JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

(Registered in U. S. Patent Office) (© Copyright, 1957, by the American Chemical Society)

VOLUME 79 FEBRUARY 27, 1957 NUMBER 4

PHYSICAL AND INORGANIC CHEMISTRY

[CONTRIBUTION FROM GIBBS LABORATORY, DEPARTMENT OF CHEMISTRY, HARVARD UNIVERSITY]

The Optical Rotatory Dispersion of Polypeptides and Proteins in Relation to Configuration[1]

BY JEN TSI YANG AND PAUL DOTY

RECEIVED AUGUST 31, 1956

The rotatory dispersion of poly-γ-benzyl-L-glutamate in solvents other than those already reported is given. Again this dispersion is found to be simple when the chain configuration is that of a random coil and complex when it is helical. The assessment of solvent effects shows that these are moderate for the helical configuration and rather large for the randomly coiled form, but they are always secondary to the much larger differences due to chain configuration. The fit of the dispersion data of the helical form by Moffitt's equation and other evidence demonstrates that the helical form exists with only one screw sense, probably right-handed. For a residue weight of 110, the average in proteins, the specific rotation $[\alpha]_D$ for the helical form has slightly positive values which lie approximately 100° above that of the randomly coiled or non-helical form. Dispersion measurements on seven proteins in aqueous solution, extending earlier observations, show that (1) the dispersion is simple, (2) values of the dispersion constant λ_c are near 220 mμ for denatured or non-globular proteins and range from 230 to 270 mμ for native proteins, and (3) values of $[\alpha]_D$ range from −30 to −70° for native proteins and can be lowered by 20 to 60° upon denaturation. The rotatory behavior of the two configurations present in polypeptides (α-helix and randomly coiled form) suffices to explain these three characteristics of proteins since it is found that hypothetical mixtures of the two forms in the range of 0 to 40% do exhibit a simple dispersion and show a progressive increase in λ_c from about 212 to 268 mμ accompanied by an increase in $[\alpha]_D$ of about 40°. On this basis it is estimated that for globular proteins in solution the fraction of residues in the helical configuration for proteins examined thus far ranges from about 15% (ribonuclease) to 40% (ovalbumin). According to the present limited data on proteins, the α-helix appears to be the only recurring structural feature which is reflected by a consistent and substantial rotational shift. Dispersion measurements on proteins in non-aqueous solvents indicate that the helical content can be shifted reversibly and over a wider range than in aqueous solutions. This is explored in four specific cases. Silk fibroin and the B chain of insulin can be brought nearly completely into the right-handed α-helical configuration with the attending display of complex dispersion. The helical content of both the oxidized and native forms of ribonuclease and bovine serum albumin can be increased somewhat but probably not to more than 50% since the dispersion remains simple. Since the presence or absence of the cystine bridges had little effect in these last two cases it appears that particular side group interactions alone are sufficient to prevent the more complete development of the helical configuration in proteins, at least in solution.

Introduction

Recent work[2-5] has shown that poly-γ-benzyl-L-glutamate, poly-L-glutamic acid and several other polypeptides can exist in either a helical or randomly coiled configuration in solution. The stable configuration is decided by the choice of solvent, temperature and molecular weight of the polypeptide. Throughout these investigations it was evident that the helical configuration exhibited a positive rotation (sodium-D line) whereas the random coil form was always attended by a substantial negative rotation. When rotatory dispersion measurements were carried out the differences were amplified because it was found that the random coil configuration displayed simple dispersion whereas that of the helical configuration was complex.[6] This behavior indicated rather clearly that the differences in rotation were not merely the result of a solvent effect or vicinal actions of surrounding groups but were due in large part to the contribution of the helical core itself to the rotation. Since the helical configuration could exist in two possible forms (right- and left-handed helices) this conclusion implied that one or the other of the forms was predominantly and probably ex-

(1) This paper is Polypeptides. XI. The preceding member of the series is published in THIS JOURNAL, **78**, 4810 (1956).

(2) P. Doty, J. H. Bradbury and A. M. Holtzer, *ibid.*, **78**, 947 (1956).

(3) E. R. Blout and A. Asadourian, *ibid.*, **78**, 955 (1956).

(4) P. Doty, A. Wada, J. T. Yang and E. R. Blout, presented at the International Symposium on Macromolecules, Rehovoth, Israel, April 2-5, 1956; *J. Polymer Sci.*, January, 1957.

(5) K. Imahori, R. H. Karlson, E. R. Blout and P. Doty, unpublished work.

(6) P. Doty and J. T. Yang, THIS JOURNAL, **78**, 498 (1956).

clusively present. Experiments with polymers made with varying amounts of D- and L-isomeric residues[7] show clearly that the stable form of the helix for polypeptides composed of L-residues contributed roughly +50° to $[\alpha]_D$, the specific rotation, for poly-γ-benzyl-L-glutamate.

While this work was in progress theoretical investigations of the optical properties of helical macromolecules were under way. Moffitt[8] showed that under favorable conditions that existed in the α-helical configuration the rotatory dispersion should exhibit marked departures from the simple one-term Drude equation. It was soon found that Moffitt's equation accurately fitted the dispersion measurements.[6,9] Moreover, Moffitt[10] was able to calculate the values of the constants characterizing the dispersion apart from that of the Drude term. That is, in his equation (M_0 is the molecular weight per residue and n is the refractive index of the solvent)

$$[\alpha]_\lambda = \left(\frac{100}{M_0}\frac{n^2+2}{3}\right)\left[\frac{a_0\lambda_0^2}{(\lambda^2-\lambda_0^2)} + \frac{b_0\lambda_0^4}{(\lambda^2-\lambda_0^2)^2}\right] \quad (1)$$

the values calculated for a right-handed α-helix were $b_0 = -580°$ and $\lambda_0 = 2000$ Å. In an analysis of the dispersion data on poly-γ-benzyl-L-glutamate in five solvents and poly-L-glutamic acid in two solvents in which the helical configuration was stable, mean values of $b_0 = -630°$ and $\lambda_0 = 2120$ Å. were obtained. This agreement offered strong support to the contention that only a single helical configuration is present and, in addition, tentatively identified the helical configuration present as the right-handed one. It is important to note that no attempt was made to evaluate the constant a_0 since this simple dispersion term is made up of contributions from the helical configuration as well as from the intrinsic residue rotation and in practice is apt to be subject to large solvent effects.

At about the same time Fitts and Kirkwood[11] applied the earlier theory of optical rotation devised by Kirkwood[12] to the problem of calculating the specific rotation of a helical polypeptide in excess of that due to the intrinsic activity of the residues alone. They obtained

$$[\alpha]_D - [\alpha_0]_D = 567\alpha_1^2\beta^2(n^2+2)/(3M') \quad (2)$$

where α_1 is the polarizability per turn, β the anisotropy ratio and M' the molecular weight per helix turn. For poly-γ-benzyl-L-glutamate the calculated value was found to be +31°. Fitts and Kirkwood[13] sought to test their prediction by comparing this value with the difference in rotation we have observed between solutions in which the configuration is helical and in which it is randomly coiled. While this gave good agreement for the first published cases, further results included in this report show that there is, as might be expected, a moderate solvent effect on the helical configuration and a pronounced one on the randomly coiled configuration. Thus between chloroform and trifluoroacetic acid solutions, the difference is 57°, while between dimethylformamide and dichloroacetic acid solutions it is 22°. Consequently, this test can hardly be conclusive but the prediction for the behavior of a right-handed helix does fall within its rather broad perimeter.

In their more recent communication[13] it is stated that the same type of dispersion shown in equation 1 is obtained if the quantity $\alpha_1\beta$ can be represented by a single Drude dispersion term. However, the equivalence of this to Moffitt's analysis is not yet apparent because Moffitt's complex dispersion term is negative whereas Fitts and Kirkwood's proposal would lead to the complex term having a positive sign, that is, the same as the term they have already evaluated. For example, fitting equation 1 to typical data for poly-γ-benzyl-L-glutamate (ethylene dichloride) in the helical form gives

$$[\alpha]_D = 18.9 - 8.7 = +10.2$$

which shows that the complex dispersion term (the second term) is negative although the sum is positive. While these apparent difficulties and others[10] may be resolved when fuller details are available, we will for the present employ equation 1 in the treatment of our data, keeping in mind the reservations that were set forth in its derivation.[9,10]

It is against the background of these developments that we have examined further the rotatory dispersion of synthetic polypeptides and some proteins and protein fragments. With polypeptides the dispersion in other solvents and the dispersion of a third configuration resembling the β-form of the solid state have been examined. By combining the dispersion characteristics of helical and coiled forms, a suggestion is made concerning the determination of the amount of helical configurations in proteins and the origin of the changes in λ_c and $[\alpha]$ upon denaturation. Finally, the extent to which the helical configuration can be developed in some proteins or protein fragments is explored and in some instances the helix-coil transition as reflected in optical rotation is studied and compared with that observed in synthetic polypeptides.

Experimental Details

Optical rotations were measured with a Rudolph precision ultraviolet polarimeter, model 80S, using a zirconium arc and a mercury (type SH, Hanovia) quartz lamp as light sources. The usual wave length range extended from 313 to 750 mμ and it was covered using 2 dm. water-jacketed tubes. Accuracy was reduced at both extremes of the range because of the necessity of using wide slit widths. Water from a circulating water-bath was used to control the temperature in the range of 8 to 60°. Unless otherwise specified, measurements were made at 20°. The concentration was in the range of 0.5 to 1 g./dl. unless indicated otherwise. All solutions were clarified, if necessary, by filtering through medium sintered-glass.

The polypeptides were selected from a number made by Dr. E. R. Blout and co-workers.[14,15] The sample numbers permit reference to be made to other work done on the same samples in related publications. Silk fibroin (*Bombyx mori*) was a gift from Dr. C. H. Bamford. Salmine sulfate was obtained through the courtesy of Dr. W. Carroll. The samples of bovine plasma albumin, bovine ribonuclease,

(7) (a) P. Doty and R. D. Lundberg, THIS JOURNAL, **78**, 4810 (1956); (b) E. R. Blout, P. Doty and J. T. Yang, *ibid.*, **79**, 749 (1957); (c) P. Doty, R. D. Lundberg and J. T. Yang, *Proc. Natl. Acad. Sci.*, in press.

(8) W. Moffitt, *J. Chem. Phys.*, **25**, 467 (1956).

(9) W. Moffitt and J. T. Yang, *Proc. Natl. Acad. Sci.* (*Wash.*), **42**, 596 (1956).

(10) W. Moffitt, *ibid.*, **42**, 736 (1956).

(11) D. D. Fitts and J. G. Kirkwood, *ibid.*, **48**, 33 (1956).

(12) J. G. Kirkwood, *J. Chem. Phys.*, **5**, 479 (1937).

(13) D. D. Fitts and J. G. Kirkwood, THIS JOURNAL, **78**, 2650 (1956).

(14) E. R. Blout and R. H. Karlson, *ibid.*, **78**, 941 (1956).

(15) E. R. Blout and M. Idelson, *ibid.*, **78**, 498 (1956).

bovine fibrinogen, lysozyme and insulin (zinc-free) were obtained from Armour Company. A second insulin sample was given by the Boots Drug Company, Ltd. The ovalbumin was from a sample prepared by Dr. S. Katz while a member of this Laboratory. For preparing insulin solutions in dimethylformamide the insulin was first made amorphous by lyophilizing a dilute hydrochloric acid solution.

The solvents were all of reagent grade. Dichloroacetic acid, dioxane and dimethylformamide were further purified by distillation before use. The hydrazine employed contained 5% water.

I. Rotatory Dispersion and Configuration in Polypeptides

Solvent Effect on the Dispersion for Helix and Coil Configurations.—Although examples of the characteristic dispersion for the two principal configurations of synthetic polypeptides in solution has been given elsewhere[4,6] the effect of varying the solvent while maintaining one or the other configuration has not been reported. Such results are shown in Fig. 1 for a sample of poly-γ-benzyl-L-glutamate (No. 421) having a weight average molecular weight of 130,000.[2] Intrinsic viscosity or flow birefringence measurements were used to establish that the configuration in the five solvents where positive rotations are found (at long wave lengths) is that of the α-helix.[2] Similarly, it was found that in dichloroacetic acid and hydrazine (as well as trifluoroacetic acid which is not shown) the configuration is randomly coiled and the rotations in these cases are negative at all wave lengths.

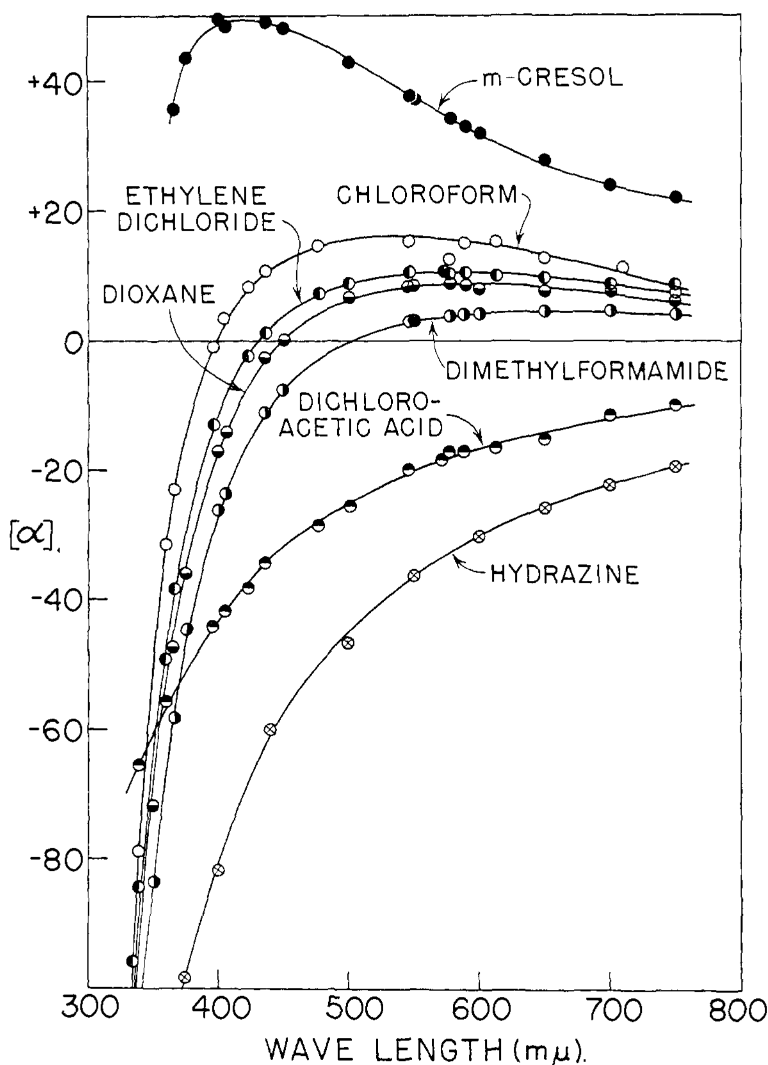

Fig. 1.—Rotatory dispersion of poly-γ-benzyl-L-glutamate (M_w = 130,000) in several solvents.

These measurements make clear that both configurations are subject to considerable solvent effect. The case of *m*-cresol is particularly striking and may be the result of a solvation of the benzyl groups in such a way that they take up a more ordered arrangement on the surface of the helix core and thereby contribute a new chromophoric group with an optical activity conferred by the helical arrangement.[16] The dispersion found for the case of the other four solvents in which the helical configuration is stable cluster together but still differ from each other by much more than experimental error and the effect of refractive index (*i.e.*, the factor $(n^2 + 2)/3$). The dispersion for the cases where the configuration is the random coil spread much more. The values of $[\alpha]_D$ for dichloroacetic acid, hydrazine and trifluoroacetic acid are, respectively, -17.5, -32 and $-46°$. Since the interaction of the solvent with the polypeptide can be so much more extensive in the coil configuration this greater solvent effect is not unexpected.

(16) In a flow birefringence study (soon to be published) we have observed that the intrinsic birefringence in *m*-cresol is positive in contrast to negative values for the other solvents. If the configuration of the main-chain is unchanged this is indicative of a substantially different orientation of the side chains.

Modified Dispersion Plot.—The most frequent representation of rotatory dispersion is that given by Drude

$$[\alpha] = \sum \frac{k_i}{(\lambda^2 - \lambda_i^2)} \qquad (3)$$

where k_i and λ_i are known as the rotation and dispersion constants, respectively. In many cases dispersion data can be fitted with a single term of this sum. This is generally not due to there being a single excited state but rather to the circumstances that can arise when the lowest lying electronic states lie close together.[9,17] In this case one usually writes

$$[\alpha] = \frac{k}{(\lambda^2 - \lambda_c^2)} \qquad (4)$$

Note that λ_c is used for normal dispersion and is to be distinguished from λ_0 used when data are fitted by equation 1. Following Lowry's suggestion, it has become customary to plot $1/[\alpha]$ against λ^2 and obtain the value of λ^2_c from the intercept and k from the slope. When, as in the present case, the

(17) J. A. Schellman, *Compt. rend. trav. lab. Carlsberg, Ser. chim.*, in press.

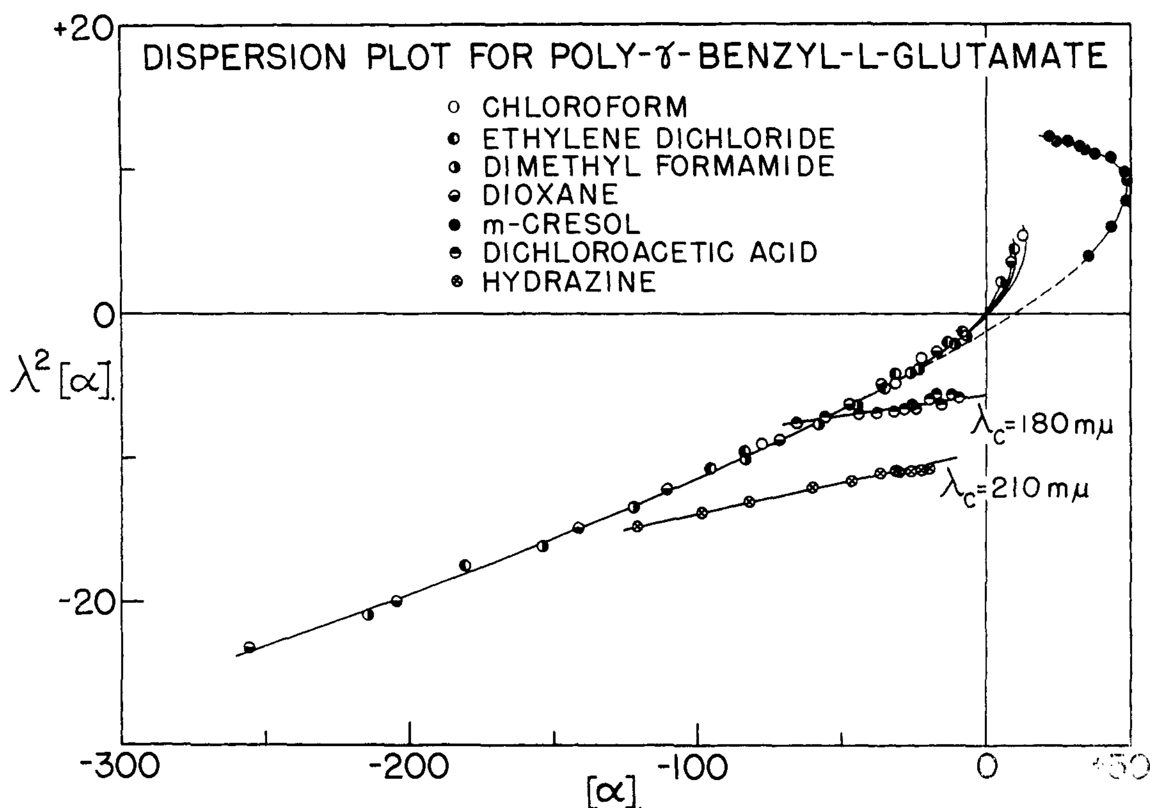

Fig. 2.—Modified dispersion plot of data appearing in Fig. 1.

value of λ_c is desired with the greatest precision permitted by the data, we have found it much more advantageous to plot $\lambda^2[\alpha]$ against $[\alpha]$ and derive λ_c from the square root of the slope of the straight line and k from the intercept. This procedure eliminates two disadvantages of the Lowry plot: (a) the determination of λ^2_c from an intercept near zero where a greatly enlarged plot would be necessary to adequately represent the data and (b) the concentration of the short wave length data in a small region of the plot. The reliable evaluation of λ_c to within a 2% probable error was important for many of the cases investigated and reported below. When data are fitted by equation 4 the specification of $[\alpha]$ at one wave length is equivalent to stating the value of k. Consequently, we will generally employ $[\alpha]_D$ rather than k.

The dispersion data already presented (Fig. 1) are shown in this type of plot in Fig. 2. The data in dichloroacetic acid and hydrazine fall on straight lines, showing simple dispersion, characterized by λ_c values of 180 and 210 mμ, respectively. The remaining dispersion data, with the exception of cresol solutions, fall on the same curved line over most of the wave length range. Separation is evident only for positive rotations. Although the results in *m*-cresol stand somewhat apart it is assumed that they constitute an exaggerated type of deviation from what appears to be a characteristic dispersion pattern subject to only limited solvent effects.

It is thus clear that the two configurations of the polypeptide chain have unique and characteristic rotatory dispersions. In addition to the examples given here we may add the data on the same polypeptide in the mixed solvent: ethylene dichloride and dichloroacetic acid (80:20).[6] In this case a shift in temperature sufficed to produce one configuration or the other and the dispersion shifted accordingly. Moreover, the same two characteristic dispersions have been found for poly-L-glutamic acid[4] when *p*H was employed to produce a similar transition. In all of these cases of complex dispersion of solutions having the helical configuration an excellent fit can be made by means of equation 1.[9]

It should be pointed out, however, that the complex dispersion data can be fitted with a two-term Drude equation

$$[\alpha] = \frac{k_1}{\lambda^2 - \lambda_1^2} + \frac{k_2}{\lambda^2 - \lambda_2^2} \qquad (5)$$

For example, poly-γ-benzyl-L-glutamate (No. 421) in ethylene dichloride is fitted with $k_1 = 13.9$, $k_2 = -8.0$, $\lambda_1 = 0$ and $\lambda_2 = 0.282$ μ. Poly-L-glutamic acid (No. 17513) in dioxane–water (2:1) at *p*H 4.7 is fitted with $k_1 = 23.3$, $k_2 = -17.4$, and $\lambda_1 = 0$ and $\lambda_2 = 0.272$ μ. We prefer the use of equation 1 to equation 5 because of the existence of a theoretical basis for the former and the use of fewer constants.

Preliminary dispersion measurements in the range of 248 to 313 mμ on some polypeptides have been made. These results indicate a deviation from the values expected on the basis of either equations 1 or 5 and are the subject of further investigation.

Rotatory Dispersion of Low Molecular Weight Polypeptides.—Previous infrared spectral studies[8]

and observations of optical rotation (sodium-D line)[18] had indicated that at sufficiently low molecular weights the helical configuration did not exist, but that one observed the randomly coiled configuration in strongly solvated solvents and an intermolecularly bonded form (*i.e.*, an aggregated state similar to the β-form known in the solid state) in less strongly interacting solvents. The intermolecularly bonded and hence aggregated form is converted to the solvated randomly coiled form upon dilution.

Dispersion measurements were made on a very low molecular weight sample[19] of poly-γ-benzyl-L-glutamate in several concentrations in chloroform and in chloroform–formamide (chloroform saturated with formamide) and dichloroacetic acid. The results are shown in Fig. 3 and plotted in the modified form in Fig. 4. At high concentrations in chloroform positive rotations are found and these increase with diminishing wave length, unlike any previous observations on higher molecular weight samples of this polypeptide. At the low concentration of 0.4 g./dl., however, negative rotations similar to those in dichloroacetic acid are found. This striking dependence of dispersion behavior on concentration is further evidence of the micelle formation indicated in previous studies.

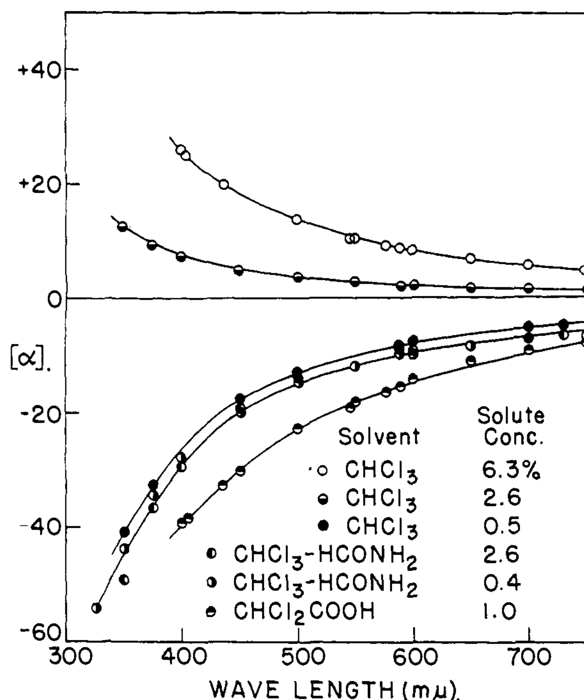

Fig. 3.—Rotatory dispersion of very low molecular weight poly-γ-benzyl-L-glutamate in several solvents at different concentrations.

When chloroform–formamide (about 0.5% formamide) is used as solvent, the dispersion typical of the coiled form is observed at all concentrations. Thus the intermolecularly hydrogen-bonded micelles are unstable in the presence of small amounts of more strongly hydrogen-bonding agents.

(18) A. E. Woodward and P. Doty, to be published soon.

(19) This sample (No. 1186B) was prepared by fractionation from a polypeptide made in dioxane with a ratio of anhydride to initiator of 4. (E. R. Blout and R. H. Karlson, unpublished work.) Its intrinsic viscosity in dichloroacetic acid was 0.062.

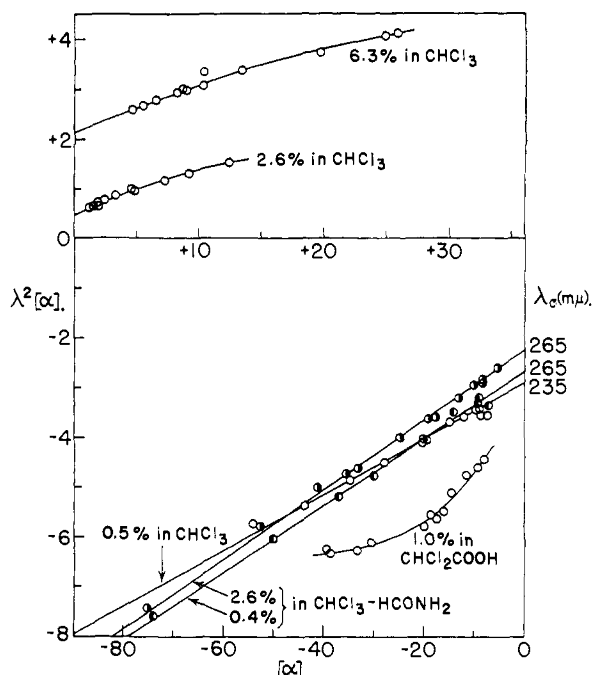

Fig. 4.—Modified dispersion plot of data appearing in Fig. 3.

II. Rotatory Dispersion and Configuration in Proteins

Rotatory Dispersion of Proteins.—The few measurements that have been made of the dispersion of protein solutions obey a one-term Drude equation.[20–23] Two other generalizations emerge from the study of the rotation and dispersion of proteins. One is the widely recognized dependence of optical rotation on denaturation[24]: most globular proteins exhibit specific rotations $[\alpha]_D$ in the range of -30 to $-70°$ in the native state and this is lowered by 20–60° upon denaturation. The other is the observation of a similar correlation between λ_c and the state of the protein: Linderstrøm-Lang and Schellman[21] have shown that in several cases λ_c is lowered from values of 250–260 mμ for the native state to 220–230 mμ for the denatured state. The problem with which we are concerned in this section is how to reconcile these three generalizations on proteins with the behavior of polypeptides described in the previous section. Before taking up this problem we shall summarize measurements we have made on several proteins which add further support to the general behavior of proteins just described.

(20) L. Hewitt, *Biochem. J.*, **21**, 216 (1927).

(21) K. Linderstrøm-Lang and J. A. Schellman, *Biochim. Biophys. Acta*, **15**, 156 (1954).

(22) C. Cohen, *J. Biophys. Biochem. Cytology*, **1**, 203 (1955).

(23) J. A. Schellman and C. S. Schellman, *Arch. Biophys. Biochem.*, in press. Some measurements below 313 mμ on bovine plasma albumin do show deviations from a one-term Drude equation but consideration of this exception can perhaps be postponed since the dispersion is simple in the commonly studied region where we wish to make comparisons.

(24) See, for example, B. Jirgensons, *Arch. Biochem. Biophys.*, **39**, 261 (1952); **41**, 333 (1952); R. B. Simpson and W. Kauzmann, THIS JOURNAL, **75**, 5139 (1953); C. Cohen, *Nature*, **175**, 129 (1955).

Our measurements of the dispersion of seven proteins under various conditions in aqueous solvents are summarized in Table I where the constants characterizing their simple rotatory behavior over the range of 300–750 mμ are given together with $[\alpha]_D$. An example is given in Fig. 5 where the dispersion data for bovine plasma albumin at different pH values are plotted in the modified fashion previously described. It is seen that the plots are linear and that the λ_c values diminish somewhat as the pH is lowered below 4.0 in keeping with the reversible expansion accompanying the denaturation noted by several investigators.[25,26] Ovalbumin, the second entry in Table I, does not undergo denaturation at pH 2.2 until heated, whereupon the change in both λ_c and $[\alpha]_D$ is very substantial. It is interesting to observe that the dispersion of a synthetic combination of half native and half denatured ovalbumin remains simple (one-term Drude equation) with intermediate values of the constants. Fibrinogen and lysozyme exhibit typical values of the dispersion constants whereas the λ_c value for ribonuclease is exceptionally low[27] for a globular protein. Still lower values of λ_c are found for salmine, a protein that is probably in the randomly coiled form. Silk fibroin in concentrated LiBr solution (54 g./100 g.) has an intermediate value of λ_c, a point to which we return later. Thus in each of these cases the lowering of λ_c is accompanied by a lowering (in an absolute sense) of $[\alpha]_D$. In every respect these measurements substantiate the generalizations stated at the beginning of this section.

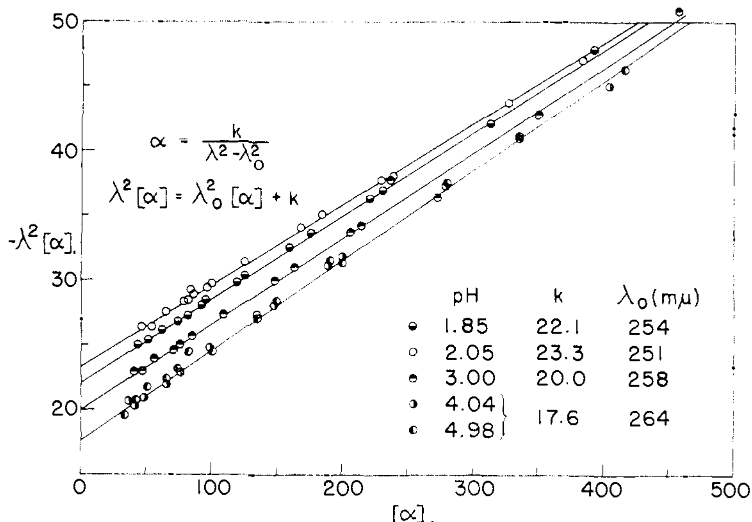

Fig. 5.—Modified plot of dispersion data for bovine serum albumin at different pH (no added salt).

TABLE I

ROTATORY DISPERSION OF SOME PROTEINS IN AQUEOUS SOLUTIONS

Protein[a]	Solvent	pH	λ_c (mμ)	$-k$	$[\alpha]_D$
Serum albumin	Water	5.0	264	17.6	−63.2
(bovine)	HCl	4.0	264	17.6	−63.2
	HCl	3.0	258	20.0	−71.0
	HCl	2.1	251	32.3	−82.0
	HCl	1.9	254	22.1	−77.6
Ovalbumin	Water	4.7	270	8.7	−31.4
	HCl (a)	2.2	270	9.0	−32.7
(5′ at 100°)	HCl (b)	2.2	243	15.4	−53.3
	(a) + (b)		253	12.4	−43.0
Fibrinogen (bovine)	0.1 M KCl	6.3	256	13.8	−48.6
Lysozyme	2 M NaCl	I.P.	254	13.5	−47.1
Ribonuclease (bovine)	0.2 M NaCl	I.P.	231	21.0	−71.5
Salmine sulfate	Water	6.4	222	26.0	−88.7
	NaOH + 2 M NaCl	11.1	213	23.7	−81.5
Silk fibroin	LiBr–H_2O		247	6.9	−23.7

[a] The concentration of serum albumin was determined spectrophotometrically using E = 6.7 at 279 mμ. The concentration of the other proteins was determined by Kjeldahl nitrogen analyses: ovalbumin, N 16.0%; fibrinogen, N 16.9%; lysozyme, N 18.6%; ribonuclease, N 16.0% and salmine, N 18.6%.

As a result of the studies reported in the last part of this paper we have made numerous observations on most of these same proteins in non-aqueous solutions and these results are summarized in Table II. In the case of insulin it is seen that the four solvents used show a progressive lowering of λ_c and $[\alpha]_D$ ending with formic acid on which there is infrared evidence of its producing a denatured or β-form configuration.[28] With silk fibroin it is seen that trifluoroacetic acid produces a very low value of $[\alpha]_D$ (particularly considering its low mean residue weight) and only a moderately low value of λ_c and that the addition of ethylene dichloride raises both. A similar situation at somewhat higher absolute values is found when dichloroacetic acid replaces trifluoroacetic acid. A corresponding shift is found for ribonuclease as well. This brief survey indicates that the pattern of dispersion behavior in non-aqueous solutions of proteins is similar to that in aqueous ones: however, with the greater variety of organic solvents, a larger range of values of λ_c and $[\alpha]_D$ is observed. As a consequence, the experience with polypeptides in organic solvents may not be as remote from the problems of the rotation and configuration of proteins as might have been thought.

A Proposal for Relating Dispersion Constants to Configuration.—We are now confronted with the general result that the specific rotation is lowered upon denaturation by as much as 40° (or more in cases not studied here) whereas the lowering observed in going from the helical configuration to the random coil in poly-L-glutamic acid (a polypeptide whose residue weight is close to that of the

(25) J. T. Yang and J. F. Foster, THIS JOURNAL, **76**, 1588 (1954).

(26) C. Tanford, J. G. Buzzell, D. G. Rands and S. A. Swanson, *ibid.*, **77**, 6421 (1955).

(27) The value given is the same as that found in reference 23 and apparently should replace the higher value previously given.[21]

(28) E. J. Ambrose and A. Elliott, *Proc. Roy. Soc.* (*London*), **A208**, 75 (1951).

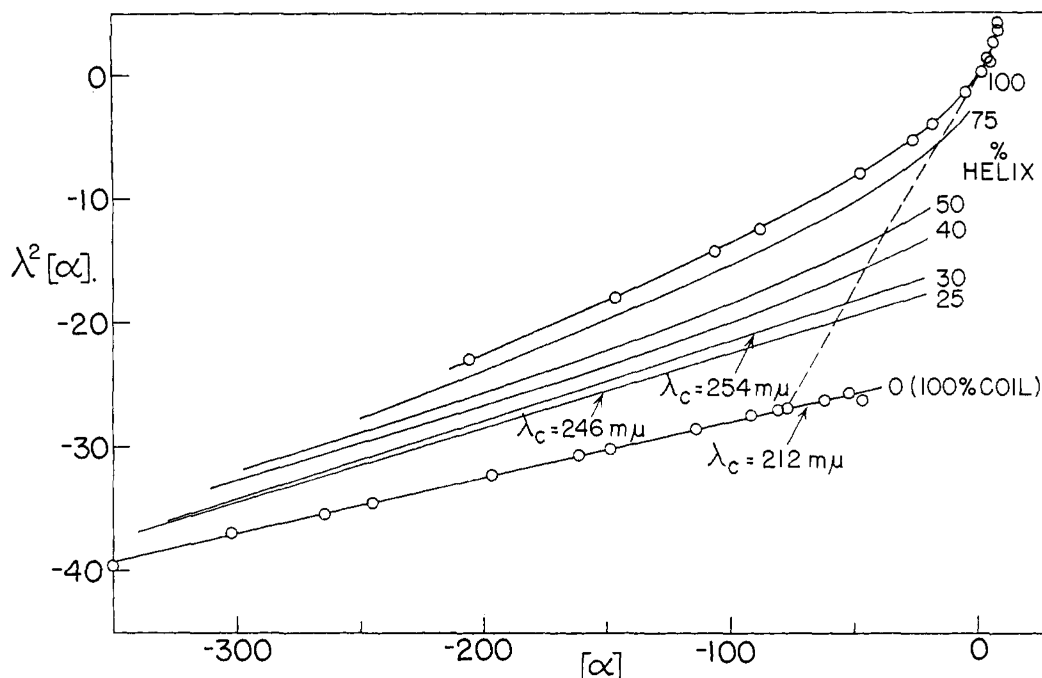

Fig. 6.—Modified plots of dispersion data for poly-L-glutamic acid in the helical and randomly coiled forms together with combinations thereof in varying proportions. The dashed line is drawn through the values at λ = 5893 Å.

proteins studied) is about 80°. Although both the helical and random coil configurations of polypeptides are subject to considerable solvent effects it is clear that the consistent behavior found in both proteins and polypeptides could not arise from this random effect. In the absence of competing proposals it appears likely that the substantially higher rotations observed for native proteins relative to their denatured counter parts do indeed arise from the helical configuration present in the native protein. To pursue this proposition we must assume that the helical configuration is essentially the α-helix as found in synthetic polypeptides and that the form that is stable for polypeptides made of L-residues (presumably the right-handed form) is dominant in proteins. Then it would follow from the observed changes in specific rotation that the helical configuration is present to the extent of only 25 to 50% in most globular proteins thus far studied. Now returning to the model behavior furnished by the synthetic polypeptides we must ask what kind of dispersion and λ_c values would be found for such mixtures of helical and randomly coiled configurations. (It should be kept in mind that any breakdown of a helically configured region of a protein will bring about a loss in the augmentation of rotation and hence the random coil can serve as the model for the non-helical region even though it may fail to represent the spatial distribution of the peptide segments in denatured regions.)

TABLE II

ROTATORY DISPERSION OF SOME PROTEINS IN NON-AQUEOUS SOLUTIONS

Protein[a]	Solvent[b]	λ_c (mμ)	$-k$	$[\alpha]_D$
Insulin	EDC-DCA(92:8)	300	5.8	−20.0
	DMF	250	10.0	−35.4
	EG	243	18.7	−65.1
	EDA	237	19.0	−62.0
	DCA	236	17.6	−60.7
	FA (88%)	230	24.8	−84.0
Serum albumin	DCA	242	19.0	−66.0
(bovine)	EDC-DCA(3:1)			−22.5
Ribonuclease	EDC-DCA(4:1)	290	10.8	−41.1
(bovine)	DCA	240	18.0	−62.2
	FA-EDC(1:1)	275	10.3	−38.0
	FA	245	15.8	−55.1
Salmine	DCA	242	14.8	−50.8
Silk fibroin	EDC-DCA(1:1)	279	6.6	−24.8
	DCA	255	14.0	−50.4
	EDC-TFA(1:1)	242	18.1	−62.9
	TFA	234	22.3	−77.2
Serum albumin	DCA	246	17.2	−59.9
(oxidized)	EDC-DCA(1:1)	265	13.4	−49.1
	EDC-DCA(4:1)	325	5.8	−23.1
Ribonuclease	DCA	233	15.5	−52.8
(oxidized)	EDC-DCA(3:1)	290	8.6	−31.8
	FA	238	20.1	−69.1
	FA-EDC(1:1)	278	12.9	−48.0

[a] Protein concentrations were determined by direct weighing and then corrected to correspond to Kjeldahl nitrogen analyses on aqueous solutions of the same samples. Solvent compositions are given as volume ratios. [b] Abbreviations: DMF, dimethylformamide; EG, ethylene glycol; EDA, ethylenediamine; FA, formic acid; DCA, dichloroacetic acid; EDC, ethylene dichloride; TFA, trifluoroacetic acid.

To answer this question the dispersion data for poly-L-glutamic acid (No. 175B) in the 300–750 mμ region has been plotted in Fig. 6 for both the helical (pH 4.7) and randomly coiled (pH 6.6) forms in dioxane–0.2 M NaCl (1:2).[4] Then dispersion curves have been calculated by combining these two results in various proportions as shown. Three interesting conclusions immediately follow. (1) With normal experimental accuracy the rotatory

dispersion is indistinguishable from the simple type (linear behavior) up to helical contents of 30 to 40%. (2) In the range of helical content where the dispersion remains simple the value of λ_c increases steadily from 212 mμ for no helical content to 254 mμ for 30% and 268 mμ for 40% (ignoring in the last case the small amount of curvature present). (3) The value of $[\alpha]_D$, shown by the dashed line, increases progressively from -77 to $-46°$ over the same interval. For a given protein these rotations should be multiplied by the ratio of the residue weight of glutamic acid (129) to the average residue weight of the protein. These three conclusions from Fig. 6 bear a striking resemblance to the three generalizations regarding the rotatory dispersive behavior of proteins discussed in the preceding section. The similarity in rotatory dispersive behavior is so close that it immediately suggests that the observations on proteins can be adequately explained in terms of a single variable, the fraction of the residues in the α-helical configuration. Thus the values of $[\alpha]_D$ and λ_c should each lead to a similar estimate of the helical content of a given protein and the maximum extent of denaturation it can undergo would then correspond to the disordering of this particular amount of helical configuration.

Using the calibration furnished by poly-L-glutamic acid in Fig. 6, a couple of examples using data in Tables I and II illustrate this point. Ribonuclease in saline solution is estimated to be 14% helical from its λ_c value and 20% by its $[\alpha]_D$ value. However, in formic acid–ethylene dichloride (1:1) the estimated helical content is increased to 44% judged by λ_c and 58% as judged by $[\alpha]_D$. Using published[21] data for insulin at *p*H 3 ($\lambda_c = 265$ and $[\alpha]_D = 32°$) one estimates the helical content to be 38 and 61% from λ_c and $[\alpha]_D$, respectively. From the data in Table II one would conclude that in dichloroacetic acid the helical content of insulin is lowered to 16% as judged by both λ_c and 24% by $[\alpha]_D$. However, by adding ethylene dichloride to the dichloroacetic acid solution to the limit of precipitation (92:8) the helical content can be increased to 62% (λ_c) and 75% ($[\alpha]_D$). These examples show substantial agreement and commend this procedure as a semi-quantitative means of estimating the amount of helical configuration subject to certain limitations. Moreover, an additional check is provided in that the simple dispersion should disappear at about 40% helical content. These points are explored further in the next section. It is best to discuss first the limitations that can be clearly foreseen.

The simplest form of the conclusions reached by comparing the optical rotatory behavior of proteins and polypeptides is that in both cases the optical dispersion is the resultant of the intrinsic residue rotations (always negative for L-amino acids) and the rotatory effects of that part of the polypeptide chain that is in the helical configuration (always positive provided that only right-handed helices are present). The intrinsic residue rotations must, of course, vary from one residue to another. This naturally prevents the precise identification of a given value of $[\alpha]_D$ with a certain helical content. Only insofar as the average of the intrinsic residue rotations of a given protein approach that of the polypeptide used for calibration (poly-L-glutamic acid in the above case) can a relation between $[\alpha]_D$ values and helical content be expected to hold. The experience gained in many studies of optical rotation of proteins suggests strongly that the average residue rotations of proteins, as judged from the values for what appear to be completely denatured forms, is generally restricted to the range of -80 to $-110°$. Thus the base line from which the helical content is to be measured in terms of $[\alpha]_D$ values is subject to considerable variations. The choice of $-77°$ based on poly-L-glutamic acid and adjusted by multiplying by the ratio of residue weights appears to be a suitable standard procedure but the use of the $[\alpha]_D$ value for the completely denatured protein is preferable when it is available.

The average specific residue rotation is also subject to alteration by solvent effects. While these are probably small in most aqueous solutions they have been shown to be as high as 20° for poly-L-glutamic acid when the ionic strength of neutral solutions is varied over the range of 0 to 4 *M* NaCl.[29] On the basis of polypeptide experience solvent effects in organic solvents can be even larger. However, some cancellation would be expected in the case of most proteins because of the diversity of residues. As a consequence, this variation in the average intrinsic residue rotation and the solvent effects it appears that for the present an uncertainty of about 20° must be attached to the value of $[\alpha]_D$ when it is being considered for assignment of helical content. Of course, the *relative* helical content in a given protein can be estimated with less certainty since the base line is then subject only to solvent effects since the intrinsic residue rotations remain constant.

The dispersion constant λ_c arises principally from the main-chain and is presumably less affected by changes in residue composition. It shows some solvent effect when very strongly interacting solvents are used but in aqueous solutions it appears to depend only on helical content as judged from the existing data. It is likely, therefore, that it is subject to less uncertainty than $[\alpha]_D$ as a basis of estimating helical content in aqueous solutions. Again when changes in helical content of a given protein are desired it should provide values of the relative amount of helix that compare in accuracy with those obtained from $[\alpha]_D$ values. This is illustrated by the data in Table I for bovine serum albumin where it can be seen that the differences between the λ_c values are precisely reflected by proportional differences between the corresponding $[\alpha]_D$ values.

We can conclude this discussion by stating that λ_c and $[\alpha]_D$ appear to offer equivalent ways of assessing the amount of helical configuration within a given protein but that for estimating the absolute amount the value of λ_c is preferred. Both means of estimating the helical content are subject to larger errors when strongly interacting solvents are used.

The present calibration of these two methods using the poly-L-glutamic acid data is, of course,

(29) A. Wada and P. Doty, unpublished results.

subject to considerable improvement as similar data on other polypeptides and particularly those containing a variety of groups becomes available. To make the best of what is now available the poly-L-glutamic acid data should be modified to apply to an average residue weight of 110 rather than the value of 129 characteristic of this polypeptide. If so, the values of the extreme ends of the scale of $[\alpha]_D$ should be multiplied by 129/110. Thus the values of -77 to $+5°$ found for poly-L-glutamic acid should be shifted to -90 to $+6°$ for typical proteins. The uncertainty in the use of this calibration is so great that it would be well for convenience to place the limits at $-90°$ for the completely non-helical form and at $+10°$ for the complete helix so that one has a scale of 100 units. It must be kept in mind that this scale of 100 units is subject to being raised or lowered by as much as 20° and of being expanded or contracted by a similar amount as a result of the estimates we can now make of the variations in the sum of the intrinsic residue rotations in various proteins and of the solvent effects.

A similar scale for measuring helical content can be constructed for λ_c but it only runs as high as about 50% since higher helical contents should give rise to abnormal dispersion in which case λ_c is not defined. From the limited evidence thus far available the scale based on λ_c should be fairly independent of the protein composition, in contrast to the one based on $[\alpha]_D$, but it is subject to some solvent effect especially in some organic solvents. The lower limit for λ_c corresponding to the completely non-helical form appears to be located at about 212 mμ in aqueous solution and to shift upward somewhat in the case of more strongly interacting solvents. Within the limitations of the present data it appears to increase linearly with helical content reaching a value of about 282 mμ for 50% helix. Beyond this, curvature in the $\lambda^2[\alpha]$ against $[\alpha]$ plot should be observed and the estimate should be based on fitting the data to equation 1. The ratio of the value of b_0 obtained from such a fit divided by 600[9] would provide the estimate of the fraction of residues in helical form.

Because the scale based on λ_c does not appear to be compositional dependent whereas the one based on $[\alpha]_D$ is, albeit within rather narrow limits, it would appear that in general the λ_c scale offers an estimate of helical content with less uncertainty. However, if the $[\alpha]_D$ value for the completely denatured protein is known in a specific case, it would be preferable to locate the lower end of the scale of 100 units at this value. When this is possible the two methods of estimation should be equally reliable.

Finally, it should be noted that helical configurations having the opposite screw sense of the helix associated with L-residues in synthetic polypeptides would cancel out the positive configurational rotation of an equivalent amount of normal helix. Such a situation would lower both the apparent value of λ_c and the specific rotation, and the foregoing procedure would lead to an underestimate of the helical content. Although this feature may occur to a limited extent, it can be argued that it is not a common feature of protein structure and indeed does not occur to the extent of having equal amounts of the oppositely directed helices in any protein thus far studied, because its frequent occurrence would wipe out the correlations between λ_c and $[\alpha]_D$. Moreover, if the abnormal form of the helix were dominant in any protein, very low values of λ_c and $[\alpha]_D$ would be found for what would be clearly recognizable as a globular, undenatured protein: this has not occurred. A similar remark can be made concerning the occurrence of the β-form in solution. The rotatory behavior of this form, reported at the end of Part I, is such that the specific rotation would increase upon denaturation: this has not been observed in protein solutions.

III. Helix-coil Transitions in Proteins

In the study of polypeptides we have found that the configuration in solution depends primarily upon the hydrogen-bonding ability of the solvent used. One class of solvents, dichloroacetic and trifluoroacetic acids, hydrazine and ethylenediamine, solvate the peptide chain so as to prevent the helical configuration from existing. Another group of solvents, dimethylformamide, chloroform–formamide, ethylene dichloride and *m*-cresol, permit the helical configuration to exist. Water appears to occupy an intermediate position in this scale. Consequently, we have thought it possible to increase the amount of helical configuration in some proteins by turning to non-aqueous solvents. It was at once apparent, however, that solvents such as dimethylformamide, in which poly-γ-benzyl-L-glutamate, poly-L-glutamic acid and other polypeptides exist as helices, did not dissolve many proteins of interest. In this situation we have adopted an alternative of dissolving the protein in dichloroacetic acid and then adding as much as possible of a desired solvent such as ethylene dichloride. In this way we have been able to pass from the randomly coiled configuration to predominately helical configurations in some cases and in others to reach at least higher helical contents than those found under usual conditions in aqueous solutions. Results on four proteins or fragments thereof are presented here.

Silk Fibroin.—Silk fibroin is an obvious choice for this type of study because it consists of a single polypeptide chain, free of any cystine bridges that may prevent helix formation and containing as it does very few proline groups, which presumably disrupt a helical configuration. On the other hand, it has an untypical composition: 70% of its residues are glycine or alanine. The high crystal energy of the β-form prevents its dissolution in most solvents but it is readily soluble in dichloroacetic[30] and trifluoroacetic acid.[31] Upon dilution of a 2 to 3 g./dl. solution in dichloroacetic acid with ethylene dichloride, the specific rotation, $[\alpha]_D$, was found to increase from $-50°$ as shown in Fig. 7. Its value changes sign at 37% ethylene dichloride and reaches the unusually high value of $+45°$ be-

(30) E. J. Ambrose, C. H. Bamford, A. Elliott and W. E. Hanby, *Nature*, **167**, 264 (1951); B. A. Toms and A. Elliott, *ibid.*, **169**, 877 (1952).

(31) J. J. Katz, *ibid.*, **174**, 509 (1954).

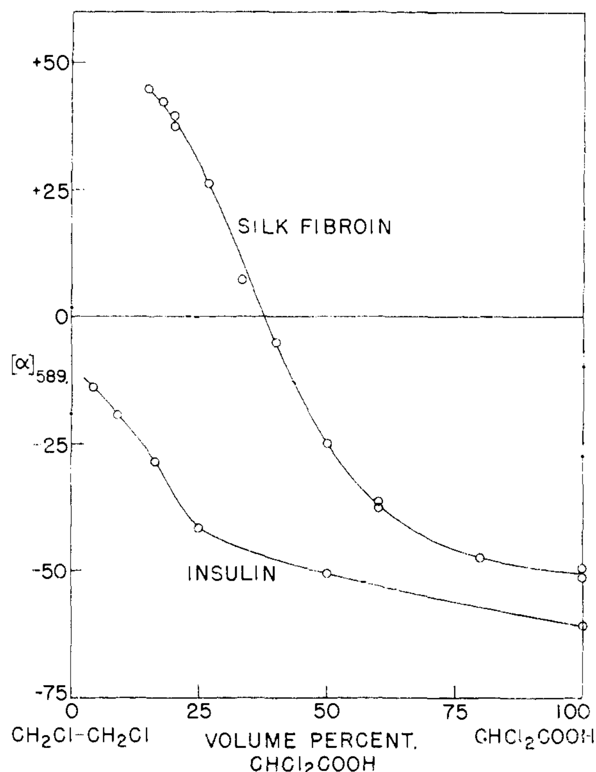

Fig. 7.—Specific rotation of silk fibroin and insulin as a function of solvent composition. The dashed line refers to poly-γ-benzyl-L-glutamate (M_w 22,000) data multiplied by 3.45.

fore precipitation occurs. The observed rotations were independent of time and rapidly reversible.

The high positive value of the specific rotation reached at 15% dichloroacetic acid is considerably higher than that expected for proteins even in the completely helical configuration. The explanation in this case follows quite naturally from the high glycine content of silk fibroin (44 mole %) and the large positive rotation which we have found the α-helix to confer on the intrinsic rotation of the polypeptide chain.[7] This configurational rotation was found to be +50° for poly-γ-benzyl-L-glutamate. Upon converting this to the basis of glycine residues (219/56) and multiplying by the weight fraction of glycine in silk fibroin (0.32) a value of 62° is obtained. Since the glycine residues are otherwise optically inactive, this value of 62° represents the contribution they would make to the specific rotation of silk fibroin in the right-handed α-helix configuration. Since the optically active, non-glycine residues have an $[\alpha]_D$ value of about 10° in the helical configuration it is expected that the $[\alpha]_D$ for silk fibroin in the helical configuration is indeed approximately +69°. This is very close to the value that would be obtained by extending the curve in Fig. 7 to pure ethylene dichloride and clearly accounts for the high positive values observed. Furthermore, the randomly coiled form of silk fibroin would be expected to have a relatively high rotation due to the glycine content: the estimated value of −68° (0.68 × −100) is within the allowed range of the observed value in dichloroacetic acid. However, it recently has been observed that poly-L-alanine and poly-L-leucine exist as helices in dichloroacetic acid, but as random coils in trifluoroacetic acid.[5] Hence the values of $[\alpha]_D$ and λ_c in dichloroacetic acid more likely reflect about 20% residual helix.

With this prediction of the attainment of high helical content in this system, it is of interest to see if there has been a conversion from simple to complex rotatory dispersion. The results are shown in Fig. 8 where it is seen that there is indeed a con-

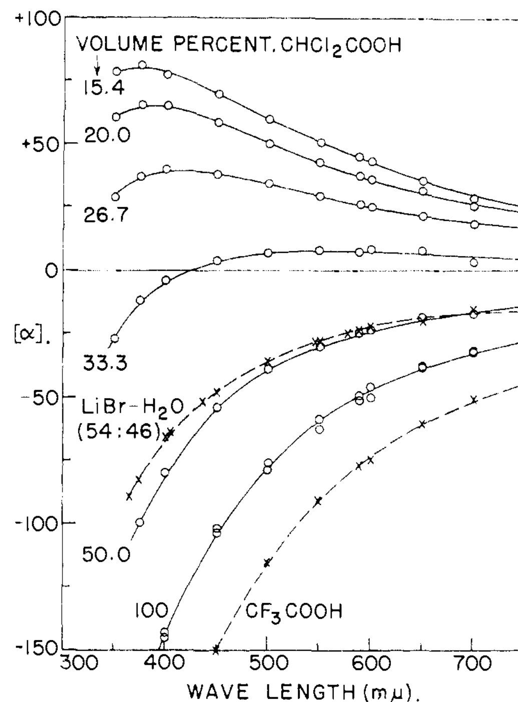

Fig. 8.—Rotatory dispersion of silk fibroin in various solvents.

tinuous change of the type expected upon passing from dichloroacetic acid to ethylene dichloride. The modified plot in Fig. 9 illustrates this more clearly. This plot shows that marked deviations from linear behavior (simple dispersion) occur at 35% dichloroacetic acid. At lower acid concentrations, equation 1 is found to fit the data with a λ_c value of 212 mμ, the same as was found for polypeptides. Moreover, the sign of b_0 is negative indicating that the helices also have the same screw sense as in the polypeptides. However, λ_c for the randomly coiled form is much larger than that previously observed and this is perhaps indicative of the large solvent effect to which this constant is subject in non-aqueous solvents.

The dispersion data for silk fibroin in concentrated lithium bromide solution (54 g. LiBr and 46 g. H_2O) is also included in Fig. 7. The dispersion in this case is simple. This together with the general location of the data in the figure indicates that the helical content is probably no greater than 40%. However, infrared studies[30] have been interpreted as showing that in films cast from such solutions or in much more concentrated solutions, the complete helical configuration is attained. If this is so, then it appears that there is further con-

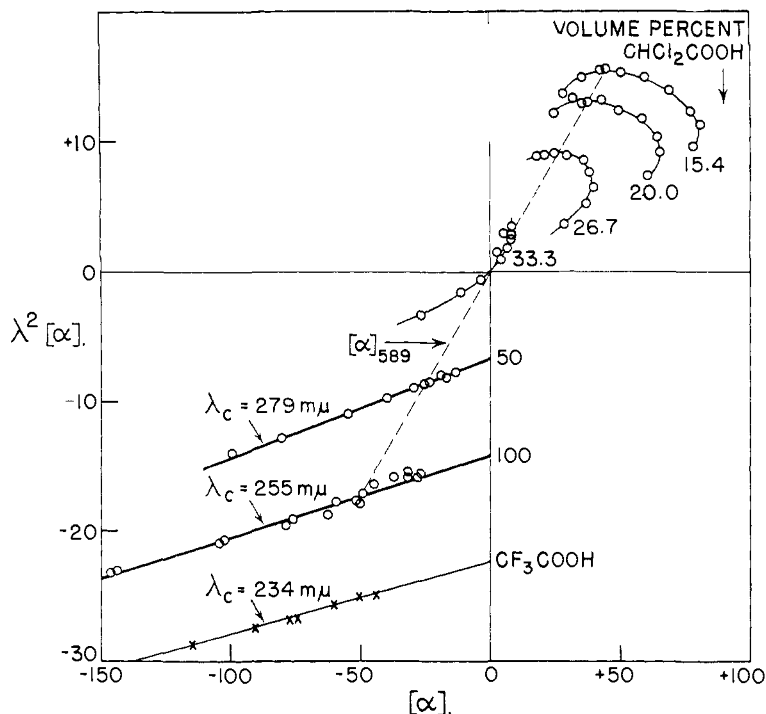

Fig. 9.—Modified dispersion plot for silk fibroin in ethylene dichloride–dichloroacetic acid mixtures.

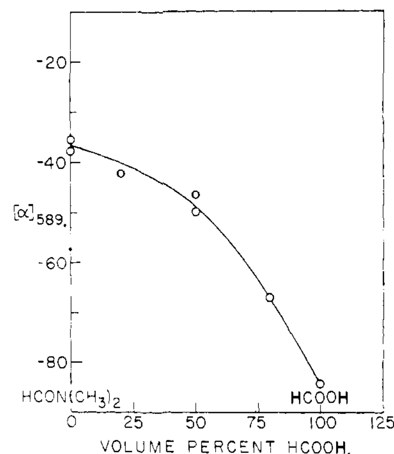

Fig. 11.—Specific rotation of insulin as a function of solvent composition.

version to the helical configuration as water is removed.

A fairly detailed study with several physical methods would be necessary to establish the macromolecular properties of the silk fibroin molecules in the ethylene dichloride–dichloroacetic acid system. The major difficulty arises because the few proline residues present would probably prevent the formation of a completely rod-like helical configuration which alone could be easily characterized. Such a study has not yet been carried out but it may be of interest to record here the way in which the intrinsic viscosity varies with solvent composition. In units of 100 cc./g. the values found were 2.7, 2.5, 2.4 and 2.1 in 0, 40, 60 and 80% ethylene dichloride. Evidently the molecular weight is such that there is not much dependence of viscosity on configuration. This same situation prevails in polypeptide systems in the vicinity of 50,000 molecular weight.

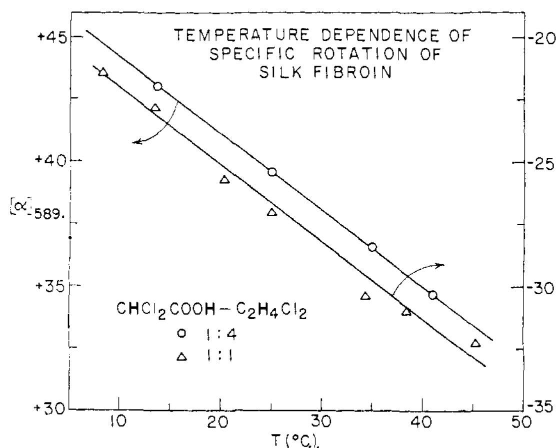

Fig. 10.—The temperature dependence of the specific rotation of silk fibroin in ethylene-dichloride mixtures.

In contrast to the helix-coil transitions observed in polypeptide systems as a function of solvent composition[6,32] that of silk fibroin (Fig. 7) is quite gradual. This point has been explored further by examining the effect of temperature on the transition.

The effect of temperature on the specific rotation of silk fibroin in 1:1 and 1:4 dichloroacetic acid–ethylene dichloride mixtures is shown in Fig. 10. In both cases there is a shift to the randomly coiled form upon raising the temperature. This is opposite to that found for poly-γ-benzyl-L-glutamate in the same solvent system, but it is in the direction normally encountered in protein denaturation.

The temperature study of the transition resembles the solvent composition dependence (Fig. 7) in that in both there appears to be only a small driving force. In Schellman's terminology[33,34] the heat of unfolding, calculated from the slope of the lines in Fig. 10, is only 5,500 cal. and the corresponding entropy of unfolding is 18 e.u. These values are an order of magnitude less than those found for similar transitions in poly-γ-benzyl-L-glutamate[6,32] or for those generally found in protein denaturation. Since it is fairly clear that nearly all the peptide residues in the silk fibroin molecules are participating in this transition, the interpretation of these results in terms of the transition involving only a small fraction of the residues

(32) J. T. Yang and P. Doty, manuscript in preparation.

(33) J. A. Schellman, *Compt. rend. trav. lab. Carlsberg, Ser. chim.*, **29**, No. 15 (1955).

(34) W. F. Harrington and J. A. Schellman, *ibid.*, **30**, No. 3 (1956).

must be rejected. Two possibilities then remain. (1) The mean heats and entropies of unfolding per residue may indeed be very small (approximately 5 cal. and 0.015 e.u.) as a result of the averaging out of much larger positive and negative values characteristic of individual residues, or (2) there may be a broad spectrum of local transition points throughout the silk fibroin chain so that formation or melting out of the helix as a function of solvent composition or temperature occurs in one small region after another, the regions presumably being differentiated by composition. At the moment a clear choice between these two possibilities can probably not be made. These observations do not mean, however, that this same transition cannot occur sharply, with denaturation-like features, in other solvent systems.

Insulin.—Since the chemical structure of insulin is completely known[35] the study of its configuration in solution is of particular interest. The results listed in Table II for insulin suggest that it possesses different helical contents in different solvents studied.[36] To investigate this a little further, we have taken the two solvents with the most different dispersion behavior, dimethylformamide and formic acid, and measured the rotation as a function of the relative amounts of these two solvents. The results are shown in Fig. 12. The observed changes

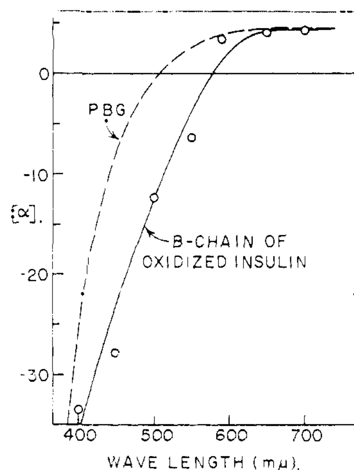

Fig. 12.—Rotatory dispersion of the B-chain of insulin in dimethyl formamide compared with poly-γ-benzyl-L-glutamate in the same solvent.

in rotation are completely and rapidly reversible. The intrinsic viscosity increases from 0.07 in dimethylformamide to 0.14 (100 cc./g.) in formic acid. Thus it appears that the regions in insulin which are helical when dissolved in dimethylformamide reversibly unfold, with a consequent swelling of the molecule, as formic acid is added. As with silk fibroin the transition is so gradual as to raise the same questions concerning the possible effects of heterogeneity of composition on the unfolding process. Of course, in this case the smaller chain lengths tend to broaden the transition.[33] Since the specific rotation in aqueous solutions from pH 4 to 10 is the same[37] as in dimethylformamide, it is probable that this same amount of helical configuration is present there as well. Indeed, the change observed by Golub and Pickett[37] in $[\alpha]_D$ from -36 to $-81°$ when the pH is increased from 10 to 12 appears to arise from the same transition we have observed upon passing from dimethylformamide to formic acid. Thus it is likely that the same transition can be observed in these three solvent systems. Using the values of $[\alpha]_D$ as a guide the region undergoing the helix-coil transition appears to represent about 45% of the residues. On the basis of the λ_c values in dimethylformamide and formic acid, however, the estimate would be only 15%. This unusually large discrepancy may possibly arise from the existence of helical regions having both screw senses.[38,39]

The rotatory behavior of insulin in the ethylene dichloride–dichloroacetic acid system is shown in Fig. 7. It appears that the same helix-coil transition is displayed here as in the solvent systems mentioned above. However, at the maximum ethylene dichloride content (91.7 volume %) the helical content appears to have increased more than in the other systems ($[\alpha]_D = -20°$). The dispersion of this solution showed some deviation from linearity in the modified plot. If this was ignored, a value of 300 mμ was obtained for λ_c. It appears that in this particular case a helical content of about two-thirds has been reached.

It may be that the further development of the helical configuration in native insulin is prevented by the three cystine bridges present. In order to test this possibility, we have prepared the B-chain by performic acid oxidation and fractionation.[35] The rotatory dispersion of this peptide fragment in dimethylformamide (Fig. 13) is that characteristic of the α-helix. For comparison, the specific rotation of poly-γ-benzyl-L-glutamate in the same solvent is shown as a dashed line. The close agreement between these two dispersion curves including the slightly positive values at high wave lengths

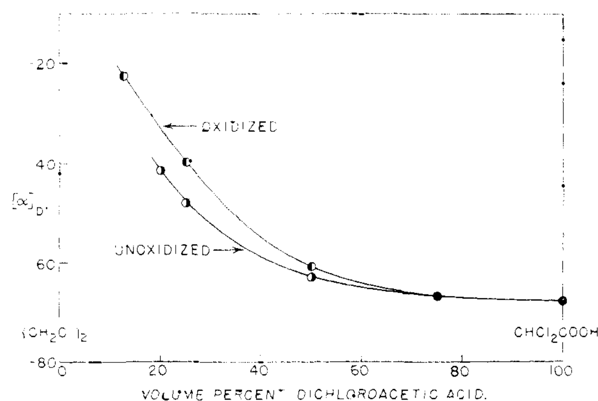

Fig. 13.—Specific rotation of native and oxidized ribonuclease as a function of solvent composition (ethylene dichloride–dichloroacetic acid).

(35) A. P. Ryle, F. Sanger, L. F. Smith and R. Kitai, *Biochem. J.*, **60**, 541 (1955); F. Sanger, *ibid.*, **44**, 126 (1949).

(36) We were aided at this point by the experiences of E. D. Rees and S. J. Singer, *Arch. Biochem. Biophys.*, **63**, 144 (1956).

(37) M. A. Golub and E. E. Pickett, *J. Polymer Sci.*, **13**, 427 (1954).

(38) H. Lindley and R. S. Rollett, *Biochim. Biophys. Acta*, **18**, 183 (1955).

(39) B. Low, "Currents in Biochemical Research 1956," ed. by D. E. Green, Interscience Publishers, Inc., New York, N. Y., p. 422.

demonstrates that the B-chain of insulin exists as an α-helix having the normal screw sense in dimethylformamide solution. This result must, however, be considered preliminary since we have not carried out either here or in the case of ribonuclease a characterization of the oxidized product employed. Hence the degree of purification achieved is uncertain but this does not affect the conclusion that the dispersion is indeed anomalous and closely approximates the type displayed by α-helices.

Ribonuclease.—Ribonuclease is another protein whose chemical structure is nearing complete elucidation.[40,41] It is known to consist of a single polypeptide chain with 4 cystine and 5 proline residues among the 130 in the chain. Here again it appears quite possible that the cystine cross-links are preventing the maximum development of the α-helix. To investigate this, ribonuclease was oxidized using the performic acid method at 0° to break the cystine bridges. The oxidized ribonuclease was insoluble in dimethylformamide (or ethylene dichloride) and consequently we proceeded as in the silk fibroin case to measure the specific rotation upon successive additions of ethylene dichloride to dichloroacetic acid solutions. The results are shown in Fig. 14 where it is seen that the specific rotation for both native and oxidized ribonuclease increases from $-67.5°$ in dichloroacetic acid to $-40°$ when three volumes of ethylene dichloride have been added. This similar behavior of the native and oxidized products, in sharp contrast to insulin, shows that the cystine bridges are not responsible for the low helical content of ribonuclease. Additional results supporting the same conclusion are included in Table II.

Bovine Serum Albumin.—Although the chemical structure of bovine serum albumin is not at all well known, we have made a preliminary study of its rotation in the native and oxidized forms identical to that just described for ribonuclease; the oxidation followed closely the procedure reported by Reichmann and Colvin.[42] The results are shown in Fig. 14 where it is seen that the rotation of the oxidized form is, in this case, about 10° higher over the accessible range of solvent composition. Although the effect is small it does appear that the development of the helix is somewhat aided by removing the cystine bridges in this case. Supporting evidence from λ_c values is given in Table II. The relatively high rotations reached at 10–15% ethylene dichloride for the oxidized material indicate that the helical content is sufficiently high for deviations from simple dispersion to be observed. Such deviations could be observed.

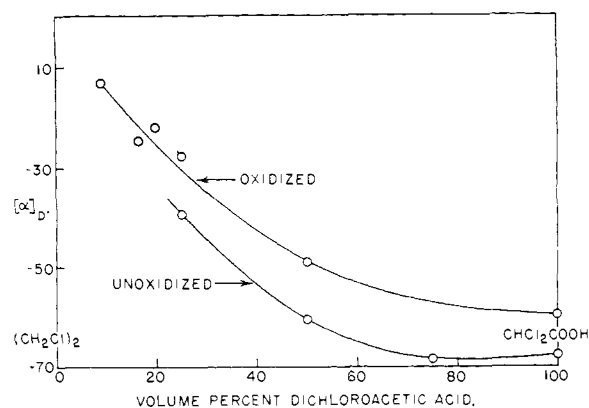

Fig. 14.—Specific rotation of native and oxidized bovine serum albumin as a function of solvent composition (ethylene dichloride–dichloroacetic acid).

Discussion

The principal conclusions reached in the foregoing presentation were based upon the premise that the rotatory dispersion of native and denatured proteins is effectively represented by the sum of the contributions from helical and non-helical regions and that the dispersion characteristics of these two forms closely resemble those observed in synthetic polypeptides. This simple hypothesis seems adequate for the interpretation of the data thus far available and it leads to the tentative conclusion that in solution, typical globular proteins have only 20–40% of amino acid residues in the helical configuration. Upon applying this interpretation to new dispersion data on proteins in non-aqueous solvents, it appeared that the helical configuration could be more fully developed than in aqueous solutions. Upon investigating this matter further, it was found that silk fibroin and the B chain of insulin could be converted almost entirely to the α-helix configuration having a single screw sense, presumably the right-handed one. The elimination of cystine bridges permitted a higher development of the helical configuration in insulin and to some extent in bovine serum albumin but had no effect in ribonuclease. The helix-coil transitions observed in these systems do not exhibit the sharpness usually associated with protein denaturation. A few remarks on these points may not be out of order.

In contrast to the relatively recent concern with the relation between the amount of helical configuration and λ_c and the nature of the rotatory dispersion, the possible relation between $[\alpha]_D$ and configuration has received attention for some time. The many investigations which have established the range of $[\alpha]_D$ values that occur for globular proteins and the extent of lowering that results from denaturation before 1950 have been briefly summarized.[43] In 1950–1952 the research group at Courtaulds Ltd. showed that synthetic polypeptides could display two configurations: the β-form in the solid state and a folded form which was identified later as being the helical form.[43] In an important note in 1951, Robinson and Bott[44] showed that, with increasing molecular weight, $[\alpha]_D$ for a series of copolymers of γ-methyl-L-glutamate and DL-phenylalanine increased from -18 to $+70°$. At about the same time, Becker and Stahmann[45]

(40) C. H. W. Hirs, William H. Stein and Stanford Moore, *J. Biol. Chem.*, **221**, 151 (1956).

(41) Christian B. Anfinsen, *ibid.*, **221**, 405 (1956).

(42) M. E. Reichmann and J. R. Colvin, *Can. J. Chem.*, **33**, 163 (1955).

(43) P. Doty and E. P. Geiduschek, "The Proteins," Chap. 5, Ed. H. Neurath and K. Bailey, Academic Press, New York, N. Y., 1953.

(44) C. Robinson and J. J. Bott, *Nature*, **168**, 325 (1951).

(45) R. R. Becker and M. A. Stahmann, THIS JOURNAL, **74**, 38 (1952).

noticed a similar increase with molecular weight in a series of poly-L-lysines: the change was from −80 to −37.6°. We note that the range covered here is similar to that now associated with the helix-coil transition and the positive values for the copolymer case find their natural explanation in the incorporation of racemic residues in a helix whose screw sense is dictated by the L-residues of γ-methyl-L-glutamate.[6] Also at this same time the careful work of Brand and his associates[46] established that the intrinsic residue rotation for several amino acids in non-terminal positions of a polypeptide chain in the randomly coiled configuration was indeed quite negative (*i.e.*, in the range of −60 to −100°).

The long recognized, qualitative relation between $[\alpha]_D$ and the degree of "native" or "uniquely organized" structure in globular proteins was considerably sharpened in 1955 by Cohen[24] who made the specific suggestion that the "rotation in the native state appears to be the result of effects of specific main-chain configuration superimposed on or otherwise changing the optical rotation from that of the independent, asymmetric carbon atoms alone." It was recognized by Cohen that the configurational effect on rotation could arise from either the persistent occurrence of a single type of main-chain configuration or from the net result of competing effects of enantiomorphic main-chain configurations. Somewhat similar views are implied in the work of Linderstrøm-Lang and Schellman.[21] This argument is carried further by the results presented here for it is our contention that the main-chain configuration responsible for changing the optical rotation is the α-helical configuration, that the magnitude of the rotational change so conferred is sufficiently large that only a minor fraction of the residues need be in this configuration to produce the observed effects and that this effect does not arise from enantiomorphic forms but rather from the existence of the α-helix in only one screw-sense (presumably right-handed). This last point rests upon (1) the compatibility of Moffitt's predictions for the dispersion of a right-handed helix with that observed for numerous polypeptides, (2) the constancy of specific rotations in homologous polypeptides of quite different molecular weight since this would not be expected if there was a small difference in stability of the two forms, (3) the optical rotation and rotatory dispersion of polypeptides containing D- and L-residues which will be reported shortly.[7] Of course, the occasional occurrence of minor regions of globular proteins in the helical form with the opposite screw sense may occur, with the result that the interpretation suggested here would underestimate the amount of helical configuration. However, there is no evidence thus far that requires the incorporation of this feature. Similarly, other helical configurations that have been suggested[47] may occasionally occur and be undetected by the procedures used here unless their values of b_0 and λ_c are quite different from that of the α-helix.

(46) E. Brand, B. F. Erlanger, J. Polatnick, H. Sachs and D. Kirschenbaum, THIS JOURNAL, **73**, 4025 (1951).

(47) See, *e.g.*, B. Low in "The Proteins," Vol. 1A, ed. by H. Neurath and K. Bailey, Academic Press, New York, N. Y., 1953.

The argument that the dispersion properties of the α-helix account for all the observations made thus far in optical rotation studies of proteins suggests the absence of other recurring structural features since these would generally be expected to make a rotational contribution. In particular it is of interest to note that the rotatory dispersion characteristic of the β-form (Fig. 4) has not been detected since its nature is such (increasingly positive values of $[\alpha]$ with decreasing λ) that substantial amounts would be noticed. This situation suggests the possibility that the non-helical regions of protein molecules are not arranged in any specific configuration but are subject to varying amounts of disorder depending upon the solvent with which they are in contact. Indeed, it is likely that the amount of helix content and the degree of order in non-helical regions may be greater in protein crystals than in aqueous solutions. Rotatory dispersion studies could possibly settle this point.

We turn finally to a brief comparison of our results with other observations on the helical content of proteins. One finds three different X-ray approaches[47] to this problem and one infrared study. All concur in finding evidence for the helical configuration in a number of globular proteins, but do not attempt a semi-quantitative estimate of the amount present. Of these the least ambiguous is the striking X-ray evidence of the α-helix in the 1.5 Å. diffraction maximum deduced by Perutz.[48] He found this spacing in α-keratin, muscle and hemoglobin. It also has been found in epidermin, tropomyosin, fibrin, and bacterial flagellae. The only point of overlap between this group and the proteins we have studied is with fibrinogen and fibrin. The amount estimated from dispersion studies (about 35%) is clearly enough to account for the X-ray observation.

The second X-ray approach has been in the attempt to interpret three-dimensional Patterson diagrams in terms of rod-like regions having the α-helix configuration.[47] This appears not to have led to the clear identification of the α-helix as a major component of crystalline proteins as was once expected. In two of the best studied cases, hemoglobin[49,50] and insulin,[51] the amount of helical content compatible with the Patterson diagrams is considered to be less than 50%. This likewise is consistent with the conclusions from dispersion studies.

The third approach has been the work of Arndt and Riley[52] on the low angle powder diffraction of polypeptides and proteins. They interpreted their work as showing the α-helix is predominant in bovine serum albumin, ribonuclease and a number of other proteins. However, with the contribution of the non-helical regions unspecified it appears to us almost impossible to make even a qualitative estimate of the amount of helical configuration present. Moreover, the interpretations of this work favor a left-handed helix in disagreement with the

(48) M. F. Perutz, *Nature*, **167**, 1053 (1951).

(49) W. L. Bragg, E. R. Howells and M. F. Perutz, *Acta Cryst.*, **5**, 136 (1952).

(50) F. H. C. Crick, *ibid.*, **5**, 381 (1952).

(51) B. Low, "McCollum-Pratt Symposium," Baltimore, Md., June 15-19, 1956, proceedings to be published.

(52) U. V. Arndt and D. P. Riley, *Phil. Trans. Roy. Soc.* (*London*), **A247**, 409 (1955).

theoretical interpretation of our dispersion work. Consequently, our observations are for the most part incompatible with those of Arndt and Riley.

Infrared spectra of crystalline insulin and other proteins[53] and of bovine serum albumin in 10% D_2O solution[54] have revealed clearly defined peaks at about 1655 cm.$^{-1}$ which are generally associated with the amide I band of the α-helical configuration. Again no estimate of the amount of helical configuration was attempted but the evidence is not incompatible with the results presented here.

We come finally to the recent evidence on the extent of intramolecular hydrogen bonding resulting from the measurements of deuterium exchange emanating from the Carlsberg Laboratory.[55-57] In these studies it has been found that all exchangeable hydrogen (or deuterium) atoms in small peptides, the A-chain of insulin and oxidized ribonuclease undergo exchange very rapidly in aqueous solution. However, for the two proteins studied, insulin and ribonuclease, a portion of the exchangeable hydrogen (33 and 21%, respectively) exchanges at a relatively slow rate.

The suggestion has been made[55] that the majority of the strongly bound hydrogen atoms are imide–hydrogen atoms in the backbone of the peptide chains that are folded into the helical structure. A comparison of this work with our estimates of the helical content is possible if the further assumption is made: that *all* the strongly bound hydrogen atoms are imide–hydrogen atoms.

In the case of insulin it is found that 23 hydrogen atoms exchange slowly and 7 more at an intermediate rate. Since there are 48 imide–hydrogen atoms this result leads to an estimate of 48 or 62% helix on the basis of the above assumption. From dispersion measurements on insulin under the same conditions[21] (λ_c = 265 mμ and $[\alpha]_D$ = $-31.7°$) the helical content is estimated at 38 and 35%[58] from the two constants, respectively. If the lower figure from the exchange data is taken, 48%, the discrepancy with our estimate is within the expected uncertainty. The higher figure, 62%, is however sufficiently different to reflect on the validity of the assumption which identified all the strongly bound hydrogen atoms with imide groups.

With ribonuclease about 50 slowly exchangeable hydrogen atoms are found. With 122 imide–hydrogen atoms in the molecule this indicates 41% helical configuration. However, using the method described here, we find from λ_c a value of 14% and from $[\alpha]_D$ a value of 20%.[59] Thus the values derived from the rotatory dispersion are only half those deduced from exchange data.

The probable explanation of the consistent difference between these two methods of estimating the amount of helical configuration probably lies in the existence of types of hydrogen bonding other than that in the α-helix. If this explanation is supported by further work, the two methods would be complementary and their joint application would permit an assessment of the numbers of hydrogen bonds in the helical configuration and in other structural arrangements.

This brings to an end a summary of other closely relevant studies, all of which indicate either directly or indirectly the existence of the α-helical configuration in proteins. The principal result of the work presented here has been to add to these evidences of the occurrence of the α-helical configuration in proteins the further evidence offered by rotatory dispersion studies. The interpretation of this newer evidence, based on the study of model polypeptide systems, indicates that it can provide a semi-quantitative estimate of the fraction of residues in the α-helical configuration for proteins in solution.

Acknowledgment.—We wish to thank Mrs. Elizabeth Klemperer for help in much of the experimental work and Professors J. A. Schellman (University of Minnesota) and S. J. Singer (Yale University) for the opportunity of reading manuscripts of theirs prior to publication. This work was made possible by support from the Milton Fund, Harvard University, and the Office of Naval Research (Nori-07654).

CAMBRIDGE, MASS.

(53) E. J. Ambrose and A. Elliott, *Proc. Roy. Soc.* (*London*), **A208**, 75 (1951).

(54) H. Lenormant and E. R. Blout, *Nature*, **172**, 770 (1953).

(55) K. Linderstrøm-Lang in "Peptide Chemistry," Report of a Symposium held by the Chemical Society, March 30, 1955, The Chemical Society, London, 1955.

(56) A. Hvidt and K. Linderstrøm-Lang, *Compt. rend. trav. lab. Carlsberg, Ser. chim.*, **29**, 385 (1955).

(57) C. N. Anfinsen, W. F. Harrington, A. Hvidt, K. Linderstrøm-Lang, M. Ottensen and J. Schellman, *Biochim. et Biophys. Acta*, **17**, 141 (1955).

(58) The value from λ_c is found as follows: $(265 - 212) \times (5/7) = 38\%$. The value based on $[\alpha]_D$ was found by using the $[\alpha]_D$ value for the oxidized chain of insulin (reference 21), $-66.7°$, as the lower limit instead of the value of $-90°$ which is used when such a scale localization as this is not possible. Consequently $(66.7 - 31.7) = 35\%$.

(59) Here also we have the means of locating the zero of the scale since an accurate estimate of the specific rotation of completely denatured ribonuclease is available (reference 34). In this case, however, this leads to a negligible shift of the scale since the value for denatured ribonuclease, $-92°$, is so close to the average value, $-90°$: $(92 - 71.5) = 20\%$.