Analysis of the Optical Rotatory Dispersion of Polypeptides and Proteins. V. A Comparison of Methods[1]

J. P. Carver, E. Shechter, and E. R. Blout

Contribution from the Department of Biological Chemistry, Harvard Medical School, Boston, Massachusetts 02115. *Received January 3, 1966*

Abstract: If quantitative information regarding the conformations of polypeptides and proteins in solution is to be obtained from optical rotatory dispersion data (ORD), certain assumptions must be made. In this paper, the extent to which these assumptions are validated by the results of model studies is considered. The methods used for interpreting the ORD of proteins and polypeptides are considered first. Then the particular case of mixtures of α-helical and random conformations is examined. Quantitative relations between the various rotatory parameters presently in use ($[R']_{233}$, $[R']_{198}$, A_{193}, A_{225}, $(A_{193} - A_{225})$, a_0, and b_0) and the rotational strengths of the α-helical and random conformation Cotton effects are derived. ORD data for poly-α-L-glutamic acid at pH 4.3 and 7.0 are used, for the purposes of illustration, to represent 100 and 0% α helix, respectively. From the relationships between rotational strengths and rotatory parameters, it is possible to determine the extent to which nonconformational alterations in the rotational strengths of the contributing Cotton effects will affect the helix content estimates from any one of the rotatory parameters. Using the preliminary results regarding the nature of side chain effects on the rotatory dispersion, described in the accompanying paper, it is concluded that $(A_{193} - A_{225})$ and b_0 give helix content estimates least sensitive to different side chains. A procedure for estimating α-helix content is suggested using either of these two rotatory parameters. Finally, some previous interpretations of the parameters of the Moffitt and modified two-term Drude equations are considered.

The optical rotatory dispersion (ORD) of a material in solution potentially contains information of a more subtle variety than that available from the common probes of structure in solution—hydrodynamic methods or absorption spectroscopy. If a chromophore, either inherently or as a result of interaction with its environment, is dissymmetric, then there will be Cotton effects associated with the electronic transitions of that chromophore. The characteristic parameters of a Cotton effect are shown in Figure 1; they are: (1) rotational strength[2] (R_{λ_i}); (2) position of its center (λ_i) (for an isolated Cotton effect its crossover point); and (3) its half-width (Δ_i) (approximately half the separation between the extrema). The position of the Cotton effect should be close to that of the maximum of the absorption band of the corresponding electronic transition.[3] The rotational strength is the imaginary part of the dot product of the induced electric dipole transition moment and the induced magnetic dipole transition moment and, as such, as directly related to the asymmetry of the chromophore—whether inherent or induced.[3]

For proteins and polypeptides in general, there exists no satisfactory means of correlating the observed ORD with the molecular structure. Therefore, as an approach to interpreting the ORD of proteins and polypeptides of unknown structure, one first determines the ORD of model compounds with known structures. One then attempts to construct from the ORD curves of the known structures a weighted sum which is identical with the ORD curve of the unknown structure. If all the weights are positive and add to unity, then it is concluded that the unknown structure consists of a mixture of the known structures in proportions equal to their respective weights. It is evident that there are many assumptions implicit in such a procedure for determining structural contents. The purpose of this paper is to state the assumptions involved, to discuss the extent to which these assumptions are inadequate for real structures, and to indicate a general approach to the analysis of ORD data.

In addition to the peptide bonds, for many proteins and some polypeptides there exist other optically active chromophores, such as side chain chromophores, disulfide bonds, and prosthetic groups. The effect which the presence of these additional sources of optical activity has on the estimation of structural content must be considered. Initially, we will restrict the discussion to problems which exist with the simpler systems where the backbone peptide groups are the only optically active chromophores. These problems arise because of two general effects: (1) the distribution of the sizes of structured regions, and (2) the nonconformational alterations in rotational strengths.

(1) Polypeptides. LIII. For the previous paper in this series, see J. P. Carver, E. Shechter, and E. R. Blout, *J. Am. Chem. Soc.*, **88**, 2550 (1966). We are pleased to acknowledge the support (in part) of this work by U. S. Public Health Service Grant AM-07300-01, -02, and -03.

(2) The rotational strength of an isolated optically active transition can also be defined in terms of directly observable properties such as the residue ellipticity, $[\theta']_\lambda$ (deg cm^2 dmole^{-1})

$$R_{\lambda_i}\ (\text{erg cm}^3) \equiv \frac{hc}{48\pi^2 N}\int [\theta']_\lambda \mathrm{d} \ln \lambda \text{ (integration taken through the band)}$$

If the ellipticity has a Gaussian dependence on wavelength, *i.e.*

$$[\theta']_\lambda = [\theta']_{ext} \exp\left[-\left(\frac{\lambda - \lambda_i}{\Delta_i}\right)^2\right]$$

then, from the Kronig–Kramers relations, one may write

$$[R']_\lambda = \frac{R_{\lambda_i}\lambda_i}{\Delta_i}\left\{\exp\left[-\left(\frac{\lambda - \lambda_i}{\Delta_i}\right)^2\right]\int_0^{(\lambda-\lambda_i)/\Delta_i} \exp(y^2)\mathrm{d}y - \frac{\Delta_i}{2(\lambda + \lambda_i)}\right\}$$

Furthermore, one obtains a direct relation between the rotational strength, R_{λ_i}, and the extreme values of: (1) the residue ellipticity, $[\theta']_{ext}$

$$R_{\lambda_i} \cong 1.23 \times 10^{-42}\frac{\Delta_i[\theta']_{ext}}{\lambda_i} \text{ erg cm}^3$$

or (2) the residue rotation, $[R']_{ext}$

$$R_{\lambda_i} \cong 2.02 \times 10^{-42}\frac{\Delta_i[R']_{ext}}{\lambda_i} \text{ erg cm}^3$$

(3) A. Moscowitz in C. Djerassi, "Optical Rotatory Dispersion," McGraw-Hill Book Co., Inc., New York, N. Y., 1960, p 150.

If a polypeptide or protein solution is monodisperse, and if all molecules have either all one structure or all another, then there are only two species and a linear interpolation using a single parameter probably gives a good approximation to the relative amounts of the two species. On the other hand, if each molecule has some residues in one structure and some in another, one must consider the actual distribution of the sizes (number of residues) of the structured regions. Such considerations are important for two reasons: (1) the relative importance of end residues may increase with an increase in the number of structured regions at constant total structural content; and (2) as the number of residues in a structured segment decreases, the rotational strengths of the Cotton effects arising from the chromophores of these residues may no longer be linear functions of the number of residues in the segment. In fact, theoretical calculations[4] for the α helix indicate that the rotation (per residue) increases by about 20% as the number of residues in the segment increases from 10 to 40. Based on this calculation and considerations of the possible influence of end effects, it would appear that for a molecule such as myoglobin, where the number of residues in a helical segment varies from 7 to 24,[5] the effect on helix content estimates from rotatory measurements of the distribution of the sizes of structured regions should be significant.

For model systems such as homopolypeptides the nonconformational alterations in rotational strengths exist as two general effects: (a) different homopolypeptides often do not have identical ORD's in the same solvent under conditions in which they are believed to be entirely of one structure,[6] and (b) the same homopolypeptide, even though it is believed to be entirely of one structure, often does not have the same ORD in solvents of different index of refraction[7] and/or dielectric constant[8] and/or dipole moment.[6] These effects will be referred to below as side chain and solvent effects, respectively.

For globular proteins, as compared with homopolypeptides, there will be many more possible modes of interaction for the peptide bond which could affect the magnitudes of the rotational strengths of the peptide Cotton effects. These additional nonconformational effects may arise from the many different kinds of side chain–side chain and side chain–peptide bond interactions[9] or from differing degrees of solvation producing an effective dielectric constant which varies throughout the molecule. Even if model studies could yield quantitative estimates of these effects, one would first need to know the structure in order to apply such corrections to the observed rotational strengths. Because of the increased number of possible interactions in a globular protein, at present it appears that it will be difficult to *quantitatively* interpret the rotational strengths of the peptide bond Cotton effects of a protein directly in terms of structure.

Despite the complications noted above, it may be possible to find a rotatory parameter which is insensitive to the major part of the nonconformational effects and which, therefore, can be used to estimate structural content. In this paper we use the term "rotatory parameter" to denote either a particular rotational strength or some linear function of the rotational strengths which is used to estimate structural content. Specific examples are b_0 and $(A_{(\alpha,\rho)(193)} - A_{(\alpha,\rho)225})$. From eq 4 and 5 of the accompanying paper[10] it is evident that the rotation at any particular wavelength is a linear function of the rotational strengths of the contributing Cotton effects. It is conceivable that for some wavelengths the rotation could be independent of nonconformational alterations in rotational strengths and therefore provide valid estimates of structural content.

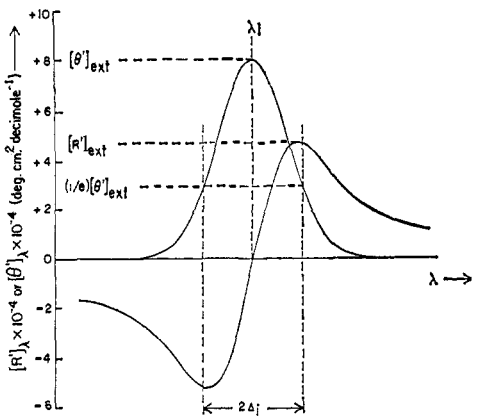


Figure 1. The optical rotatory dispersion curve and associated ellipticity band for an isolated optically active transition. The curves were calculated for a rotational strength of 10^{-38} erg cm^3 and a λ_i/Δ_i ratio of 10. For the meaning of $[R']_{ext}$ and $[\theta']_{ext}$ see ref 2.

Thus, *a priori*, the contribution of the regions of a given structure within a molecule to a rotatory parameter of that molecule will be proportional to the fraction of its residue in those regions if the following four assumptions are valid for that rotatory parameter. (I) The backbone peptide bonds are the only source of optical activity contributing to the rotatory parameter (*i.e.*, absence of extrinsic and side chain Cotton effects affecting this parameter). (II) For residues located within a given structured segment of the molecule, the rotatory parameters must be independent of the number of residues in that segment (*i.e.*, no effect of the distribution of the sizes of structured regions). (III) The rotatory parameters are determined only by the peptide bond conformations. That is, for any given peptide bond, the rotatory parameters must be insensitive to differences in side chains and to changes in the local environment other than those accompanying the transfer of that residue from one structure to another (*i.e.*, absence of nonconformational rotational effects). (IV) The experimental error in the determination of the rotatory parameter must be negligible in comparison with the maximum possible contributions of the various structures. As mentioned above, many of these assumptions will not be strictly valid for real structures—especially globular proteins. Therefore, a practical approach to the estimation of structural contents is to

(4) I. Tinoco, Jr., R. W. Woody, and D. F. Bradley, *J. Am. Chem. Soc.*, 38, 1317 (1963).
(5) J. C. Kendrew, H. C. Watson, B. E. Strandberg, R. G. Hart, D. R. Davies, D. C. Phillips, and V. C. Shore, *Nature*, 190, 666 (1961).
(6) P. Urnes and P. Doty, *Advan. Protein Chem.*, 16, 401 (1962).
(7) J. Cassim and E. W. Taylor, *Biophys. J.*, 5, 553 (1965).
(8) E. Shechter and E. R. Blout, *Proc. Natl. Acad. Sci. U. S.*, 51, 794 (1964).
(9) B. J. Litman and J. A. Schellman, *J. Phys. Chem.*, 69, 978 (1965).
(10) See ref 1.

seek the method of analysis which is least affected by the partial invalidity of these assumptions. In particular, it will be shown that solvent and side chain effects on structural content estimates (part of assumption III) can be minimized by a suitable choice of rotatory parameters.

The next section (I) of this paper consists of a discussion of the various rotatory parameters now in use. In section II, the particular system of α-helical–random conformation mixtures is examined, with particular reference to the problem of determining the relative sensitivities of the various rotatory parameters to nonconformational effects, and a procedure is suggested for helix content determinations. In the final section (III) some interpretations of several of these rotatory parameters are discussed.

I. The Interpretations of Rotatory Measurements

It was pointed out above that if assumptions I–IV are valid, then the contribution of a structure to the observed ORD of a protein or polypeptide will be in proportion to the fraction of its residues in that structure. Under these circumstances, to estimate per cent structure, one only needs parameters linearly related to rotational strengths and not rotational strengths themselves.

The methods of obtaining rotatory parameters from the observed rotatory dispersion fall into three categories: (A) those using the rotation at a particular wavelength (single-wavelength methods); (B) those using rotations over a wide wavelength range but remote from the contributing Cotton effects (dispersion methods); (C) those using not only the rotations remote from the contributing Cotton effects but also the rotations in the wavelength region in which most of these occur (curve-fitting methods).

A. Single-Wavelength Methods. The first category includes the use of $[R']_D$[6,11] (the residue rotation at 589 mμ) and the rotation at an extremum of the observed ORD. The use of $[R']_D$ requires that assumptions I–III be valid for all conformational Cotton effects contributing to the rotation at 589 mμ. If only one rotatory parameter is available, as is the case with $[R']_D$, then structural content estimates will be meaningful only if all residues are in one or other of two structures. If the rotations at several well-separated wavelengths (so that the rotations vary significantly in magnitude) meet assumptions I–IV, then, in principle, it should be possible to estimate the per cent content of as many structures as there are wavelengths plus one.

A disadvantage of using the rotation at an extremum compared to $[R']_D$ is that rotations within one half-width of the center of the Cotton effect (*i.e.*, at the extremum) are much more sensitive to changes in the half-width of the Cotton effect than rotations further removed from the center.[10]

When using the rotation at an extremum of a Cotton effect, one hopes that the contribution of the other Cotton effects at that wavelength will be relatively insignificant (or constant). If this is true, assumptions I–IV need only be valid for this one Cotton effect. If there are significant contributions from other Cotton effects, then the situation is identical with that for $[R']_D$ and assumptions I–IV must be met for all contributing Cotton effects.

B. Dispersion Methods. The one-term Drude,[12] the Moffitt[13] (ME), and the modified two-term Drude[14] (MTTDE) and other two-term Drude equations[15,16] constitute the second category of methods for the extraction of information concerning Cotton effect parameters from ORD. The general approach for the derivation of the visible and near-ultraviolet dispersion methods for ORD analysis is to make some assumptions as to the relative order of magnitude and positions of the conformational Cotton effects and then to derive a two-term expression approximating the contribution of these Cotton effects to the visible and near-ultraviolet rotations. Two-term equations are used since their parameters may be conveniently evaluated by graphical methods. Fortunately, regardless of the manner in which the wavelength-dependent parts of the Cotton effect contributions are approximated to give two terms, the coefficients of these terms are always some linear combination of the rotational strengths of the assumed Cotton effects. These coefficients, then, are the rotatory parameters. The different equations which have been proposed arise from different approximation methods dictated by different assumptions as to the nature of the contributing Cotton effects. The situation, as far as the interpretations of the rotatory parameters derived from the observed rotations is concerned, is the same as for $[R']_D$. That is to say, assumptions I–IV theoretically must be valid for all the conformational Cotton effects contributing in the visible and near ultraviolet regions.

With both the ME and the MTTDE, two helix content parameters are obtained. Thus, in principle, one can estimate the structural contents for a system in which there are three structures present. In essence, this is what one does when one compares the helix content estimates from the two parameters—agreement is interpreted as indicating that the contribution of the "third structure" is zero, and disagreement as indicating that there is a significant contribution from a "third structure."[8,14]

C. Curve-Fitting Methods. The third method is the curve-fitting approach which permits direct estimation of the Cotton effect parameters. When the ORD is available in the wavelength region where the backbone peptide bond Cotton effects occur, one can use a nonlinear, least-squares analysis to relate the observed rotations to the parameters of the component Cotton effects.[10] This method presents a very distinct advantage over A and B in that it yields rotational strength estimates directly for each contributing Cotton effect. Thus, assumption I need not be met since interpretation of the parameters of those Cotton effects which are directly assignable and understood can be made even though the interpretation of the parameters of the others may not be unequivocal. However, assumptions II–IV must be fulfilled at least for those rotational strengths which are to be interpreted in terms of structure.

(11) For reviews see E. R. Blout, ref 3, Chapter 17; and ref 6.

(12) J. A. Schellman, *Compt. Rend. Trav. Lab. Carlsberg*, 30, 363 (1958).

(13) W. Moffitt and J. T. Yang, *Proc. Natl. Acad. Sci. U. S.*, 42, 596 (1965).

(14) E. Shechter and E. R. Blout, *ibid.*, 51, 695 (1964).

(15) K. Yamaoka, *Biopolymers*, 2, 219 (1964).

(16) K. Imahori, *Kobunshi Kagaku, Suppl. 1*, 12, 34 (1963).

The foregoing discussion is sufficiently general that it should pertain to most of the possible structures for polypeptides and proteins. However, in the remainder of this paper we shall discuss specifically the application of these methods to a system in which the α-helical and the random conformation are the only conformations present.[17-19] This restriction, that only α-helical and random conformations are present, constitutes assumption V.

II. Methods for α-Helix Content Estimation

A. Rotation at a Single Wavelength. The use of $[R']_D$ for the estimation of α-helix content has been shown[20] to lead to erroneous results because of the sensitivity of the observed values of $[R']_D$ to solvent effects. That is to say, assumption III is not valid.

For partly random, partly α-helical mixtures it has been suggested[21] that the rotation at 233 mμ, the long wavelength trough in the ORD curve of α-helical synthetic polypeptides, be used as a direct measure of α-helix content. When measurements were extended to shorter wavelengths[22] and the 198-mμ peak discovered, it was suggested that this value might be used in a similar way. Using the Cotton effect parameters (rotational strengths, half-widths, and positions) obtained from the computer analysis (solution 2, Table VI; solution 3, Table IV) and eq 5 of ref 1, one may compute the contribution of each Cotton effect to the observed values of $[R']_{233}$ and $[R']_{198}$. The results of such a calculation for poly-α-L-glutamic acid (pH 4.3 and 7.0) are given in Table I. The superscripts on $[R']_{233}$ refer to the pH at which the data were obtained.

Table I.

Parameter	192 mμ[a]	208 mμ[a]	224 mμ[a]	B[b]	Sum[c]
$[R']_{233}^{4.3}$	+8050	−5,030	−17,600	...	−14,600
$[R']_{198}^{4.3}$	+41,700	+15,500	+8,370	...	+65,600
	198 mμ	217 mμ	235 mμ		
$[R']_{233}^{7.0}$	−3,430	+1,440	+67	+151	−1,770
$[R']_{198}^{7.0}$	−1,270	−1,330	+41	+283	−2,280

[a] Calculated contribution to the parameter from Cotton effects at wavelength indicated. [b] Calculated contribution from background. [c] Sum (rounded off to three significant figures) of calculated contributions from individual Cotton effects.

From the values of the calculated contributions to the rotatory parameters in Table I and the rotational strengths of the contributing Cotton effects obtained from the computer solutions one may write

$$[R']_{233}^{4.3} = 1.98R_{192} + 3.59R_{208} + 10.1R_{224} \quad (1)$$

$$[R']_{198}^{4.3} = 10.2R_{192} - 11.1R_{208} - 4.79R_{224} \quad (2)$$

$$[R']_{233}^{7.0} = 2.43R_{198} + 7.51R_{217} - 5.30R_{235} + 7.93B \quad (3)$$

$$[R']_{198}^{7.0} = 0.900R_{198} - 6.91R_{217} - 3.25R_{235} + 14.8B \quad (4)$$

where $R_{\lambda_i} \times 10^{-42}$ is the rotational strength in erg cm^3 of the Cotton effect at wavelength λ_i (mμ), and B refers to the background parameter of the calculations in ref 10.

Even though 233 and 198 mμ are one half-width away from the centers of their respective Cotton effects, there is still a significant contribution at these wavelengths from other peptide bond Cotton effects. Thus, assumptions II and III must be valid for all six peptide Cotton effects considered above. If, in addition, assumptions I, IV, and V are valid for $[R']_{233}$ and $[R']_{198}$, then one may write

$$X_{[R']_{233}} = \frac{100([R']_{233}^{\text{obsd}} - [R']_{233}^{7.0})}{([R']_{233}^{4.3} - [R']_{233}^{7.0})}$$

$$X_{[R']_{198}} = \frac{100([R']_{198}^{\text{obsd}} - [R']_{198}^{7.0})}{([R']_{198}^{4.3} - [R']_{198}^{7.0})} \quad (5)$$

where X_{P_i} is the per cent α-helix content calculated from parameter p_j.

A useful measure of the sensitivity of a given rotatory parameter to changes in the various rotational strengths is the per cent change in rotational strength R_{λ_i} causing a 1% change in α-helix content estimates from the jth parameter (assuming constant end points).[23] We denote this quantity β_{ij}. The larger $|\beta_{ij}|$ is, the less sensitive the jth parameter will be to alterations in the rotational strength of the Cotton effect at wavelength

(17) The ORD data for PGA used in the calculations below and also in the accompanying paper[1] are reproducible under the conditions of measurement now used in this laboratory but disagree with Yamaoka's data as reported in ref 14 and 15. It has been pointed out recently[18] that data on the ORD of PGA from several laboratories reveal significant differences. All of the possible causes of these discrepancies have not been investigated, although there is evidence[19] that the previous thermal history of the solution greatly influences the state of aggregation and the observed rotations for PGA. Until all the effects contributing to the discrepancies between the measurements of this and other laboratories have been eliminated and a set of standard conditions established, one must regard any ORD data for α-helical PGA as provisional. These considerations, at present, do not seem to indicate a need to change the reference values of the MTTDE for 100% α helix in aqueous solution since poly-γ-methoxyethyl-α-L-glutamate gives very similar values of A_{193} and A_{225} to those originally given for PGA (pH 4).[14]

(18) J. T. Yang and W. J. McCabe, *Biopolymers*, **3**, 209 (1965).

(19) T. M. Schuster, to be published.

(20) E. Shechter, J. P. Carver, and E. R. Blout, *Proc. Natl. Acad. Sci. U. S.*, **51**, 1029 (1964).

(21) N. S. Simmons, C. Cohen, A. G. Szent-Gyorgyi, D. B. Wetlaufer, and E. R. Blout, *J. Am. Chem. Soc.*, **83**, 4766 (1961).

(22) E. R. Blout, I. Schmier, and N. S. Simmons, *ibid.*, **84**, 3193 (1962).

(23) Equations 1-4 may be summarized by the general equation

$$p_j = \sum_{i=1}^{N} (\gamma_{ij} R_{\lambda_i} \bar{X}_{ij})/100 \quad (I)$$

where γ_{ij} = the coefficient of R_{λ_i} in the expansion of p_j in terms of rotational strengths.

$$\bar{X}_{ij} = \begin{cases} X_{pi} \text{ for } \alpha\text{-helical conformation Cotton effects} \\ (100 - X_{pi}) \text{ for random conformation Cotton effects} \end{cases}$$

R_{λ_i}, p_j and X_{pi} are defined in the text. Equations 5 may be summarized by

$$X_{pi} = 100(p_j - p_j^{7.0})/(p_j^{4.3} - p_j^{7.0}) \quad (II)$$

For constant $p_j^{7.0}$ and $p_j^{4.3}$ one may consider p_j as a function of R_{λ_i} and X_{pi} alone. Then one may write, at constant p_j

$$\left(\frac{dR_{\lambda_i}}{dX_{pi}}\right)_{p_j,R_{\lambda_k}} = -\left(\frac{\partial p_j}{\partial X_{pi}}\right)_{R_{\lambda_i},R_{\lambda_k}} \Big/ \left(\frac{\partial p_j}{\partial R_{\lambda_i}}\right)_{X_{p_j},R_{\lambda_k}}$$

where $k = 1, 2 \ldots, N$; $k \neq i$. From (I) and (II) above

$$\left(\frac{\partial p_j}{\partial R_{\lambda_i}}\right)_{X_{pi},p_i} = \gamma_{ij}\bar{X}_{ij}$$

and

$$\left(\frac{\partial p_j}{\partial X_{p_i}}\right)_{R_{\lambda_i},p_j} = (p_j^{4.3} - p_j^{7.0})$$

so that

$$\frac{100}{R_{\lambda_i}}\left(\frac{dR_{\lambda_i}}{dX_{pi}}\right)_{p_j^{4.3},p_j^{7.0},p_j,R_{\lambda_k}} = \frac{p_j^{4.3} - p_j^{7.0}}{\gamma_{ij}R_{\lambda_i}\bar{X}_{ij}} \equiv \beta_{ij}$$

λ_i and the more reliable are the helix content estimates from the jth parameter. Since nonconformational alterations in the rotational strengths (see assumption III) would be expected to be manifested as changes in rotational strengths at constant structural content, $|\beta_{ij}|$ constitutes a quantitative measure of the sensitivity of a given rotatory parameter to such effects. Table II

Table II. Calculated Values of $|\beta_{ij}|$ for Various Rotatory Parameters and Cotton Effects

	Cotton effects						
Parameter	192 mμ	208 mμ	224 mμ	198 mμ	217 mμ	235 mμ	B
$[R']_{233}$	1.6	2.5	0.73	3.7	8.9	190	85
$[R']_{198}$	1.6	4.4	8.2	54	51	1700	240
A_{193}	0.93	3.6	87	2.9	50	340	8.9
A_{225}	48	4.8	1.2	15	16	89	12
$A_{193} - A_{225}$	1.4	8.9	3.5	4.8	110	170	9.7
a_0 (λ_0 = 212 mμ)	0.28	0.70	0.48	0.76	4.6	6.1	5.7
b_0 (λ_0 = 212 mμ)	1.4	8.9	3.5	4.8	110	170	9.7

gives the numerical values of $|\beta_{ij}|$ for $[R']_{233}$ and $[R']_{198}$. The $|\beta_{ij}|$ are proportional to $1/X_{p_i}$ (for α-helical conformation Cotton effects) and to $1/(100 - X_{p_i})$ (for random conformation Cotton effects). The values of $|\beta_{ij}|$ given in Table II are for X_{p_i} (or $100 - X_{p_i}$) equal to unity and therefore represent the smallest values which $|\beta_{ij}|$ may adopt.

A similar matrix of values of $|\beta_{ij}|$ can be calculated for the other Cotton effect parameters, namely, position and half-width. These additional values were not calculated, because such calculations are more complicated than those for rotational strengths, and because the quantitative data regarding the extent of side chain and solvent effects is very limited at present.

From the results of the accompanying paper[10] it appears that the rotational strength most influenced by side-chain effects is that of the 208-mμ Cotton effect. Since this Cotton effect contributes significantly to $[R']_{233}$ and $[R']_{198}$ (see Table I), assumption III is not strictly valid for these rotatory parameters. However, the entries in Table II indicate that α-helix content estimates from $[R']_{233}$ and $[R']_{198}$ will not be appreciably affected (<5%) by changes in R_{208} on the order of 15–20% and, therefore, appear to be relatively insensitive to side-chain effects.

Of the remaining assumptions, assumption I (the absence of extrinsic Cotton effects) and assumption V (that only α-helical and random conformations are present) have to be validated for each compound. A discussion of the problems involved in attempting to establish the validity of these two assumptions appears in section IID. Assumption IV (negligible experimental error) is met by $[R']_{233}$ and $[R']_{192}$. The experimental errors in determining these quantities (for a synthetic polypeptide with a side chain not absorbing in the 180–600-mμ region) are about ±200 and ±5000 deg, respectively (or as a percentage of the change in rotation between helix and random conformations, 12,800 and 68,400 deg, respectively, ±1.5, ±7%).

B. Dispersion Methods. The ORD of polypeptides and proteins with a high α-helix content can be fitted to a one-term Drude equation only over a narrow spectral range. This equation, therefore, is not currently used for α-helix content estimation and will not be discussed.

1. Modified Two-Term Drude Equation. Yamaoka[15] found that a two-term Drude equation with λ_1 = 193 mμ and λ_2 = 226 mμ fitted the rotatory dispersion data of two α-helical synthetic polypeptides in several solvents over the wavelength range 275–700 mμ. The modified two-term Drude equation (MTTDE) was derived[14] in an attempt to relate the coefficients of the two Drude terms to rotational strengths. Recent circular dichroism measurements[24,25] and the results of the curve-fitting analysis of the ORD[10] indicate the presence of Cotton effects for the α-helix and random conformations which had not been demonstrated at the time of the original derivation of the MTTDE.

In order to determine the contribution of the peptide Cotton effects to A_{193} and A_{225}[25b] the rotations of *each* Cotton effect over the wavelength range 600–290 mμ were calculated using the same PGA parameters used for Table I. The values of A_{193} and A_{225} giving the best least-squares fit to the MTTDE were determined for each Cotton effect and are given in Table III. In all

Table III

Parameter[a]	192 mμ[b]	208 mμ[b]	224 mμ[b]	B[c]	Sum[d]	Obsd[e]
$A_{193}^{4.3}$	+3730	−960	−40	...	+2730	+2730
$A_{225}^{4.3}$	−37	−374	−1550	...	−1960	−1960
	198 mμ	217 mμ	235 mμ			
$A_{193}^{7.0}$	−1200	+69	+10.3	+388	−733	−737
$A_{225}^{7.0}$	−116	+113	−20.0	−152	−175	−174

[a] The superscripts refer to the pH at which the PGA data were obtained. [b] Calculated contribution to the parameter fom Cotton effects at wavelength indicated. [c] Calculated contribution from background. [d] Sum (rounded off to three significant figures) of contributions to the parameter from individual Cotton effects. [e] The parameter values obtained when the sum of the calculated rotations from all three Cotton effects is used.

cases the fit was good to 5%, or less, over the wavelength range 600–290 mμ. It is evident that for PGA at pH 4.3 A_{193} reflects the 192-mμ Cotton effect contribution and A_{225}, the 224-mμ Cotton effect contribution; but both have significant contributions from the 208-mμ Cotton effect. From the values of the calculated contributions to the rotatory parameters in Table III and the rotational strengths of the contributing Cotton effects obtained from the computer solutions one may write

$$A_{193}^{4.3} = 0.917R_{192} + 0.687R_{208} + 0.023R_{224} \quad (6)$$

$$A_{225}^{4.3} = -0.009R_{192} + 0.267R_{208} + 0.888R_{224} \quad (7)$$

$$A_{193}^{7.0} = 0.851R_{198} + 0.357R_{217} - 0.822R_{235} + 20.3B \quad (8)$$

$$A_{225}^{7.0} = 0.082R_{198} + 0.589R_{217} + 1.60R_{235} - 7.94B \quad (9)$$

Then, providing assumptions I–V are valid, α-helix content may be estimated by linear interpolation, *i.e.*

(24) M. Legrand and R. Viennet, *Compt. Rend.*, **259**, 4277 (1964).
(25) (a) G. Holzwarth and P. Doty, *J. Am. Chem. Soc.*, **87**, 218 (1965); (b) the rotatory parameters of the MTTDE are referred to below as A_{193} and A_{225} instead of the more cumbersome form originally adopted: $A_{(\alpha,\rho)(193)}$ and $A_{(\alpha,\rho)225}$, respectively.

$$X_{A_{193}} = 100\,\frac{(A_{193} - A_{193}^{7.0})}{(A_{193}^{4.3} - A_{193}^{7.0})}$$

$$X_{A_{225}} = \frac{100\,(A_{225} - A_{225}^{7.0})}{(A_{225}^{4.3} - A_{225}^{7.0})} \qquad (10)$$

From eq 6–10, $|\beta_{ij}|$ may be computed (Table II). As noted earlier, it appears that the major side chain effect is in R_{208}. As long as differences in R_{208} are less than 20%, the effect on helix content estimates from eq 10 should be less than 5%.

Another parameter, based on dispersion methods, which has been suggested[8] as a measure of α-helix content is the difference between the MTTDE coefficients, $(A_{193} - A_{225})$. The advantage of this rotatory parameter is its insensitivity to solvent dielectric constant; $(A_{193} - A_{225})$ may also be resolved into separate contributions from the peptide bond Cotton effects (Table IV). From the values of the calculated contributions to the rotatory parameters in Table IV and

Table IV

Parameter	192 mμ[a]	208 mμ[a]	224 mμ[a]	B[b]	Sum[c]
$A_{193}^{4.3} - A_{225}^{4.3}$	+3770	−589	+1510	...	+4690
	198 mμ	217 mμ	235 mμ		
$A_{193}^{7.0} - A_{225}^{7.0}$	−1090	−44	+30.3	+540	−562

[a] Calculated contribution to the parameter from Cotton effects at wavelength indicated. [b] Calculated contribution from background. [c] Sum (rounded off to three significant figures) of calculated contributions to the parameter from individual Cotton effects.

the rotational strengths of the contributing Cotton effects obtained from the computer solutions one may write

$$A_{193}^{4.3} - A_{225}^{4.3} = 0.926R_{192} + 0.420R_{208} - 0.865R_{224} \qquad (11)$$

$$A_{193}^{7.0} - A_{225}^{7.0} = 0.769R_{198} - 0.232R_{217} - 2.42R_{235} + 28.2B \qquad (12)$$

and if assumptions I–V are valid

$$X_{(A_{193} - A_{225})} = 100\,\frac{[(A_{193} - A_{225}) - (A_{193}^{7.0} - A_{225}^{7.0})]}{[(A_{193}^{4.3} - A_{225}^{4.3}) - (A_{193}^{7.0} - A_{225}^{7.0})]} \qquad (13)$$

The corresponding sensitivity coefficients $|\beta_{ij}|$ are listed in Table II. Helix content estimates from $(A_{193} - A_{225})$ are even less sensitive to differences in R_{208} than are estimates from A_{193} or A_{225}. In addition, $X_{(A_{193} - A_{225})}$ is less sensitive to differences in R_{192} than $X_{A_{193}}$ and less sensitive to differences in R_{224} than $X_{A_{225}}$. Thus, even though assumption III is not strictly valid for the MTTDE parameters, it appears that side chain effects have little influence on helix content estimates from A_{193}, A_{225}, and particularly from $(A_{193} - A_{225})$.

2. Moffitt Equation. The semiempirical derivation[13] of the Moffitt equation (ME) is based on the assumption that the dominant contributions to the visible and near-ultraviolet ORD of α-helical polypeptides and proteins are from transitions of the peptide bond around 185 and 148 mμ. In view of experimental and theoretical results[10,21,24,25a,26] obtained since the original formulation of the Moffitt equation, this assumption is invalid. Thus the rotatory parameters of the ME, a_0 and b_0, cannot be related to the rotational strengths of the contributing Cotton effects *via* the equations of the original derivation. However, the parameters of the ME and the MTTDE are related by the following equations

$$a_0\lambda_0^2 = A_{193}\lambda_{193}^2 + A_{225}\lambda_{225}^2 \qquad (14)$$

$$a_0\lambda_0^4 + b_0\lambda_0^4 = A_{193}\lambda_{193}^4 + A_{225}\lambda_{225}^4 \qquad (15)$$

provided that for the desired wavelength range

$$[A_{193}\lambda_{193}^2(\lambda_{193}^2 - \lambda_0^2) + A_{225}\lambda_{225}^2(\lambda_{225}^2 - \lambda_0^2)] \times (\lambda^2 - \lambda_0^2)^{-3} + \text{subsequent terms} = 0 \qquad (16)$$

The derivation of these relations, which is given in Appendix I, does not involve the Moffitt assumptions (see relations Da–c of Appendix I and accompanying discussion). The failure to use this more general derivation caused three erroneous conclusions in a previous paper.[20] These were: (1) the statement on p 1033 that, "for values of λ_0 other than 209 mμ, the linear relation of b_0 to helix content has to be checked experimentally;" (2) the statement on p 1035 that "the modified two-term Drude equation... allows more precise determinations of α-helix content by extending the ranges of measurements to shorter wavelengths" (the implicit comparison being with the ME); and (3) eq 14 of that paper. The first statement is obviously in error since from eq 15 it is evident that b_0 is linearly related to A_{193} and A_{225} and hence to helix content, independent of the particular choice of λ_0. Once this point is established, the second statement is obviously incorrect since λ_0 values can be chosen which will give a straight-line fit in a Moffitt plot over as wide a spectral range as for an MTTDE plot. Finally, the coefficients on the right-hand side of eq 14 of ref 20 are incorrect (see Appendix I). The correct equations are, for $\lambda_0 = 209$ mμ

$$a_0 = 0.853A_{193} + 1.16A_{225}$$
$$b_0 = -0.126A_{193} + 0.184A_{225} \qquad (17)$$

For $\lambda_0 = 212$ mμ

$$a_0 = 0.829A_{193} + 1.13A_{225}$$
$$b_0 = -0.142A_{193} + 0.142A_{225} \qquad (18)$$

For $A_{193} = +2900$ deg cm^3 dmole^{-1} and $A_{225} = -2050$ deg cm^3 dmole^{-1} the difference between the rotations predicted by the MTTDE and those predicted by the ME, using a_0 and b_0 values calculated either from eq 17 or from eq 18, is less than 3% of the total calculated rotation for 250 mμ $\leq \lambda \leq$ 540 mμ ($\lambda_0 = 212$ mμ) and 310 mμ $\leq \lambda \leq$ 600 mμ ($\lambda_0 = 209$ mμ). Evidently, as Urnes and Doty[6] have pointed out, values of λ_0 other than 212 mμ may be used in the ME provided that the b_0 values for 100% α helix and 0% α helix are determined for the particular λ_0 value used. It should be further noted that the wavelength range over which the ME is valid changes when different λ_0 values are used.

From Section IIB1 we know the relationship of A_{193} and A_{225} to the rotational strengths of the Cotton effects of the random and α-helical conformations (eq 6–9). Substituting in eq 18 we obtain

$$a_0^{4.3} = 0.750R_{192} + 0.871R_{208} + 1.02R_{224} \qquad (19)$$

(26) J. A. Schellman and P. Oriel, *J. Chem. Phys.*, 37, 2114 (1962).

$$b_0^{4.3} = 0.131R_{192} - 0.060R_{208} + 0.123R_{224} \quad (20)$$

$$a_0^{7.0} = 0.798R_{198} + 0.962R_{217} + 1.13R_{235} + 7.86B \quad (21)$$

$$b_0^{7.0} = -0.109R_{198} + 0.033R_{217} + 0.344R_{235} - 4.00B \quad (22)$$

Before one may interpret a_0 and b_0 in terms of helix content, assumptions I–V must be met. One may then write

$$X_{a_0} = 100\frac{a_0 - a_0^{7.0}}{a_0^{4.3} - a_0^{7.0}}; \quad X_{b_0} = 100\frac{b_0 - b_0^{7.0}}{b_0^{4.3} - b_0^{7.0}} \quad (23)$$

From these equations β_{ij} values were computed for a_0 and b_0 and are shown in Table II. As stated above, side chain effects seem to have their major influence on R_{208} (for full helix). From Table II it would appear that such effects would have little influence on b_0 but a large influence on a_0. Thus, only b_0 meets assumption III and gives helix content estimates comparable to (A_{193} − A_{225}) in insensitivity to changes in R_{208}.

Since A_{193} is linearly related to A_{225}[8,14] and a_0 and b_0 are linearly related to A_{193} and A_{225} (eq 18), a_0 and b_0 must be linearly related to each other. In fact, taking the "aqueous solvent line" (eq 9 of ref 14) and substituting in eq 17 and 18 (λ_0 = 212 mμ), one obtains upon simplification

$$b_0 = -1.05a_0 - 571 \quad (24)$$

Alternatively, using the "organic solvent line" (eq 3 of ref 8), one may write

$$b_0 = -1.05a_0 - 372 \quad (25)$$

Then why cannot deviation from eq 24 and 25 (*i.e.*, different helix content estimates from a_0 and b_0) be used as a criterion for the presence of other structures than α-helical and random conformations in the same manner as the "aqueous" and "organic" lines of the MTTDE are used? The answer is that they can, but not in precisely the same way. Prior to the development of the MTTDE many attempts had been made to devise a method for extracting useful information from the a_0 values of proteins, but the large discrepancy between helix content estimates from a_0 and b_0 (sometimes nearly 50%) has led to a general rejection of a_0 as a useful parameter for proteins. It is now possible, on the basis of the MTTDE, to show, in principle, how a_0 may be used so that the ME becomes equivalent to the MTTDE with respect to information obtained.

Applying the transformation between (A_{193}, A_{225}) values and (a_0, b_0) values for λ_0 = 212 mμ (eq 18) to the reference points of the MTTDE one obtains
(1) for high dielectric constant solvents ($D > 30$)

	A_{193}	A_{225}	a_0	b_0
helix	+2900	−2050	+90	−700
random	−750	−60	−690	+100

(2) for low dielectric constant solvents ($D < 30$)

	A_{193}	A_{225}	a_0	b_0
helix	+3020	−1900	+360	−700
random	−600	0	−500	+90

It can be shown that for a pair of (A_{193}, A_{225}) values just off the "aqueous solvent line" so that the difference in helix content estimates from A_{193} and A_{225} is y%, the corresponding difference in helix content estimates from a_0 and b_0 will be 3.2y%. Thus, one finds that for a difference in helix content estimates from A_{193} and A_{225} of 5% one obtains a difference in helix content estimates from a_0 and b_0 of 16%.

For proteins, $X_{A_{193}}$ and $X_{A_{225}}$ are generally calculated using both the high and the low dielectric constant solvent reference values. It is then assumed that the solvent dielectric constant associated with the reference values giving the best agreement between $X_{A_{193}}$ and $X_{A_{225}}$ represents the average effective dielectric constant for the peptide bond environment in the molecule. If agreement cannot be obtained to better than about 5% for either set of reference values it is concluded that the molecule contains additional structures.

In principle, the same procedure may be applied to a_0 and b_0. The only difference is that if the best agreement is not within 15%, it may be concluded that the molecule contains additional structures. Under circumstances where the difference between helix content estimates from a_0 and b_0 is less than 15%, it may be concluded that the molecule probably contains only random and α-helical conformations and that the amount of helix is correctly estimated from b_0. Thus a_0, because of its extreme sensitivity to solvent and side chain effects (see Table II), cannot yield a quantitatively meaningful estimate of helix content. However, if the appropriate end points are used (those computed from MTTDE end points), a_0 and b_0 can give a qualitative indication of the presence of other conformations similar to that given by the two MTTDE parameters.

This method of using a_0 and b_0 to determine whether additional structures are present follows as a necessity from the mathematical equivalence of the ME and MTTDE. Such an approach to the interpretation of the parameters of the ME presents no advantages whatsoever over the MTTDE, and is presented solely to illustrate the equivalence of the two equations—when used in the manner outlined above. It should be noted that only with the insight gained from the MTTDE has it been possible to determine how one may usefully interpret a_0. As will be seen below, there is a possibility of errors arising from applying the 5% MTTDE criterion to the values of $X_{A_{193}}$ and $X_{A_{225}}$ of globular proteins. Such errors may also arise in using the 15% limit for X_{a_0} and X_{b_0} estimates.

To summarize this discussion of the ME: (1) we conclude that a_0 and b_0 of the ME can be interpreted with the aid of eq 18 provided that assumptions I–V are met; (2) we point out that assumption III is valid for b_0 but not a_0 (at least with respect to side-chain effects); (3) we correct three conclusions in ref 20 with respect to the comparison between the ME and the MTTDE; (4) we demonstrate that it is possible to formulate a criterion, based on a_0 and b_0 values, for the presence of structures other than α-helical or random conformations.

3. Comparison of the Dispersion Methods. Both dispersion methods discussed above (ME and MTTDE) start with the assumption that the rotation in the visible and near ultraviolet can be adequately represented by a multiterm Drude equation—one term for each contributing Cotton effect. Now that the Cotton effect parameters are known for several synthetic polypeptides in the α-helical and random conformations, it is possible to check the validity of this assumption. In Appendix II it is shown that for the PGA (pH 4.3)

data errors of approximately 5% arise at 280 mμ as a result of this approximation. Thus, even if it were possible to obtain statistically significant coefficients from a fit of the observed ORD to a multiterm Drude equation, these coefficients would not yield precise rotational strengths. However, for two-term equations this type of error is not significant since it can be compensated for by the error terms arising in the subsequent approximations.

In the general procedure for deriving the various two-term Drude equations each term in the multiterm Drude equation is expanded as a Taylor series in $(\lambda^2 - \lambda_i^2)^{-1}$ about some appropriate origin. The choice of origin and assumptions regarding the relative magnitudes of the coefficients of the various terms are guided by the parameter values for the assumed Cotton effects.

There is a distinction to be made between the original derivations of the ME and MTTDE and the phenomenological interpretations of these equations. As is quite evident from the preceding discussion, neither derivation correctly predicts the dependence of the rotatory parameters upon the rotational strengths of the existing Cotton effects. This failure stems partly from the methods of approximation, and partly from incorrect starting assumptions regarding the contributing Cotton effects. Now that the precise parameters for the Cotton effects of a few synthetic polypeptides are known,[10] it should be possible to "derive" the correct relationships. However, because of the additional contributing Cotton effects the approximation process from many Cotton effect functions to an MTTDE or an ME is considerably more complicated and involves the determination of the conditions which minimize a sum of involved error terms. Such a derivation may be performed, but since a much more direct method is available, it seems unnecessary to attempt it. The direct method is the one used in the foregoing discussion to express the rotatory parameters in terms of rotational strengths. It has the advantage of being readily applicable to any new two-term equation which one may wish to consider.

It has been demonstrated[6] that the ME can be made to fit the ORD of α-helical synthetic polypeptides over various wavelength ranges simply by using different values of λ_0. Thus, as λ_0 is varied, one generates a family of ME's valid over different wavelength ranges. By contrast, two-term Drude equations add another degree of freedom by providing a fourth parameter to be adjusted. The addition of a fourth parameter allows a family of two-term Drude equations with different pairs of λ_1 and λ_2 values[27] to fit the ORD data over *the same wavelength range*. The significant advantage gained by this approach is that the rotatory parameters of each two-term equation in the family will have a different linear dependence on the rotational strengths of the contributing Cotton effects and therefore a different set of $|\beta_{ij}|$ values. It should be possible by extensive model studies to establish which Cotton effect parameters are most influenced by nonconformational effects. Then, a particular two-term equation can be chosen, for which the helix content estimates are least sensitive to these observed nonconformational effects. Thus, once such criteria for choice are defined by model studies, the addition of the fourth parameter allows one to select the particular two-term Drude equation meeting these criteria over the wavelength range under study. However, once this particular equation has been selected (either a two-term or, if it meets the criteria, the ME for that wavelength range), then the λ_1 and λ_2 (or λ_0) values are fixed and are *no longer parameters*. Thus, for the purpose of determining helix contents, there are *only two parameters* to be evaluated no matter which two-term equation is used.

It becomes apparent that the ME and MTTDE are two out of many possible equivalent two-term equations. Since the Cotton effect parameters are now known for several synthetic polypeptides, it should be possible, by the methods of Sections IIB1 and IIB2, to relate the parameters of any of these different two-term representations of the visible and near-ultraviolet ORD to the parameters of the actual Cotton effects. The derived parameters of these equations will all be more or less sensitive to the various sources of error in helix content estimation, and it is on the basis of such considerations that an evaluation of the different representations should rest.

C. Curve-Fitting Methods. The nonlinear, least-squares approach has been successfully applied[10] to the ORD of several synthetic homopolypeptides and would appear to be the most promising approach to the determination of quantitative limits for the magnitude of solvent and side-chain effects. As yet, this particular approach has not been successful in resolving the ORD of mixtures of two structures into their component Cotton effects.

D. The Remaining Assumptions. Much of the discussion so far in this paper has concerned the extent to which the interpretation of the currently used rotatory parameters is influenced by side-chain effects. We concluded that b_0 and $(A_{193} - A_{22\bar{5}})$ are the least sensitive to such effects and, therefore, yield the best measure of helix content for systems where side-chain effects are important. In the remainder of this section, we discuss first solvent effects (assumption III), then the backbone peptide bond assumption (assumption I), and finally the helix–random conformation limitation (assumption V).

1. Assumption III. Cassim and Taylor[7] have shown that b_0 values for poly-γ-benzyl-α-L-glutamate show a linear dependence on the index of refraction of the solvent. They suggest that the explanation lies in solvent-induced frequency shifts in the peptide bond transitions and not in conformational changes. Since the curve-fitting method of the accompanying paper[10] yields estimates of the positions (frequencies) of the contributing Cotton effects, it provides a way of testing this hypothesis. Unfortunately, almost all the solvents used by Cassim and Taylor are opaque in the far ultra-

(27) In the original derivation[14] of the MTTDE, because a Cotton effect was known to be centered near 193 mμ, one term was fixed at 193 mμ and the other varied to give a straight line over the longest wavelength range, this condition being obtained with λ_2 = 225 mμ. It was thought that under these conditions A_{193} would include contributions only from the 193- and the 198-mμ Cotton effects of the α helix and random conformations, respectively, and thus yield a parameter linearly related to conformational rotational strengths. However, as was pointed out earlier, such a parameter will be obtained no matter what value of λ_1 is used provided that, in conjunction with an appropriate value of λ_2, the resulting two-term Drude equation fits the observed ORD over a reasonably large wavelength range. Thus, many other pairs of values for λ_1 and λ_2 could have been used.[15,16] However, if other pairs are used a different set of relations analogous to eq 6–9 will have to be derived in order to relate the rotatory parameters to the rotational strengths of the assumed Cotton effects.

violet. However, several polypeptide–solvent systems are being investigated with a view to testing their hypothesis.

2. Assumptions I and V. A first step in attempting to establish the validity of assumption I is to determine whether or not there are residues in the molecule which may make "abnormal" contributions to the optical activity, *e.g.*, prolyl, cystinyl, or any of the aromatic side chain residues. It may be possible to establish that aromatic side chains are contributing by examining the circular dichroism or ORD of concentrated solutions of the protein in the 250- to 300-mμ region. However, failure to observe optically active bands in this region does not ensure that side chain chromophores are not contributing in the region of peptide bond absorption. It is possible to determine whether a large proportion of the backbone peptide bonds are in conformations other than α helical or random (failure of assumption V) by examination of the far-ultraviolet ORD or, if this is not accessible, from an analysis of the infrared spectrum. However, it is generally not possible to determine the presence of small amounts (on the order of 10 or 20%) of structures other than the α-helical or the random conformation by such techniques.

It would appear that a sufficient criterion for the validity of assumptions I and V is the ability to construct the observed ORD for the protein from the ORD curves of model α-helical and random conformation polypeptides. However, because of possible nonconformational effects a failure to meet such a condition cannot be taken to mean that assumptions I and V are not valid.

It should also be emphasized that agreement between $X_{A_{193}}$ and $X_{A_{225}}$ to within 5%, X_{a_0} and X_{b_0} to within 15%, may be a *necessary* but is not a *sufficient* condition for the presence of only α-helical and random conformations. The 5% limit was based on the observed scatter in an A_{192} *vs.* A_{225} plot for various synthetic polypeptides and copolypeptides in solvents of similar dielectric constant.[8,13] This limit, therefore, includes side chain effects and possibly effects of other solvent parameters (index of refraction, dipole moment) on A_{193} and A_{225} as well as experimental error. However, because of the greater variety of nonconformational alterations of peptide rotational strengths present in globular proteins, as compared with homopolypeptides, the values of $X_{A_{193}}$ and $X_{A_{225}}$ for proteins can conceivably differ by more than 5% even when there are only α-helical and random conformations. Conversely, the values of $X_{A_{193}}$ and $X_{A_{225}}$ for proteins may differ by less than 5% even in the presence of small amounts of structures other than the α-helical or random conformations or of small extrinsic Cotton effects. Thus, the application of the 5% criterion to the $X_{A_{193}}$ and $X_{A_{225}}$ of proteins could be misleading. An example is elastin. The visible and near-ultraviolet ORD of this protein in 0.01 *N* HCl gives $X_{A_{193}}$, $X_{A_{225}}$, and $X_{(A_{193} - A_{225})}$ values of 13, 16, and 15%, respectively,[28] from the low dielectric constant reference values. Applying the reasoning used for apomyoglobin[29] one concludes that elastin contains 15% α helix and that the regions of helix are buried within the molecule. Yet the ORD curve in the far ultraviolet differs significantly from that for a mixture of 15% α-helical and 85% random conformations. Such a deviation might be expected since 12% of the residues of this molecule are either prolyl or hydroxyprolyl and 33% are glycyl.[30] Evidently, a difference between $X_{A_{193}}$ and $X_{A_{225}}$ of less than 5% is not a sufficient condition for the presence of only random and α-helical conformations.

Thus, for proteins, there is at present no adequate *general* method either for determining whether chromophores other than peptide bonds are contributing to the observed ORD (assumption I) or for establishing that only α-helical and random conformations are present (assumption V). For synthetic polypeptides, however, on the basis of the data available.[8,14] a difference between $X_{A_{193}}$ and $X_{A_{225}}$ of more than 5% may be considered to indicate the presence of structures other than the α-helical or random conformations.

3. α-Helix Content Estimation. There remain, therefore, at least four potential sources of error in α-helix content determination using the best current methods, b_0 or $(A_{193} - A_{225})$, for proteins. They are: (a) the possible contributions to the ORD from optically active chromophores other than the backbone peptide bonds, (b) the possible effect of the distribution of helical lengths, (c) the possible effects of the local environment on peptide bond optical activity, *i.e.*, solvent effects and fixed conformation effects, (d) the possible presence of *small* amounts of structures other than α-helical or random conformations.

Because of our current ignorance of the magnitude of these effects, one cannot with confidence place even an upper limit on the uncertainty in α-helix content estimates for globular proteins. However, it should be possible, with further model studies of the type described above and in the accompanying paper, to determine such limits. On the assumption that the four factors listed above are introducing negligible errors in helix content estimates from b_0 and $A_{193} - A_{225}$, it would appear that these parameters currently represent the most reliable estimates of helix content.

We therefore suggest the following procedure for the estimation of α-helix content of proteins. (1) Measure the ORD over as wide a wavelength range as possible, at several concentrations, and in a variety of solvents, for the purpose of determining whether there are contributions from chromophores other than the backbone peptide bonds. In addition, examine the far-ultraviolet ORD, if possible, in order to eliminate the possibility of there existing large amounts of structures other than the α-helix or the random conformation. (2) Evaluate A_{193}, A_{225}, and $(A_{193} - A_{225})$ from a modified two-term Drude plot. If a straight line is not obtained over the range 600–280 mμ then one of the sources of error considered above is significant and α-helix content estimates should be considered qualitative at best. (3) Estimate helix content using the following reference values for 100 and 0% α-helix

	A_{193}		A_{225}		$(A_{193} - A_{225})$
	High D	Low D	High D	Low D	*D* independent
100%	+2900	+3020	−2050	−1900	+4930
0%	−750	−600	−60	0	−650
range	+3650	+3620	−1990	−1900	+5580

(28) J. P. Carver and J. Gross, unpublished results.
(29) S. C. Harrison and E. R. Blout, *J. Biol. Chem.*, **240**, 299 (1965).
(30) S. M. Partridge, *Advan. Protein Chem.*, 17, 227 (1962).

(4) If agreement is obtained between $X_{A_{193}}$ and $X_{A_{225}}$ to within 5% then $X_{(A_{193} - A_{225})}$ gives the α-helix content *provided that none of the sources of error listed above are contributing significantly.*

III. Some Comments on the Interpretation of the ME and MTTDE Parameters

Recently, the usefulness of the MTTDE has been questioned.[31] It is not our intention to discuss that paper at length, but only those criticisms which are relevant to the application of the equation will be considered.

It was claimed[31] that the values of A_{193} and A_{225} are not independent because of the method of plotting, and an alternative method which should yield them independently is given. This second method had been investigated by us prior to the publication of the first paper in this series. We found that it yields the same results as the method of plotting originally suggested.[13] Therefore, A_{193} and A_{225} are obtained independently by both methods.

It was pointed out[31] that the ORD's of some structures other than the random or α-helical conformations are fitted by the MTTDE from which it was concluded that $\lambda_2 = 225$ mμ and $\lambda_1 = 193$ mμ are "not unique"—presumably to α-helix–random conformation mixtures. In our paper[8] we did not state that Drude terms at 193 and 225 represented an expression which would *only* be fitted by the ORD of mixtures of random and α-helical conformations. What was claimed[8] was that the parameter values derived from the MTTDE would not be *linearly related according to either eq 9 of ref 14 or eq 3 of ref 8* when conformations other than random and/or α-helical were present. With the reservations of the earlier discussion, this would appear still to be true.

Another criticism[31] was that the MTTDE reference values for the random conformation are not sufficiently well defined. To prove the point, ORD data for PGA (pH 7.3) at various salt concentrations are given for the wavelength range 260–190 mμ. It is assumed that the molecules are in the random conformation for all salt concentrations and that the significant differences observed in the ORD curves constitute polyelectrolyte effects. The unsubtantiated premise that PGA (pH 7.3) in 6 *M* KF is representative of the random conformation is then used as a basis for the assertion that the A_{193} and A_{225} values observed for this system (−70, −390) are as reliable reference values for 0% helix as those observed for PGA in water at pH 7.3 (−880, −100). Since A_{193} and A_{225} helix content estimates, using these two sets of parameters as representing 0% helix, are self-consistent but different for the two solvent conditions, it is concluded that agreement between A_{193} and A_{225} helix content estimates is no guarantee of the accuracy of the estimates. However, the basic premise seems false—the Cotton effect curves shown and the corresponding A_{193}, A_{225}, and b_0 values are entirely consistent with the formation of 10–20% helix on going from 1 to 6 *M* KF. On the assumption of partial helix formation in 6 *M* KF, the origin of the consistency of the $X_{A_{193}}$ and $X_{A_{225}}$ estimates is self-evident. The foregoing may or may not be the correct interpretation of the data. The fact remains, however, that both b_0 and $(A_{193} - A_{225})$ give equivalent helix content estimates of about 15% for PGA, in 6 *M* KF at pH 7.3. Thus any conclusions regarding the relevance of such "solvent" effects with respect to the validity of helix content estimates from the MTTDE must hold equally well for the ME.

Now let us consider a viewpoint regarding the interpretation of a_0 values of the ME. It has been held[31-33] that part of the a_0 values obtained for proteins and synthetic polypeptides is a constant contribution from the asymmetric α carbons, which can be represented as $(\Sigma a_0^R)\lambda_0^2/(\lambda^2 - \lambda_0^2)$ (where a_0^R is the part of a_0 depending "on the intrinsic residue rotation (asymmetric carbon)"[32] and the summation is taken "over the a_0^R values characteristic of each residue in the protein"[32]). There is little doubt that the asymmetry of the α carbon contributes to the interaction potentials of the peptide groups in both α-helical and random regions and in this sense each *rotational strength* may be considered to be partly arising from the "influence" of the asymmetric α carbons. However, as Kauzmann pointed out in 1957,[34] there are two other sources for the optical activity of polypeptides besides the contributions of the asymmetric α carbons (inherent optical activity of the side chains) and the contribution of the peptide backbone; namely, (1) the interactions of side chains with the backbone peptide bonds and (2) the interactions between side chains. Both of these sources of interaction are considerably altered upon going from an α-helical to a random conformation. It is therefore incorrect to attempt to resolve the rotational strength of the peptide transitions into two components one of which is assumed to remain constant upon transition from α-helical to random conformation and the other, although changing, is assumed to adopt the same values for *all* α-helical–random polypeptides and solvent systems at a given helix content. Such a resolution was originally suggested by Doty[32] and is still being proposed by Yang[31,33] as a basis for his modified Moffitt equation.

A major source of much of the present dilemma regarding the proper methods for the estimation of the α-helix content of polypeptides and proteins is the criterion for 100% of a structure. Unfortunately, neither of the hydrodynamic methods utilized to date[7,35,36] (flow birefringence and viscosity) is capable of distinguishing 90 from 100% helix content. The most common resort[7,35,36] is to an internal "criterion"—that is, the attainment of a plateau in a plot of the helix content parameter *vs.* a parameter of the transition causing perturbation. However, the only valid conclusion which may be derived under such circumstances is that the maximum or minimum amount of structure *for that system* has been reached.

Summary

In the following we summarize the conclusions of the foregoing discussion.

(31) J. T. Yang, *Proc. Natl. Acad. Sci. U. S.*, **53**, 438 (1965).

(32) P. Doty, Proceedings of the 4th International Congress of Biochemistry, Vol. VIII, Vienna, 1958, p 8.

(33) J. T. Yang in "Polyamino Acids, Polypeptides and Proteins," M. A. Stahmann, Ed., University of Wisconsin Press, Madison, Wis., 1962, p 225.

(34) W. Kauzmann, *Ann. Rev. Phys. Chem.*, **8**, 413 (1957).

(35) E. Iizuka and J. T. Yang, *Biochemistry*, **4**, 1249 (1965).

(36) J. Y. Cassim and E. W. Taylor, *Biophys. J.*, **5**, 573 (1965).

(1) Several assumptions regarding the nature of the contributions to the observed ORD are inherent in all current methods of utilizing the ORD for the estimation of α-helix content of polypeptides and proteins. The general validity of these assumptions for real structures is discussed. It is concluded that although many of these assumptions may not be strictly valid, in some cases a suitable choice of rotatory parameters allows correct α-helix content estimates to be made.

(2) The mathematical equivalence of the ME and MTTDE has been demonstrated; thus, their rotatory parameters, in principle, are equally useful with respect to the estimation of α-helix content and the detection of the presence of structures other than the α-helix or the random conformation.

(3) The previously demonstrated[8,14] utility of the MTTDE parameters for the detection of the presence of structures in synthetic polypeptides other than the α-helix and random conformations has been extended to the ME.

(4) It is concluded that it should be possible to obtain many more equivalent two-term representations of the visible and near-ultraviolet ORD of α-helical and random conformation mixtures.

(5) The curve-fitting method described in the accompanying paper yields a preliminary indication of the nature of "side chain effects" and should prove useful in the further investigation of the extent of nonconformational alterations in the peptide bond rotational strengths. Thus, when more is known about the quantitative effects of nonconformational alterations it may be possible to select an optimal two-term equation along the lines used above to evaluate the present methods.

(6) Of the present methods of estimating α-helix content it is shown that $[R']_{233}$, $[R']_{198}$, b_0, and $(A_{193} - A_{225})$ are all relatively insensitive to variations in R_{208}. Because of the probable sensitivity of $[R']_{233}$ and $[R']_{198}$ to changes in the half-widths of the Cotton effects, it would appear that b_0 and $(A_{193} - A_{225})$ best meet the basic assumptions for helix content estimation discussed in this paper.

(7) Finally, it is pointed out that a major obstacle to the improved definition of reference values for α-helix content parameters is the lack of adequate criteria for the presence of either 100% α-helical or 100% random conformation.

Acknowledgment. We wish to thank Drs. P. Urnes and P. Doty for useful discussions during the course of this work. We are also indebted to Mr. J. R. Parrish, Jr., for helpful criticism.

Appendix I

The derivation of the precise relationship between the ME and MTTDE rotatory parameters is a direct application of the approach used both by Schellman[12] and by Moffitt[13] in the derivations of the "one-term Drude" equation and the Moffitt equation, respectively.

Consider the MTTDE in simplified notation

$$[R']_\lambda = \frac{A_{193}\lambda_{193}^2}{\lambda^2 - \lambda_{193}^2} + \frac{A_{225}\lambda_{225}^2}{\lambda^2 - \lambda_{225}^2} \qquad (A)$$

then expand the first term as a Taylor series in $(\lambda^2 - \lambda_{193}^2)^{-1}$ (considered as a function of λ_{193}^2) and the second term as a Taylor series in $(\lambda^2 - \lambda_{225}^2)^{-1}$ (considered as a function of λ_{225}^2) and choose a common origin λ_0^2 about which to expand. One obtains

$$[R']_\lambda = \frac{a_0\lambda_0^2}{\lambda^2 - \lambda_0^2} + \frac{b_0\lambda_0^4}{(\lambda^2 - \lambda_0^2)^2} \qquad (B)$$

provided that

$$\text{(a)}\quad a_0\lambda_0^2 = A_{193}\lambda_{193}^2 + A_{225}\lambda_{225}^2 \qquad (Ca)$$

$$\text{(b)}\quad a_0\lambda_0^4 + b_0\lambda_0^4 = A_{193}\lambda_{193}^4 + A_{225}\lambda_{225}^4 \qquad (Cb)$$

$$\text{(c)}\quad 0 = [A_{193}\lambda_{193}^2(\lambda_{193}^2 - \lambda_0^2) + A_{225}\lambda_{225}^2(\lambda_{225}^2 - \lambda_0^2)](\lambda^2 - \lambda_0^2)^{-3} + \text{subsequent terms} \qquad (Cc)$$

If it is further assumed that

$$\text{(a)}\quad \lambda_0 = (\lambda_{193} + \lambda_{225})/2 \qquad (Da)$$

$$\text{(b)}\quad |A_{193} + A_{225}| \ll |A_{193}| \qquad (Db)$$

$$\text{(c)}\quad |\lambda_{225} - \lambda_{193}| \ll \lambda_0 \qquad (Dc)$$

then one can simplify the equations in (C) to give

$$a_0 \simeq (A_{193} + A_{225}) + b_0$$
$$b_0 \simeq (A_{193} - A_{225})(\lambda_{193} - \lambda_{225})/\lambda_0 \qquad (E)$$

This type of simplification was introduced by Moffitt[13] on the assumption that the original two terms in equations of the form (A) arose as oppositely signed components of an exciton band and, for such a system, are perfectly valid. In a recent paper,[20] we used eq E to relate the rotatory parameters of the ME and the MTTDE, taking $\lambda_0 = 209$ mμ, the value given by eq Da. Subsequently, it was found that the relationships obtained from eq C and from eq E for α-helical synthetic polypeptide parameter values differed significantly. On examination of assumptions D it was evident that assumption Db is not valid for such a system. The correct relationships are given, therefore, by the more general eq C—see eq 14–18 of the text. The only restriction on the value of λ_0 is that it minimize the higher order terms in the expansion (eq Cc).

Appendix II

As was shown in the accompanying paper,[1] the error in representing a Moscowitz term by a Drude term is given to a first approximation by $A_i\Delta_i^2\lambda_i/4(\lambda^2 - \lambda_i^2)^3$ which falls off quite rapidly for an isolated Cotton effect, becoming less than 2% for $(\lambda - \lambda_i)/\Delta_i$ greater than 5.3. However, with some systems, the α helix in particular, the error in representing the ORD as the sum of the corresponding Drude terms is much greater. For α-helical polypeptides the rotations observed in the visible and near ultraviolet are small compared to the contribution of each term due to the fact that the contributions partially cancel. However, as indicated below, the errors do not cancel to the same extent. As a consequence, the per cent error becomes very large. At 280 mμ, for example (5.3 half-widths from the 224-mμ band), the contribution to the rotation of the 192-, 208-, and 224-mμ Cotton effects are approximately +3300, −1520, and −2810°, respectively. The errors introduced by the three terms are 0.6, 1.0, and 2.1% of their respective contributions, that is, +19.6, −14.7, and −58.0°, or a total error of −53°. This error is associated with a net rotation

of $-1040°$ and therefore corresponds to a 5.1% error. At longer wavelengths the error is reduced. At 350 mμ the errors arising from the 192-, 208-, and 224-mμ Cotton effects are +3.4 (0.2%), −2.0 (0.3%), and −5.1° (0.5%), respectively. The actual rotation at 350 mμ is −184°, so that the error has fallen to 2% at this wavelength. At wavelengths longer than 350 mμ, the error varies slowly reaching a minimum of 1.3% near 500 mμ and increasing to 1.7% by 600 mμ. Thus a multiterm Drude equation (as proposed by Yang[31]) will not yield true rotational strengths. However, the general derivation of a two-term dispersion equation to approximate the visible and near-ultraviolet ORD introduces additional error terms. A particular approximation such as the MTTDE can usually be chosen so that the additional errors introduced minimize the total error and thus the approximation can be used over the range 600–280 mμ despite the inadequacies of the multiterm Drude approximation.

Studies of the Chymotrypsinogen Family. V. The Effect of Small-Molecule Contaminants on the Kinetic Behavior of α-Chymotrypsin[1]

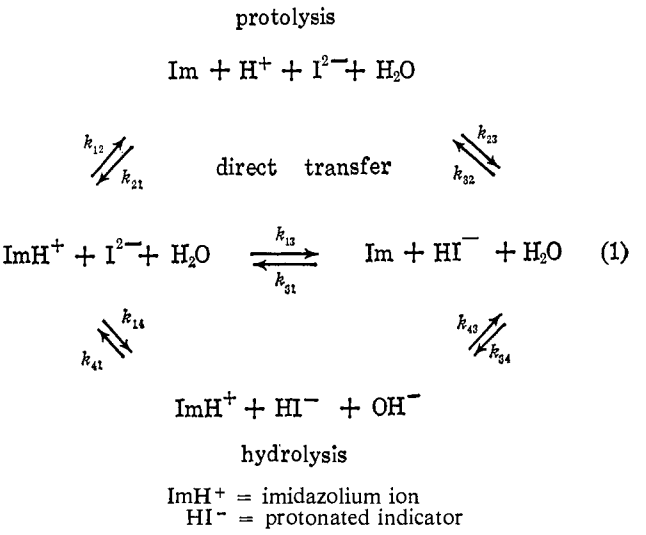
Anthony Yapel, Moon Han, Rufus Lumry, Andreas Rosenberg, and Da Fong Shiao

Contribution from the Laboratory for Biophysical Chemistry, Chemistry Department, University of Minnesota, Minneapolis, Minnesota 55455. *Received October 27, 1965*

Abstract: All commercial preparations of α-chymotrypsin investigated contained contaminants of one or two types. The first type apparently consists of autolysis products and interferes with quantitative study of the kinetics of chymotryptic catalysis by high-speed methods. The second contaminant, of unknown nature, is tightly bound in such a way as to prevent substrate binding or the binding of competitive inhibitors. Its effect is formally noncompetitive and thus reduces the concentration of participating enzyme. Reliable measurement of the type and extent of contamination is at present possible only with temperature-jump methods but a simple, completely satisfactory purification procedure has been developed. Steady-state kinetics study provides unreliable tests for contamination because of the slow, time-dependent dissociation of the second type of contaminant at low-protein concentrations. The N-*trans*-cinnamoylimidazole test is also unreliable since the contaminants do not block the reaction. Quantitative data previously reported for α-chymotrypsin (CT) as the result of steady-state kinetics investigation must be considered unreliable until verified using pure protein. The dissociation equilibrium constant for indole binding to CT is found to be 2.73×10^{-4} *M* at pH 7.5 and 3°.

During the past several years, relaxation methods have assumed increasing importance as tools in the elucidation of complex enzymic mechanisms. The temperature-jump method, in particular, has been successfully used to determine not only the rate constants in enzymic systems but also to detect the presence of intermediates which would not ordinarily be observable in conventional steady-state kinetics.[2] The method is now proving to be most useful in studies of the active site of chymotrypsin, particularly as regards the protonic ionization of the two imidazole groups in this enzyme.

In a previous communication[3] we reported that the two imidazole groups of α-chymotrypsin (CT), in the presence of a pH indicator (phenol red), are the source of a single large relaxation effect in the neutral pH range. The results of studies of the protonic behavior of these imidazoles as a function of pH, indicator concentration, and enzyme concentration were consistent with the following mechanism.[4]



The over-all relaxation time for the above system is given by eq 2 and 3, derived using a steady-state treatment for $[H^+]$ and $[OH^-]$.

$$\tau^{-1} = \tau_{DT}^{-1} + \tau_{P}^{-1} + \tau_{H}^{-1} \quad (2)$$

(1) This is paper No. 30 from this laboratory. Please request reprint by this number.
(2) G. G. Hammes and P. Fasella, *J. Am. Chem. Soc.*, **84**, 4644 (1962); G. G. Hammes and P. Fasella, *ibid.*, **85**, 3929 (1963); R. Cathou and G. G. Hammes, *ibid.*, **86**, 3240 (1964).
(3) A. Yapel and R. Lumry, *ibid.*, **86**, 4499 (1964).
(4) M. Eigen, G. G. Hammes, and K. Kustin, *ibid.*, **82**, 3482 (1960).