatest disagreement with the experiment, failing to reproduce the shape of the far-UV spectrum. In general, theory predicts a larger amplitude than experimentally observed, and the long-wavelength shoulder is predicted better than the negative band at 200 nm.

DISCUSSION

We have developed a set of internally consistent transition parameters for the tyrosine and phenylalanine side-chain chromophores and combined them with the previously described peptide transitions and disulfide transitions in the origin-independent matrix method to compute the CD spectra of BPTI and eight mutants of BPTI in which an aromatic residue (either tyrosine or phenylalanine) is replaced by leucine. The crystal structure of the wild-type BPTI was used in the calculations, and the mutant BPTI structures were obtained by replacing the appropriate residue by leucine. Energy-minimized structures were used for computing the CD spectra. The experimental CD spectra of the wild type and the eight mutants of BPTI were compared with the computed CD spectra.

In general, reasonable agreement between experimental and theoretical CD spectra was obtained. Our calculations emphasize the importance of aromatic contributions to the CD of BPTI. The CD spectrum calculated from only the peptide transitions differed a great deal from the experimental CD. The inclusion of aromatic transitions, particularly the B transitions, improves the theoretical spectrum and decreases the disagreement with experiment. The CD spectra were also computed for the three crystal forms of BPTI. The peptide contributions to CD in the three crystal forms are similar, due to very similar main-chain conformation in the crystal structures (rms differences less than 0.5 Å). Slightly different side-chain orientations of the aromatic residues in the three structures, due to different packing interactions, leads to different calculated CD for the three structures of BPTI. This shows the sensitivity of CD to relative orientations of the aromatic side chains in BPTI. Even small deviations in the aromatic side-chain positions can lead to noticeable changes in the computed CD spectra. The structure of form II gave the CD spectrum that is most similar to experiment and was used to generate the mutant structures. This is probably because this is a superior structure obtained at atomic resolution and including hydrogen atom positions.

In these calculations we have considered the four lowest energy $\pi-\pi^*$ transitions of the benzenoid chromophores of the tyrosine and phenylalanine side chains. Of these, the L_b transition is in the near UV where peptide contributions are insignificant. The L_b transition of tyrosine occurs at 278 nm and that of phenylalanine at 257 nm. The L_b of tyrosine is

the stronger transition and, together with the disulfides, is mainly responsible for the near-UV CD of BPTI. The experimental CD spectra of wild type and the mutants of BPTI examined here have a negative band about 280 nm. Computed CD spectra also predict a negative band in this region.

The L_a and B transitions of tyrosine and phenylalanine are in the far-UV region, overlapping with the peptide transitions. The L_a transition of tyrosine is at 229 nm and that of phenylalanine is at 208 nm. A net positive rotational strength was predicted for the wild-type BPTI due to these two transitions: tyrosine L_a, 1.23 DBM, and phenylalanine L_a, 3.65 DBM. However, the Gaussian band shapes and the bandwidths used in this study render the CD due to L_a transitions hidden under the CD due to peptide $n\pi^*$ and $\pi\pi^*$ transitions. The B transitions are located about 192 and 188 nm, respectively, for tyrosine and phenylalanine and are much stronger transitions (transition dipole moments $\approx$ 6 D). They couple strongly with the peptide transitions, due to their near degeneracy with the peptide $\pi\pi^*$ transition. This makes it impossible to resolve the peptide and aromatic B transitions from the protein CD spectrum. According to our calculations, only upon including the B transitions do we observe any similarity with the experimental data, suggesting that the far-UV CD in BPTI is significantly influenced by the B transitions.

BPTI presents a good model for examining the various aspects of the aromatic contributions to protein CD. It is a small protein with a high percentage of aromatic residues that are in close proximity, enhancing the coupling of aromatic transitions. The CD spectrum of BPTI, with a small positive band at 190 nm, a sharp negative band at 200 nm, and a shoulder at 225 nm, is atypical of proteins, and the standard methods of CD analyses do not give satisfactory results. Proteins for which aromatic contributions interfere with protein CD analyses are known, and our calculations indicate the importance of aromatic contributions to BPTI CD.

To examine the origin of aromatic coupling in BPTI and to understand the effects of point mutations of aromatic residues, we carried out an analysis of the distances between aromatic rings in BPTI. On the basis of distances between centers of aromatic rings less than 9 Å, the aromatic residues were grouped into three clusters. Cluster 1 comprised residues F4, Y23, and F45; residues Y10, F33, and Y35 formed cluster 2; and cluster 3 was formed by the remaining two aromatic residues Y21 and F22. The residues in clusters 1 and 2 are separated by distances greater than 12 Å. Cluster 3 acts as a link between clusters 1 and 2, with at least one residue belonging to clusters 1 and 2 at a distance less than 9 Å from F22 or Y21, belonging to cluster 3.

Each aromatic residue has a different environment; the interactions of each aromatic side chain with other aromatic side chains are unique. Residues F4 and Y10 are at one end of the molecule, interacting with the least number of aromatic side chains. At the other end is residue F22 interacting with the largest number of aromatic side chains. The aromatic rings of Y35 and F33 are in close proximity (6−7 Å), as are the rings of F33 and F22 (5−6 Å). These aromatic residues have varying degrees of solvent exposure. The solvent accessibility values given by DSSP (*32*) for residues Y23, F45, Y35, F33, and F22 are 15, 16, 17, 19, and 22 Å^2, respectively. These residues are least exposed to solvent. The most exposed aromatic residues are Y10 and Y21, with solvent accessibility surfaces of 85 and 78 Å^2, respectively. Residue F4 has an intermediate solvent exposure.

Our calculations were performed on a single structure obtained by energy minimization, whereas the experimental data correspond to the time-averaged structure of an ensemble of structures in solution. The structure used in the calculation of wild-type CD corresponds to the solid-state structure as seen by X-ray diffraction, which could be assumed to be similar to the solution structure. The same cannot be said of the structures used in calculating the CD of mutant BPTI, as they were derived from the wild-type crystal structure for lack of experimental structures for the mutants. This is where the environment of the mutated residue becomes important. If a residue exposed to solvent is mutated, the local changes in the structure expected may be small. On the other hand, mutation of a buried residue may lead to some local rearrangements of the structure. Our calculations of the CD spectra of different crystal forms of BPTI suggest that small differences in the orientations of the aromatic side chains could lead to some changes in the CD spectra of the protein. Changes in the local structure of the protein upon mutation of an aromatic residue may lead to changes in the CD spectra. The differences between the wild-type and the mutant CD spectra would then be the sum of the changes in CD due to local structural rearrangements and those due to the different interactions between aromatic side chains in the mutant, assuming that the main-chain conformations are similar in both the wild type and the mutant.

The calculated CD spectra of the mutant BPTIs agree reasonably well with experiment for the mutants Y10L, Y23L, F45L, and F4L. The calculated intensities were much larger than experiment in F33L and Y35L. The worst agreement between experiment and theory was obtained for F22L and Y21L. In all cases the agreement in the wavelength range 220−250 nm was much better than that in the range 190−220 nm, which is consistent with the earlier observation that the B transitions have a greater influence on the CD spectra of BPTI.

The differences in the experimental and the theoretical CD spectra of the mutants could be rationalized from the environment of the mutated residue and the resulting aromatic interactions. The residues F22 and Y21, mutants of which give the worst agreement between experimental and theoretical CD spectra, form a cluster and have the largest number of aromatic interactions. Mutation of one of these residues may alter the local structure, and our minimization protocol may not have reproduced this change. The residue pair Y35/F33 also presents a similar situation. Because of their close proximity, the structure used in the CD calculation may be in error. Microsecond time scale conformational changes in Y35G BPTI have been detected by ^{15}N transverse relaxation rates (*39*, *40*), suggesting that the crystal structure (*41*) may represent one of the multiple conformations in solution. Use of a single structure in our calculations is only an approximate representation of the solution state. A better approach to generate the mutant structures would be a restrained molecular dynamics simulation, allowing for local conformational changes surrounding the mutated residue. The molecular dynamics of the wild-

type crystal structure may also give a better basis for CD calculation, as it would incorporate the effects of protein dynamics in the calculation.

CONCLUSIONS

We have presented a model for calculating the contributions of tyrosine and phenylalanine side chains to the protein CD spectra. The chromophores responsible for CD were modeled by *p*-ethylphenol and ethylbenzene, and the transition parameters were calculated from CNDO/S wave functions. The transitions were incorporated into the matrix method along with those for the peptide groups and disulfides, and CD spectra were calculated. The partially allowed L transitions were modeled either by transition charge densities at the center of bonds (L_b) or by vibronic components similar to those of benzene (L_a). The transition parameters obtained from CNDO/S were scaled to match the experimental data, where available, and different methods of scaling were examined. Our model for side-chain chromophores works reasonably well in reproducing experimental spectra for BPTI and its mutants.

There are other factors influencing the CD spectrum, such as environmental factors, contributions from other side chains, high-energy transitions that we have not considered in this calculation, and the dynamic nature of the proteins. We do not have any quantitative estimates of these effects, either from experiment or from theory. We know from theory that the crystal field affects the transition moments of chromophores (*42*), and such effects are to be expected in cases where the chromophore is hydrogen-bonded either intramolecularly or with water molecules. Incorporation of such environmental effects into the calculation framework should improve the comparison between theory and experiment.

REFERENCES

1. Nakanishi, K., Berova, N., and Woody, R. W., Eds. (1994) *Circular Dichroism−Principles and Applications*, VCH Publishers Inc., New York.
2. Sreerama, N., and Woody, R. W. (1994) *J. Mol. Biol. 242*, 497−507.
3. Woody, R. W., and Dunker, A. K. (1996) in *Circular Dichroism and the Conformational Analysis of Biomolecules* (Fasman, G. D., Ed.) pp 109−157, Plenum Press, New York.
4. Bayley, P. M., Nielsen, E. B., and Schellman, J. A. (1969) *J. Phys. Chem. 73*, 228−243.
5. Manning, M. C., and Woody, R. W. (1991) *Biopolymers 31*, 569−586.
6. Hirst, J. D. (1998) *J. Chem. Phys. 109*, 782−787.
7. Woody, R. W., and Sreerama, N. (1999) *J. Chem. Phys.* (in press).
8. Bode, K. A., and Applequist, J. (1998) *J. Am. Chem. Soc. 120*, 10938−10946.
9. Woody, R. W. (1978) *Biopolymers 17*, 1451−1467.
10. Woody, R. W. (1994) *Eur. J. Biophys. 23*, 253−262.
11. Platt, J. R. (1949) *J. Chem. Phys. 17*, 484−495.
12. Manning, M. C., and Woody, R. W. (1989) *Biochemistry 28*, 8609−8613.
13. Goux, W. J., and Hooker, T. M., Jr. (1980) *J. Am. Chem. Soc. 102*, 7080−7087.
14. Woody, R. W. (1973) *Tetrahedron 29*, 1273−1283.
15. Del Bene, J., and Jaffé, H. H. (1968) *J. Chem. Phys. 48*, 1807−1813.
16. Brown, A., Kemp, C. M., and Mason, S. F. (1971) *J. Chem. Soc. A*, 751−755.
17. Nielsen, E. B., and Schellman, J. A. (1971) *Biopolymers 10*, 1559−1581.
18. Goldenberg, D. P. (1988) *Biochemistry 27*, 2481−2489.
19. Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 488−492.
20. Zhang, J.-X., and Goldenberg, D. P. (1997) *Protein Sci. 6*, 1552−1562.
21. Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Brice, M. D., Rogers, J. R., Kennard, O., Shimonouchi, T., and Tasumi, M. (1977) *J. Mol. Biol. 112*, 535−542.
22. Giessner-Prettre, C., and Pullman, A. (1967) *Theor. Chim. Acta 11*, 159−164.
23. Shillady, D. D., Billingsley, F. P., and Bloor, J. E. (1971) *Theor. Chim. Acta 21*, 1−8.
24. Ponder, J. W., and Richards, F. M. (1987) *J. Mol. Biol. 193*, 775−791.
25. Sasisekharan, V., and Ponnuswamy, P. K. (1971) *Biopolymers 10*, 583−592.
26. Woody, R. W. (1968) *J. Chem. Phys. 49*, 4797−4806.
27. Murrell, J. N., and Pople, J. A. (1956) *Proc. Phys. Soc. A69*, 245.
28. Murrell, J. N. (1963) *The Theory of the Electronic Spectra of Organic Molecules*, p 9, John Wiley and Sons Inc., New York.
29. Kurapkat, G., Krüger, P., Wollmer, A., Fleischhauer, J., Kramer, B., Zobel, E., Koslowski, A., Botterweck, H., and Woody, R. W. (1997) *Biopolymers 41*, 267−287.
30. Deisenhofer, J., and Steigemann, W. (1975) *Acta Crystallogr. Sect. B 31*, 238−250.
31. Wlodawer, A., Walter, J., Huber, R., and Sjolin, L. (1984) *J. Mol. Biol. 180*, 301−329.
32. Wlodawer, A., Nachman, J., Gilliland, G. L., Gallagher, W., and Woodward, C. (1987) *J. Mol. Biol. 198*, 469−480.
33. Wlodawer, A., Deisenhofer, J., and Huber, R. (1987) *J. Mol. Biol. 193*, 145−156.
34. Kabsch, W., and Sander, C. (1983) *Biopolymers 22*, 2577−2637.
35. Kosen, P. A., Creigton, T. E., and Blout, E. R. (1981) *Biochemistry 20*, 5744−5754.
36. Greenfield, N. J. (1996) *Anal. Biochem. 235*, 1−10.
37. Sreerama, N., Venyaminov, S. Y., and Woody, R. W. (1999) *Protein Sci. 8*, 370−380.
38. Hansen, A. E. (1966) *Theor. Chim. Acta 6*, 341−349.
39. Beeser, S. A., Goldenberg, D. P., and Oas, T. G. (1997) *J. Mol. Biol. 269*, 154−164.
40. Beeser, S. A., Oas, T. G., and Goldenberg, D. P. (1998) *J. Mol. Biol. 284*, 1581−1596.
41. Housset, D., Kim, K.-S., Fuchs, J., Woodward, C., and Wlodawer, A. (1991) *J. Mol. Biol. 220*, 757−770.
42. Sreerama, N., Callis, P. R., and Woody, R. W. (1994) *J. Phys. Chem. 98*, 10397−10407.

BI990516Z