Reexamination of the Conformation of Muscle Proteins by Optical Activity†

Cheun-Shang C. Wu and Jen Tsi Yang*

ABSTRACT: The circular dichroism and optical rotatory dispersion of muscle proteins are reexamined. By the method of Chen et al. (Chen, Y. H., Yang, J. T., and Chau, K. H. (1974), *Biochemistry 13,* 3350), the estimated helical contents of myosin (78%), heavy meromyosin (HMM) (70%), subfragment 1 (SF-1) (60%), and G-actin (45%) are higher than hitherto reported. Tropomyosin (TM) and light meromyosin fraction I (LMM Fr. I) possess more than 90% helix in agreement with the values based on the b_0 method. HMM, SF-1, and G-actin also contain about 8, 16, and 27% β form. The three troponins (TN) and three light chains (LC) of myosin have moderate amounts of helices (29 to 51%) and some β form (13 to 23%). If the light chains are intact in HMM and SF-1, myosin would have 3–5% β form, which is difficult to detect with the present method. For comparison, the predictive method based on amino acid sequence gives similar estimates for TM, G-actin, and TN-C with bound calcium, but slightly higher helical contents than our results for TN-I and the light chains.

Circular dichroism (CD[1]) and optical rotatory dispersion (ORD) have been widely used for the studies of conformation and conformational changes of proteins, but the early methods of analysis based on the Moffitt equation, the mean residue rotation at 233 nm, and mean residue ellipticity at 222 nm ignore the rotatory contributions of the β conformation, which is equally abundant as the helices in most proteins, as the x-ray diffraction studies have now revealed (cf. Blake et al., 1967; Blow, 1969; Wyckoff et al., 1970; Becker et al., 1975). Recent attempts have been focused on fitting the ORD and CD spectra of proteins by those constructed with mixtures of helix, β, and unordered forms. The assumptions are that proteins contain only three conformations and their optical activities are additive. Either the results of synthetic polypeptides as model compounds are used for such curve fitting (Greenfield et al., 1967; Greenfield and Fasman, 1969; Straus et al., 1969) or the reference values of the three conformations are computed from the CD spectra of proteins of known structures (Saxena and Wetlaufer, 1971; Chen et al., 1972). Chen et al. (1974) have also considered the chain length dependence of the CD and ORD of helices, although the determination of the average number of residues per helical segment and helical segments in each molecule is not yet perfected.

The helical contents of muscle proteins range from low to moderate for globular proteins such as SF-1 (Murphy, 1974), G-actin (Kay, 1960; Nagy and Jencks, 1962; Nagy, 1966; Murphy, 1971; Nagy and Strzelecka-Golaszewska, 1972), myosin light chains (Lowey and Holt, 1972), and troponins (Staprans and Watanabe, 1970; Murray and Kay, 1972; Mani et al., 1973, 1974b), to moderately high for myosin and HMM, and almost completely helical for rodlike proteins such as TM and LMM Fr. I (Cohen and Szent-Gyorgyi, 1957; Lowey and Cohen, 1972; Oikawa et al., 1968; Staprans and Watanabe, 1970; Murphy, 1974). Most of these estimates were based on the b_0 values of the Moffitt equation. Only limited CD studies are available. With improved instrumentation as well as methods of protein purification, we therefore reexamined the CD and ORD of muscle proteins. The recent methods of analysis (Chen et al., 1974) enable us to estimate both the helical and β contents of these muscle proteins.

Experimental Section

Preparation of Proteins. All proteins used were prepared from the back muscle of rabbit at 4 °C, unless stated otherwise. Myosin was extracted by the procedure of Wu and Sayre (1971) and purified on a DEAE-Sephadex A-50 column (Richards et al., 1967). HMM and LMM Fr. I were prepared by the method of Lowey and Cohen (1962) with the following modifications. After adding 2 mg of trypsin inhibitor per mg of trypsin to stop the digestion of myosin, the protein solution was diluted with 10 times its volume of cold distilled water to precipitate LMM and unreacted myosin. After centrifugation, HMM in the supernatant was salted out at 3 °C with 40–60% saturated ammonium sulfate, redissolved, and dialyzed against 40 mM sodium pyrophosphate (pH 7.5). It was passed through a Sephadex G-200 column using the same buffer as the eluent. LMM and unreacted myosin in the precipitate were redissolved in 0.5 M KCl–0.05 M potassium phosphate (pH 6.2) and digested another 10 min with $\frac{1}{200}$ their weight of trypsin at 23 °C. The reaction was stopped with two times excess trypsin inhibitor. LMM Fr. I was precipitated with ethanol and redissolved in buffer (Szent-Gyorgyi et al., 1960), and passed through a Sephadex G-200 column with 40 mM sodium pyrophosphate. SF-1 was prepared from myofibrils and purified on a Sephadex G-200 column (Cooke, 1972).

G-actin was prepared according to Spudich and Watt (1971). Frozen TM, a gift of Professor S. Watanabe, was purified on a Sephadex G-200 column (Azuma and Watanabe, 1965).

Troponin was extracted from the muscle and purified by the salting-out method of Staprans et al. (1972). The three com-

† From the Cardiovascular Research Institute and Department of Biochemistry and Biophysics, University of California, San Francisco, California 94143. *Received November 11, 1975.* This work was aided by U.S. Public Health Service Grants GM-10880 and HL-06285.

[1] Abbreviations used are: CD, circular dichroism; ORD, optical rotatory dispersion; HMM, heavy meromyosin; SF-1, subfragment 1; LMM Fr. I, light meromyosin fraction I; TM, tropomyosin; TN-C, TN-I, and TN-T, troponin C, I, and T; LC_1, LC_2, and LC_3, light chains 1, 2, and 3 of myosin in the order of decreasing molecular weight; Nbs_2, 5,5′-dithiobis(2-nitrobenzoic acid); EGTA, ethylene glycol bis(β-aminoethyl ether)-*N,N′*-tetraacetic acid.

ponents TN-C, TN-I, and TN-T were then fractionated on a DEAE-cellulose column with 6 M urea, 20 mM Tris-HCl, 1 mM EDTA, or EGTA at pH 8.0 and a linear gradient of KCl up to 0.4 M. Urea was dialyzed out immediately after fractionation.

The light chains of myosin were dissociated from myosin with 4 M urea (Weeds and Lowey, 1971). The Nbs_2 light chain (LC_2), however, was separated by Nbs_2 treatment prior to urea addition. The light chains were fractionated on a DEAE-cellulose column containing 20 mM Tris-HCl and 0.2 mM dithioerythritol (pH 8.0) with a linear gradient of KCl up to 0.3 M. (We designated the subscripts 1, 2, and 3 in the order of decreasing molecular weight.)

Homogeneity of Preparations. The purities of all proteins except the trypsin-digested myosin (HMM, LMM Fr. I, and SF-1) were further tested with sodium dodecyl sulfate disc gel electrophoresis (Weber and Osborn, 1969). (Bovine serum albumin, pepsin, trypsin inhibitor, myoglobin, and insulin were used as molecular weight markers.) Myosin which dissociated into two subunits showed one thick band at 200 000 daltons (heavy chain) and three bands at 26 000, 18 000, and 16 000 daltons (light chains). Two or three very thin bands immediately ahead of the heavy chain could not be removed by rechromatography. They were as yet unidentified and one of them was probably the C-protein (Offer et al., 1973). Judged from the intensity of these bands, our myosin preparation may contain a maximum impurity of about 5%. G-actin and the two subunits of TM showed a single band at 45 000 and 35 000 daltons, respectively. The purity of HMM, LMM Fr. I, and SF-1 was determined chromatographically (trypsin digestion is expected to split a few additional susceptible bonds and show a complicated gel electrophoresis pattern) (Weeds and Lowey, 1971). However, our CD analysis (see Results) indicated that such nonhomogeneity, if any, was insignificant.

Troponins and especially the light chains of myosin are known to be more difficult to purify than other muscle proteins. TN-I (23 000 daltons) that emerged before the salt gradient was applied was often contaminated with faint bands of proteins of smaller molecular weights and sometimes TM-like protein. TN-T (37 000 daltons) was occasionally mixed with a small amount of TN-C (17 000 daltons). All these samples were discarded and only preparations that showed a single band were used.

The light chains of myosin that were well resolved by chromatography often showed more than a single band by gel electrophoresis, probably because they were easily aggregated (Lowey and Holt, 1972). Again these preparations were discarded and new batches prepared until homogeneous samples were obtained.

Preparation of Solution. Except TM, all proteins were freshly prepared and used within 2 weeks. The preparations of myosin, LMM Fr. I, TM, and troponins were dialyzed against 40 mM sodium pyrophosphate (pH 7.5) (the troponin solutions also contained 0.5 mM EGTA). KCl of high ionic strength was not used because of the absorbance of chloride ions below 200 nm, which interferes with optical measurements. Sodium phosphate (10 mM, pH 7.0) or 40 mM sodium pyrophosphate was used for HMM (we observed no difference in CD spectra in the two buffers). Sodium phosphate (10 mM) was also used for SF-1 and the three light chains. The G-actin solution was dialyzed exhaustively against 2 mM sodium phosphate (pH 7.0) to remove free ATP. All protein solutions were filtered through a 3 μm Millipore filter before use.

Determination of Concentration. All protein concentrations were determined by the micro-Kjeldahl method. The nitrogen contents and the number of amino acid residues (N) were either taken from the literature or calculated from amino acid compositions: myosin, 17.1% (Chung et al., 1967), 4060 (Lowey and Cohen, 1962); HMM, 16.7% (Kominz et al., 1954), 2700 (Lowey and Cohen, 1962); LMM Fr. I, 16.7% (Kominz et al., 1954), 1010 (Lowey and Cohen, 1962); SF-1, 16.5%, 1040 (Young et al., 1965); TM, 16.6%, 568 (Stone et al., 1974); G-actin, 16.2%, 374 (Elzinga et al., 1973); TN-C, 15.5%, 158 (Collins et al., 1973); TN-I, 18.3%, 179 (Wilkinson and Grand, 1975); TN-T, 17.9%, 312 (Wilkinson, 1974); LC_1, 16.4%, 190 (Frank and Weeds, 1974); LC_2, 16.5%, 169 (Collins, 1976); and LC_3, 16.2%, 149 (Frank and Weeds, 1974).

CD and ORD Measurements. CD spectra were measured on a JASCO J-10 spectropolarimeter and ORD spectra on a Cary 60 spectropolarimeter both under nitrogen flush at 25 °C. Protein concentrations used were 0.2–0.4 mg/ml for measurements in the ultraviolet region with cells of 0.1, 0.5, and 2 mm path lengths and 3–6 mg/ml in the near ultraviolet and visible region with 5-, 10-, and 20-mm cells. The cells were calibrated by the known rotations of the National Bureau of Standards sucrose on Cary 60, which had been calibrated according to the procedure of Samejima and Yang (1964). The JASCO J-10 had been calibrated with *d*,10-camphorsulfonic acid by the method of Cassim and Yang (1969). The spectra of each protein were determined at least three times with two or more preparations and averaged. The reproducibility was within 3% for spectra above 205 nm and about 7% below 205 nm. The data were expressed in terms of mean residue ellipticity, $[\theta]$, and mean residue rotation, $[m]$.

Method of Analysis. We computed the percents of helix and β form from the CD and ORD data by the method of Chen et al. (1974). At any chosen wavelength, λ, the experimental CD can be represented as:

$$[\theta] = f_H[\theta]_H^{\infty}(1 - k/\bar{n}) + f_\beta[\theta]_\beta + f_R[\theta]_R \qquad (1)$$

Here the subscripts H, β, and R refer to the helix, β, and unordered forms. The f's are the fractions of the three conformations; the sum of three f's equals one and all f's ≥ 0. $[\theta]_H^{\infty}$ is the reference value for a helix of infinite length, $\bar{n}$ is the average length of the helical segments, and k is a wavelength-dependent constant. $\bar{n}$ is related to the number of helical segments, i, by $i = f_H N/\bar{n}$ (N = total number of amino acid residues). Replacing $[\theta]_H^{\infty}$ by three Gaussian bands for the $n\text{-}\pi^*$, $\pi\text{-}\pi_\parallel^*$, and $\pi\text{-}\pi_\perp^*$ transitions, eq 1 becomes (Chen et al., 1974):

$$\begin{aligned}[\theta] = f_H\{&-37\,300(1 - 2.50/\bar{n})\exp[-(\lambda - 223.4)^2/\Delta_{n\text{-}\pi^*}^2] \\ &- 37\,200(1 - 3.50/\bar{n})\exp[-(\lambda - 206.6)^2/\Delta_{\pi\text{-}\pi_\parallel^*}^2] \\ &+ 101\,000(1 - 2.50/\bar{n})\exp[-(\lambda - 193.5)^2/\Delta_{\pi\text{-}\pi_\perp^*}^2]\} \\ &+ f_\beta[\theta]_\beta + (1 - f_H - f_\beta)[\theta]_R \qquad (2)\end{aligned}$$

where the Δ's are the half-bandwidths in nanometers.

Since the rotational strength, R_j, of a jth CD band is:

$$R_j = 1.23 \times 10^{-42}[\theta_j^0]\Delta_j/\lambda_j \qquad (3)$$

where λ_j is the wavelength at the extremum $[\theta_j^0]$. Equation 2 then becomes (see Chen et al. (1974) for the R_j values based on the CD spectrum of myoglobin):

$$\begin{aligned}[\theta] = f_H\{&-1800(223.4/\Delta_{n\text{-}\pi^*})(1 - 2.50/\bar{n}) \\ &\times \exp[-(\lambda - 223.4)^2/\Delta_{n\text{-}\pi^*}^2] - 1600(206.6/\Delta_{\pi\text{-}\pi_\parallel^*}) \\ &\times (1 - 3.50/\bar{n})\exp[-(\lambda - 206.6)^2/\Delta_{\pi\text{-}\pi_\parallel^*}^2] \\ &+ 4370(193.5/\Delta_{\pi\text{-}\pi_\perp^*})(1 - 2.50/\bar{n}) \\ &\times \exp[-(\lambda - 193.5)^2/\Delta_{\pi\text{-}\pi_\perp^*}^2]\} + f_\beta[\theta]_\beta \\ &+ (1 - f_H - f_\beta)[\theta]_R \qquad (4)\end{aligned}$$

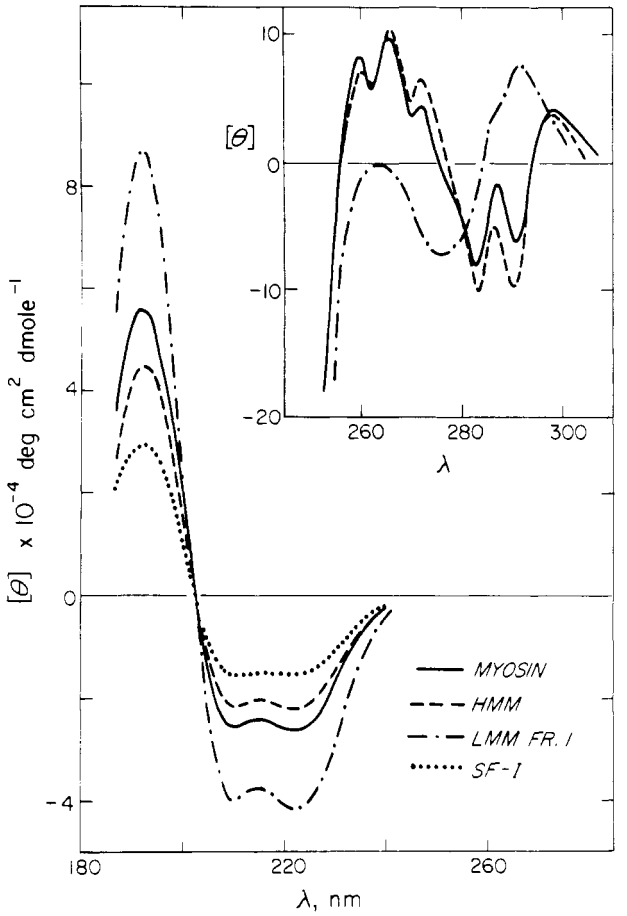

FIGURE 1: CD spectra of myosin, HMM, LMM Fr. I, and SF-1 at 25 °C. Solvent: 40 mM sodium pyrophosphate (pH 7.5) for myosin and LMM Fr. I; 10 mM sodium phosphate (pH 7.0) for HMM and SF-1.

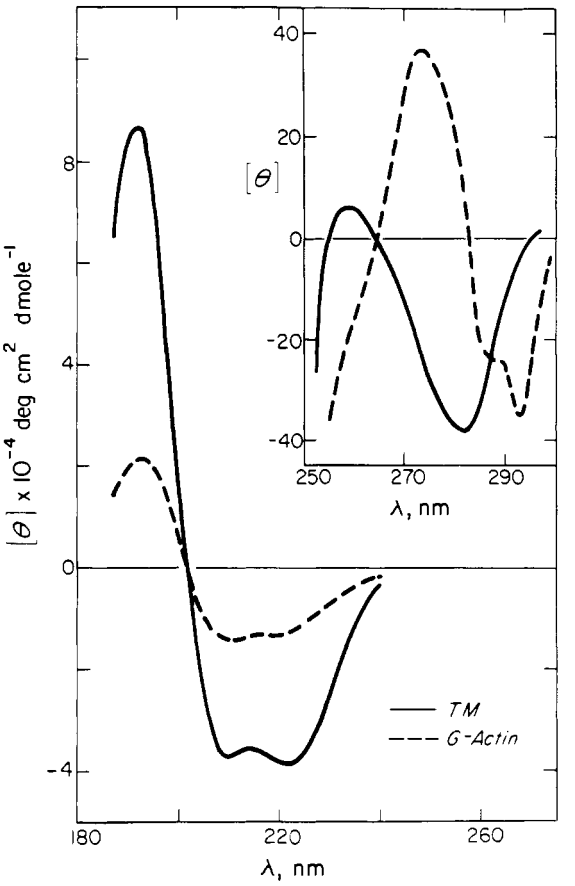

FIGURE 2: CD spectra of G-actin and tropomyosin at 25 °C. Solvent: 2 mM sodium phosphate (pH 7.0) for G-actin; 40 mM sodium pyrophosphate (pH 7.5) for tropomyosin.

Both eq 2 and 4 have six parameters ($f_H, f_\beta, \bar{n}$, and 3 Δ's). Chen et al. (1974) proposed to eliminate $\bar{n}$ from eq 2 by combining it with the mean residue rotation at 233 nm:

$$[m]_{233} \simeq [-18\,700(1 - 3.10/\bar{n}) + 1820]f_H - 1820 \quad (5)$$

(method 1). Likewise, we can combine eq 4 with eq 5 to eliminate $\bar{n}$. Next, we solve the five parameters (f_H, f_β, and 3 Δ's) with more than 50 data points (the $[\theta]$'s at 1-nm intervals) by the nonlinear least-squares method, using a BMD07RT, UCLA computer program. (Chen et al. termed the solution of eq 4 with six parameters as method 2. The solution of combined eq 4 and 5 (method 3) was inadvertently termed as a combination of methods 1 and 2. J. T. Yang was responsible for this mistake.)

Actually, the combined eq 2 and 5 is a *constant θ and variable Δ* method and the combined eq 4 and 5 a *constant $\theta\Delta$ (or constant R)* method. A third possibility is a *constant Δ and variable θ (or R)* method. In this case eq 2 can be rewritten as (see Chen et al. (1974) for the values of Δ's based on the CD spectrum of myoglobin):

$$\begin{aligned}[\theta] = {} & f_H\{[\theta^0]_{n\text{-}\pi^*}^{\infty}(1 - 2.50/\bar{n}) \\ & \times \exp[-(\lambda - 223.4)^2/(10.8)^2] + [\theta^0]_{\pi\text{-}\pi\|^*}^{\infty}(1 - 3.50/\bar{n}) \\ & \times \exp[-(\lambda - 206.6)^2/(8.9)^2] + [\theta^0]_{\pi\text{-}\pi_\perp^*}^{\infty}(1 - 2.50/\bar{n}) \\ & \times \exp[-(\lambda - 193.5)^2/(8.4)^2]\} \\ & + f_\beta[\theta]_\beta + (1 - f_H - f_\beta)[\theta]_R \quad (6)\end{aligned}$$

Equation 6 can be combined with eq 5 to reduce one parameter. There are several other alternatives. All are devised as a compromise. They usually give good results of f_H and f_β, but the estimated $\bar{n}$ or i is often not satisfactory. There is also no a priori reason that the three λ_j's of the CD bands of a helix must be prefixed. Whether small variations in λ_j's would improve the computed results remains to be investigated.

Results

Ultraviolet CD Spectra. Figures 1 and 2 show the CD of myosin, HMM, LMM Fr. I, G-actin, and TM. All spectra have a double minimum at 222 and 210 nm, except that G-actin is blue-shifted from 222 to 219 nm, and a maximum at 192 nm. These results suggest the presence of helical conformation in the six proteins. G-actin differs from the other five proteins in that the ratio of $[\theta]_{219}/[\theta]_{210}$ is smaller than one, whereas the ratio of $[\theta]_{222}/[\theta]_{210}$ is larger than one for the others. This is probably due to a relatively large amount of the β form in G-actin. Removal of free ATP from G-actin solution did not affect its CD and ORD. The magnitudes of the CD extrema of TM are close to those of helical synthetic polypeptides, which can vary slightly with the nature of the side chains. The magnitudes of LMM Fr. I are actually larger than those of helical poly(L-glutamic acid) (Cassim and Yang, 1970) or the $[\theta]_H^{\infty}$'s computed from proteins of known structure (Chen et al., 1974). Both TM and LMM Fr. I are known to form a two-stranded coiled-coil structure (Woods, 1967). Whether this superhelix would enhance the CD of a single-stranded helix is difficult to ascertain at present.

Figures 3 and 4 show the CD spectra of troponins and light chains of myosin. These six proteins have relatively smaller extrema and thereby lower helical contents than those shown

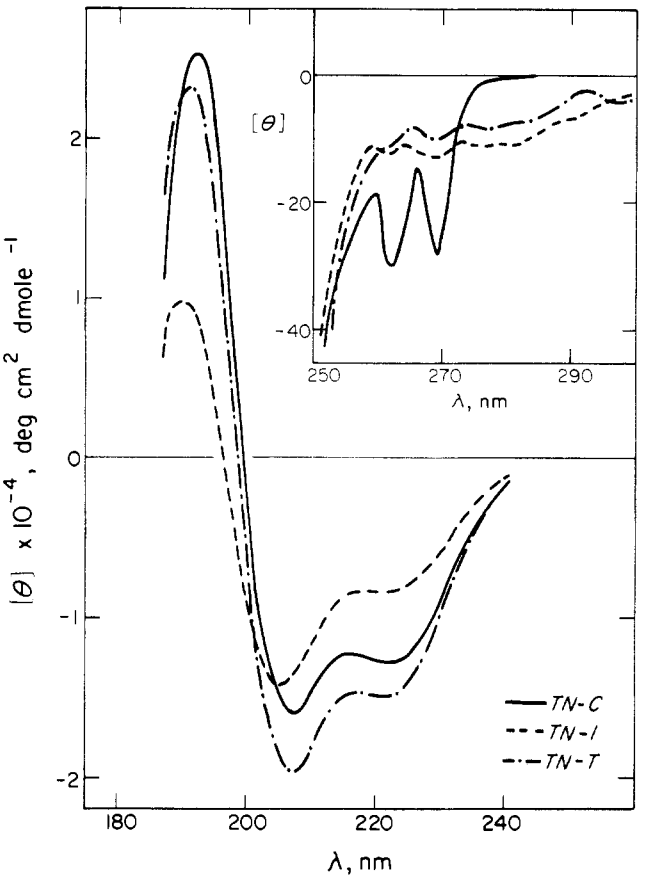

FIGURE 3: CD spectra of troponins at 25 °C. Solvent: 40 mM sodium pyrophosphate and 0.5 mM EGTA (pH 7.5).

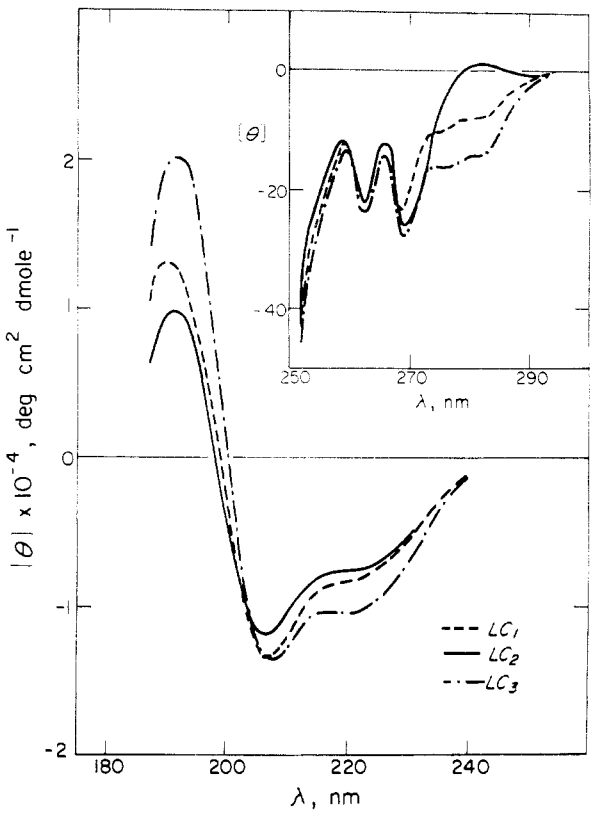

FIGURE 4: CD spectra of light chains of myosin at 25 °C. Solvent: 10 mM sodium phosphate (pH 7.0).

in Figures 1 and 2. In all cases, the π–$\pi_{\parallel}$* transition band is blue shifted by 2 to 5 nm from 210 nm and has a higher intensity than the n–π* transition band. For TN-I and LC_1, the 192-nm band is also blue shifted to 190 nm. As the intensity of the spectrum decreases, the crossover point also tends to move toward the shorter wavelength.

Table I lists the extrema of the CD and ORD of the 12 proteins. Our CD results of G-actin resemble those of Murphy (1971), but the magnitudes of the extrema are about 30% larger than those reported by Nagy and Strzelecka-Golaszewska (1972). Our [θ]'s of myosin and HMM are about 10–15% larger (in magnitude) than those of Oikawa et al. (1968), but again are comparable to Murphy's values (1974). The CD values of SF-1, however, are about 30% higher than those of Murphy (1974). Our ORD for myosin, HMM, and LMM Fr. I are close to those of McCubbin et al. (1966). Our CD of TM and LMM Fr. I agree with those reported by Oikawa et al. (1968) and Staprans and Watanabe (1970). Our CD extrema of TN-C and TN-T are higher than those reported by Kay and co-workers but those for TN-I are slightly lower (Murray and Kay, 1972; Mani et al., 1973, 1974b). Our CD extrema for the light chains of myosin are comparable to those of Lowey and Holt (1972). In all cases we have extended the CD spectra down to 187 nm. The data at these low wavelengths are necessary for an accurate determination of the percents of helix and β form. No ORD of troponins and light chains of myosin are available in the literature.

Near-Ultraviolet CD Spectra. Multiple Cotton effects in the near-ultraviolet region represent the CD of nonpeptide chromophores (Figures 1–4, insets). In all cases, their intensities are extremely small as compared with those of the peptide chromophores and therefore should not affect our method of analysis significantly. The close similarities in profile, and even magnitude, of CD between myosin, HMM, and SF-1 (not shown) (Figure 1, inset) suggest the predominant rotatory contribution of HMM and thereby SF-1 in the parent molecule, which overshadows the CD spectrum of LMM Fr. I that is quite different. The negative 283-nm band of TM (Figure 2, inset) is probably due to the tyrosine residues (the CD of phenylalanine chromophore is weak), since TM lacks tryptophan and cystine residues (Kominz et al., 1954; Woods, 1967).

TN-C and the light chains are known as homologous proteins of parvalbumin that bind calcium ions (Collins et al., 1973; Weeds and McLachlan, 1974; Tufty and Kretsinger, 1975; Collins, 1976). Their near-ultraviolet CD spectra are strikingly similar (Figures 3 and 4, insets). The 263- and 269-nm bands were also found for carp parvalbumin (C. S. C. Wu and J. T. Yang, unpublished data), which has only phenylalanine as the sole aromatic residue. Both TN-I and TN-T have undefined and small dichroic bands in this region, although the aromatic residue contents are not unusual (Wilkinson, 1974).

Determination of f_H and f_β. We used both the constant θ (eq 2 and 5) and constant $\theta\Delta$ (eq 4 and 5) methods for the analysis of the CD spectra. Table II lists the averaged values of computed f_H, $\bar{n}$, i, and f_β for the 12 proteins together with the estimates reported by others. In general, the estimated f_H and f_β by both methods agree well. The estimated $\bar{n}$ and i are uncertain in several cases.

The helical content of G-actin is much higher than that estimated from the Drude and Moffitt equations (Kay, 1960; Nagy and Jencks, 1962) or from the CD spectrum with poly(L-lysine) as the reference compound (Murphy, 1971). Our estimated β form is about the same as that obtained by Murphy (1971).

TABLE I: Extrema of the CD and ORD Spectra of Muscle Proteins below 250 nm at 25 °C.[a,b]

Protein[c]	CD						ORD			
	λ (nm)	[θ]	λ (nm)	[θ]	λ (nm)	[θ]	λ (nm)	[m]	λ (nm)	[m]
G-Actin	219	−12 900	210	−14 000	192	20 800	230	−5 800	200	28 000
Myosin	222	−26 200	210	−25 100	192	56 200	233	−12 900	199	60 100
HMM	222	−22 000	210	−21 200	192	44 800	233	−10 500	199	51 100
SF-1	222	−15 200	210	−15 300	192	29 500	233	−7 300	199	33 300
LMM Fr. I	222	−39 800	210	−38 200	192	87 100	233	−19 400	199	100 000
Tropomyosin	222	−38 800	210	−37 100	192	87 400	233	−18 500	199	85 000
Troponins										
TN-C	222	−12 800	207	−16 100	192	25 300	233	−6 400	199	34 100
TN-I	221	−9 000	205	−15 500	190	10 700	232	−5 600	199	20 400
TN-T	222	−14 900	207	−19 700	191	23 200	232	−7 600	199	36 300
Light chains										
LC_1	220	−8 200	207	−13 000	190	12 900	233	−4 700	199	21 000
LC_2	221	−7 500	207	−11 800	191	9 800	233	−4 600	199	20 200
LC_3	221	−10 100	208	−13 200	192	20 000	233	−5 300	199	23 000

[a] Solvent: 2 mM sodium phosphate (pH 7.0) for G-actin; 40 mM sodium pyrophosphate (pH 7.5) for myosin, HMM, LMM Fr. I, and tropomyosin; 40 mM sodium pyrophosphate and 0.5 mM EGTA (pH 7.5) for troponins; 10 mM sodium phosphate (pH 7.0) for SF-1 and light chains. [b] Dimension of [θ] and [m]: deg cm^2 $dmol^{-1}$. [c] M_0 used for CD and ORD calculations (in the descending order): 112, 115, 115, 115, 115, 115, 113, 117, 118, 109, 110, 111.

TABLE II: Estimated f_H, $\bar{n}$, i, and f_β of Muscle Proteins from Optical Activity.

Proteins	f_H	$\bar{n}$	i	f_β	Previous Work			
					Method	f_H	f_β	Ref
G-Actin	0.45	7	24	0.27	ORD	0.08–0.30		*c*
	(0.37)[a]	(8)	(17)	(0.25)	CD	0.26–0.30	0.26	*d,e*
Myosin	0.78	23	154	0	ORD	0.58–0.60	0.12	*f*
					CD	0.58–0.72	0.10–0.16	*g,h*
HMM	0.70	13	147	0.07	ORD	0.44–0.51	0.17	*f*
					CD	0.45–0.58	0.18–0.22	*g,h*
SF-1	0.60	8	85	0.16	CD	0.32	0.08–0.10	*h*
LMM Fr. I	0.94	505	2[b]	0	ORD	0.97		*f,i*
					CD	1		*g*
Tropomyosin	0.96	284	2[b]	0	ORD	>1		*i,j*
	(0.98)			(0)	CD	0.92–0.94		*j*
Troponins								
TN-C	0.51	8	11	0.13				
	(0.64)	(11)	(9)	(0.11)				
TN-I	0.29	12	5	0.20				
	(0.53)	(11)	(9)	(0.14)				
TN-T	0.38			0.14				
Light chains of myosin								
LC_1	0.37	7	11	0.23	CD	0.40–0.50		*k*
	(0.53)	(14)	(7)	(0.21)				
LC_2	0.33	7	8	0.22	CD	0.40–0.50		*k*
	(0.41)	(9)	(8)	(0.28)				
LC_3	0.48	6	11	0.19	CD	0.40–0.50		*k*
	(0.56)	(14)	(6)	(0.26)				

[a] The values inside parentheses are based on the predictive method from amino acid sequence (Chou and Fasman, 1974); the predicted β turns are not listed. [b] Assume $i = 2$; see text. [c] Kay (1960); Nagy and Jencks (1962); Nagy (1966). [d] Nagy and Strzelecka-Golaszewska (1972). [e] Murphy (1971). [f] McCubbin et al. (1966). [g] Oikawa et al. (1968). [h] Murphy (1974). [i] Cohen and Szent-Gyorgyi (1957); Szent-Gyorgyi et al. (1960). [j] Staprans and Watanabe (1970). [k] Lowey and Holt (1972).

The estimated f_H of myosin is again larger than that from ORD methods (Lowey and Cohen, 1962; McCubbin et al., 1966). The CD analysis by the early procedure of Chen et al. (1972) indicates the presence of both helix and β forms in myosin (Murphy, 1974). However, the reference [θ] values have since been refined because of updated x-ray diffraction results of the five proteins used by Chen et al. (1974). Our results show no detectable β form in myosin (see the results on light chains below). Myosin has about 4000 amino acid residues (N), making it difficult to determine i or $\bar{n}$ accurately. This is more easily seen from the chain-length dependent factor in eq 1, where $f_H(1 - k/\bar{n}) = f_H - ik/N$. The ik/N term is only about 10% of the f_H for myosin (k varies from 2.5 to 3.5 for the three CD bands of a helix). The estimated f_H of HMM is also

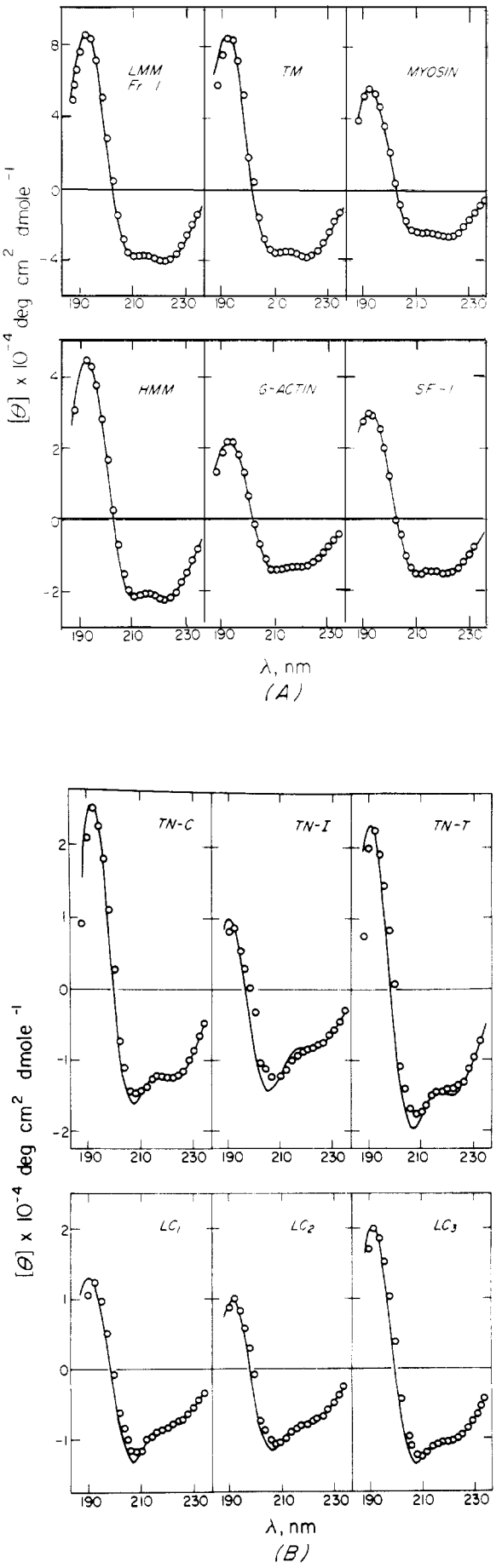

FIGURE 5: (A and B) Comparison of experimental and computed CD spectra of 12 proteins: solid line, experimental; circles, computed.

larger than that reported by McCubbin et al. (1966), Oikawa et al. (1968), and Murphy (1974). The f_H of SF-1 is twice as high as that of Murphy, but the f_β is comparable. The computed f_β's of 0.07 for HMM and 0.16 for SF-1 vary inversely with their molecular weights and probably arise from the β forms in light chains. Both TM and LMM Fr. 1 have high helical contents (close to 90%), but the computed $\bar{n}$ was negative and obviously erroneous. Since the two proteins are known to be a two-stranded helix (Lowey and Cohen, 1962; Woods, 1967), we therefore fixed $i = 2$, which defines $\bar{n}$. Recomputation led to a helical content of more than 90% in both proteins.

To check further the consistency of the estimated f_H's, we compared the helical content of myosin (mol wt 470 000) with that calculated from its subfragments, HMM (mol wt 320 000) and LMM Fr. I (mol wt 120 000). Neglecting the loss of 6–7% in mass due to trypsin digestion, the calculated f_H(myosin) has an average value of 0.77, which agrees well with the computed f_H of 0.78 in Table II.

The b_0 values of the Moffitt equation based on visible ORD of the first six proteins in Table II are listed elsewhere (Yang et al., 1976). Since the b_0 (helix) is also chain length dependent (Chen et al., 1972), a reference value less negative than −630 should be used if the helical contents of G-actin, myosin, HMM, and SF-1 are estimated by the b_0 method.

Among three troponins, TN-I has the smallest f_H and the largest f_β. TN-T appears to have a smaller f_H than TN-I, even though TN-T has larger ellipticities around the double minimum but a slightly smaller ellipticity for the positive band (We were unable to obtain reasonable estimates of $\bar{n}$ for TN-T.) Among the light chains of myosin, LC_3 has the largest f_H, as would be expected from their spectra (Figure 2). f_β is about the same for all three proteins. The helix contents of 33 to 48% are comparable to those estimated by Lowey and Holt (1972). In general, the troponins and light chains have less helical structures than the first six proteins in Table II.

Assuming all the light chains (19–23% β form) are intact in heavy meromyosin (HMM) and subfragment 1 (SF-1), the weight ratio of HMM to LC's would be 320 000:(20 700 + 19 000 + 16 500) ≃ 6:1, and that of SF-1 to LC's would be 120 000:(20 700 + 19 000 + 16 500) ≃ 2:1. The contributions of the β form of light chains to HMM and SF-1 then become about 4 and 10%, which are comparable to those estimated in HMM and SF-1 (8% for HMM; 16% for SF-1). This in turn would lead to about 3% β form in myosin. However, a small amount of β form in heavy chain itself cannot be completely ruled out. In either case, a 3–5% β form in myosin is difficult to detect with the present method of analysis.

With the computed average f_H, f_β, and $\bar{n}$ (Table II), three Δ's (not listed), and $[\theta]_H^\infty$, $[\theta]_\beta$, and $[\theta]_R$ (taken from Chen et al., 1974), we can reconstruct the CD spectra of the 12 proteins (Figures 5A and B); the agreement between the experimental and reconstructed curves is good for the first six proteins and fair for the others. Since both methods have five parameters, a good fit does not necessarily indicate a unique solution, but a poor fit points up the imperfection of the methods. In most cases our reconstructed curves are consistent with the experimental observations.

Discussion

Muscle proteins are not beautiful, crystalline materials; each has an intractability of its own and some are much easier to purify and handle than others (a Board Member's comment). Accordingly, the homogeneity of the preparations should be rigorously tested prior to any meaningful physical measure-

ments (see Experimental Section). Most of our CD and ORD results agree with some of the published ones, but there are differences in other cases. Part of these discrepancies could be attributed to the degree of purity of the preparations. In our case, the maximum impurity amounts to about 5% for myosin, which is as purified as can be achieved at present.

Another possible source of error is the determination of protein concentrations by various methods such as ultraviolet absorptivity (Nagy and Strzelecka-Golaszewska, 1972; Murphy, 1974; Mani et al., 1974b), colorimetry (Murphy, 1971, 1974; Mani et al., 1973), and synthetic boundary by analytical ultracentrifuge (Murray and Kay, 1972). For instance, TN-C was reported to have an $E^{1\%}_{277.5}$ of 2.3 (Murray and Kay, 1972) or $E^{1\%}_{280}$ of 1.4 (van Eerd and Kawasaki, 1973); $E^{1\%}_{280}$ for TN-I ranges from 2.9 (van Eerd and Kawasaki, 1973), to 3.97 (Wilkinson, 1974), to 6.6 (Mani et al., 1974a), and that for TN-T is 4.4, 5.0, and 3.7 by the same workers (although Wilkinson and Mani et al. used the same Rayleigh fringe numbers per unit concentration). We used the micro-Kjeldahl method for nitrogen analysis; the percent nitrogen was calculated from amino acid composition including the amide nitrogen whenever available. Unfortunately, the latter values for SF-1 and TN-T are not reported in the literature. This would make the apparent concentration slightly larger than the true concentration and the calculated $[\theta]$ and $[m]$ could be off by a few percent.

A third source of error is the CD measurements. Recently, Konno et al. (1975) reported that an accurately weighed standard sample (*d*,10-camphorsulfonic acid) sent to several laboratories gave the CD extremum at 290 nm that varied as much as 20%. Clearly, the instrument must be properly calibrated (Cassim and Yang, 1969) and personal errors minimized so as to obtain reproducible measurements.

The methods of analysis of CD spectrum used in this work differ from the earlier methods such as the b_0, $[m]_{233}$, $[\theta]_{222}$, which are only used to estimate the helical content, but not the β form. All these methods by necessity involve many assumptions and therefore are empirical. Whether such estimates are good remains to be seen when the x-ray structures of these proteins become available. However, muscle proteins have not yet been crystallized. The structure of myosin having a high molecular weight is difficult to study by x-ray, even if it can be crystallized. The next best approach is then to compare the CD analysis with some other empirical method that has no bearing with optical activity. Recently, numerous methods of predicting the secondary structure from the primary structure of proteins have been proposed but with varying degrees of success. We choose the method of Chou and Fasman (1974) purely for its simplicity; according to these authors, this method has a 70–80% accuracy of prediction. The very simplicity could have an adverse effect; it tempts us to use the method indiscriminately and to overlook its shortcoming. In some cases the prediction can favor either the helix or β form for a segment, whereas in other cases exceptions to the rules must be introduced and the assignment becomes ambiguous. Thus, the method is by no means straightforward.

With this caution in mind we determine the predicted conformation of several muscle proteins, since the amino acid sequences of G-actin (Elzinga et al., 1973; Collins and Elzinga, 1975), the α chain of TM (Stone et al., 1974), TN-C (Collins et al., 1973), TN-I (Wilkinson and Grand, 1975), LC_1 (Frank and Weeds, 1974), LG_2 (Collins, 1976), and LC_3 (Frank and Weeds, 1974) are available. For G-actin, our estimated f_H of 0.45 and f_β of 0.27 compare favorably with the predicted values of 0.37 and 0.25, respectively. The α chain of TM seems to have several helix breaking tetrapeptides, but the extension of helical segments from both sides of the tetrapeptides makes it into one uninterrupted long helix except at the C terminal. Since the Chou–Fasman parameters are based on the x-ray structures of globular proteins, the good prediction of $f_H = 0.98$ is more than can be expected.

The predicted f_H of TN-C is considerably higher than that based on CD analysis. However, the addition of calcium ion (1 mM) would raise the estimated f_H to 0.66, which is close to the predicted value (Wu and Yang, 1975). Thus, the negatively charged carboxylate ions at the binding sites might destabilize some helices that have been predicted by the sequence method. This electrostatic effect is also observed during protonation, as evidenced by the enhancement of the CD band at 222 nm upon lowering the pH of the protein solution (Lehrer and Leavis, 1974). The predicted f_H of TN-I is almost twice that by our estimation. One plausible explanation is the exposure of the proteins to urea during fractionation, which might have partially denatured TN-I (but it seems to have no effect on TN-C). On the other hand, the amino acid sequence of TN-I shows a concentration of charged residues in certain segments, neglect of which could lead to an overprediction by the Chou–Fasman method.

The amino acid sequences of LC_1 and LC_3 are identical at the C-terminal 141 residues; LC_1 has 49 and LC_3 has 8 additional residues at the N-terminal end (Frank and Weeds, 1974). Their predicted f_H values are similar, but higher (especially LC_1) than those obtained by the optical methods, although f_β values are comparable (Table II). The use of urea during separation of troponins and light chains might have altered the protein conformation. From their antibody studies, Lowey and Holt (1972) concluded that the conformations of bound alkali light chains (LC_1 and LC_3) are very different from that of dissociated ones, but the Nbs_2 light chain (LC_2) is probably unchanged. The predicted f_H and f_β of LC_2 ($f_H = 0.41$, $f_\beta = 0.28$) (Collins, 1976) are slightly higher than those obtained by the CD method (Table II). Our results on the light chains of myosin and troponins should be viewed with some reservation because of the relatively poor agreement between reconstructed and experimental CD curves for these proteins (Figure 5B). The reference proteins used for determining $[\theta]_\beta$ and $[\theta]_R$ have relatively low β form contents (Chen et al., 1974). Because of the smaller rotatory contribution of β form than helix, the resulted $[\theta]_\beta$'s are probably less certain. In low helix-containing proteins this contribution might weigh more than has been expected.

The amino acid sequence of TN-T is underway (Pearlstone et al., 1975). With the success of the crystallization of TN-C (Mercola et al., 1975), it will not be long before the first x-ray data of muscle proteins become available. Our CD method of analysis will then be subject to a crucial test.

In the absence of x-ray diffraction results, the predictive methods complement the CD analysis of native proteins. A reasonable agreement in results from the two entirely different approaches would increase our confidence in the use of these empirical treatments. On the other hand, serious disaccord can compel us to look for possible sources of errors in either or both of the two tools. In spite of their shortcomings, CD and ORD remain powerful physical methods of studying the conformation and conformational changes of proteins in solutions.

Acknowledgment

We thank Professor S. Watanabe for his suggestions and for a gift of tropomyosin preparations and Dr. Y. H. Chen for his discussion about the computer programs.

References

Azuma, N., and Watanabe, S. (1965), *J. Biol. Chem. 240,* 3847–3851.

Becker, J. W., Reeke, G. N., Jr., Wang, J. L., Cunninghum, B. A., and Edelman, G. M. (1975), *J. Biol. Chem. 250,* 1513–1524.

Blake, C. C. F., Maer, G. A., North, A. C. T., Phillips, D. C., and Sarma, V. R. (1967), *Proc. R. Soc. London, Ser. B 167,* 365–377.

Blow, D. M. (1969), *Biochem. J. 112,* 261–268.

Cassim, J. Y., and Yang, J. T. (1969), *Biochemistry 8,* 1947–1951.

Cassim, J. Y., and Yang, J. T. (1970), *Biopolymers 9,* 1475–1502.

Chen, Y. H., Yang, J. T., and Martinez, H. M. (1972), *Biochemistry 11,* 4120–4131.

Chen, Y. H., Yang, J. T., and Chau, K. H. (1974), *Biochemistry 13,* 3350–3359.

Chou, P. Y., and Fasman, G. D. (1974), *Biochemistry 13,* 222–245.

Chung, C. S., Richards, E. G., and Olcott, H. S. (1967), *Biochemistry 6,* 3154–3161.

Cohen, C., and Szent-Gyorgyi, A. G. (1957), *J. Am. Chem. Soc. 79,* 248.

Collins, J. H. (1976), *Nature (London) 259,* 699–700.

Collins, J. H., and Elzinga, M. (1975), *J. Biol. Chem. 250,* 5915–5920.

Collins, J. H., Potter, J. D., Horn, M. J., Wilshire, G., and Jackman, N. (1973), *FEBS Lett. 36,* 268–272.

Cooke, R. (1972), *Biochem. Biophys. Res. Commun. 49,* 1021–1028.

Elzinga, M., Collins, J. H., Kuehl, W. M., and Adelstein, R. S. (1973), *Proc. Natl. Acad. Sci. U.S.A. 70,* 2687–2691.

Frank, G., and Weeds, A. G. (1974), *Eur. J. Biochem. 44,* 317–334.

Greenfield, N. J., Davidson, B., and Fasman, G. D. (1967), *Biochemistry 6,* 1630–1637.

Greenfield, N. J., and Fasman, G. D. (1969), *Biochemistry 8,* 4108–4116.

Kay, C. M. (1960), *Biochim. Biophys. Acta 43,* 259–267.

Kominz, D. R., Hough, A., Symonds, P., and Laki, K. (1954), *Arch. Biochem. Biophys. 50,* 148–159.

Konno, T., Meguro, H., and Tuzimura, K. (1975), *Anal. Biochem. 67,* 226–232.

Lehrer, S. S., and Leavis, P. C. (1974), *Biochem. Biophys. Res. Commun. 58,* 159–165.

Lowey, S., and Cohen, C. (1962), *J. Mol. Biol. 4,* 293–308.

Lowey, S., and Holt, J. C. (1972), *Cold Spring Harbor Symp. Quant. Biol. 37,* 19–28.

Mani, R. S., McCubbin, W. D., and Kay, C. M. (1973), *FEBS Lett. 29,* 243–247.

Mani, R. S., McCubbin, W. D., and Kay, C. M. (1974a), *Biochemistry 13,* 5003–5007.

Mani, R. S., McCubbin, W. D., and Kay, C. M. (1974b), *FEBS Lett. 38,* 357–360.

McCubbin, W. D., Kay, C. M., and Oikawa, K. (1966), *Biochim. Biophys. Acta 126,* 597–600.

Mercola, D., Bullard, B., and Priest, J. (1975), *Nature (London) 254,* 634–635.

Murphy, A. J. (1971), *Biochemistry 10,* 3723–3728.

Murphy, A. J. (1974), *Arch. Biochem. Biophys. 163,* 290–296.

Murray, A. C., and Kay, C. M. (1972), *Biochemistry 11,* 2622–2627.

Nagy, B. (1966), *Biochim. Biophys. Acta 115,* 498–500.

Nagy, B., and Jencks, W. P. (1962), *Biochemistry 1,* 987–996.

Nagy, B., and Strzelecka-Golaszewska, H. (1972), *Arch. Biochem. Biophys. 150,* 428–435.

Offer, G., Moos, C., and Starr, R. (1973), *J. Mol. Biol. 74,* 653–676.

Oikawa, K., Kay, C. M., and McCubbin, W. D. (1968), *Biochim. Biophys. Acta 168,* 164–167.

Pearlstone, J. R., Carpenter, M. R., and Smillie, L. R. (1975), *Fed. Proc., Fed. Am. Soc. Exp. Biol. 34,* 539.

Richards, E. G., Chung, C. S., Menzel, D. B., and Olcott, H. S. (1967), *Biochemistry 6,* 528–540.

Samejima, T., and Yang, J. T. (1964), *Biochemistry 3,* 613–616.

Saxena, I. P., and Wetlaufer, D. B. (1971), *Proc. Natl. Acad. Sci. U.S.A. 68,* 969–972.

Spudich, J., and Watt, S. (1971), *J. Biol. Chem. 246,* 4866–4871.

Staprans, I., and Watanabe, S. (1970), *J. Biol. Chem. 245,* 5962–5966.

Staprans, I., Takahashi, H., Russell, M. P., and Watanabe, S. (1972), *J. Biochem. 72,* 723–735.

Stone, D., Sodek, J., Johnson, P., and Smillie, L. B. (1974), *Fed. Eur. Biochem. Soc. Meet., [Proc.], 9th, 31,* 125–136.

Straus, J. H., Gordon, A. S., and Wallach, D. F. H. (1969), *Eur. J. Biochem. 11,* 201–212.

Szent-Gyorgyi, A. G. (1953), *Arch. Biochem. Biophys. 42,* 305–320.

Szent-Gyorgyi, A. G., Cohen, C., and Philpott, D. E. (1960), *J. Mol. Biol. 2,* 133–142.

Tufty, R. M., and Kretsinger, R. H. (1975), *Science 187,* 167–169.

van Eerd, J.-P., and Kawasaki, Y. (1973), *Biochemistry 12,* 4972–4980.

Weber, K., and Osborn, M. (1969), *J. Biol. Chem. 244,* 4406–4412.

Weeds, A. G., and Lowey, S. (1971), *J. Mol. Biol. 61,* 701–725.

Weeds, A. G., and McLachlan, D. M. (1974), *Nature (London) 252,* 646–649.

Wilkinson, J. M. (1974), *Biochim. Biophys. Acta 359,* 379–388.

Wilkinson, J. M., and Grand, R. J. A. (1975), *Biochem. J. 149,* 493–496.

Woods, E. F. (1967), *J. Biol. Chem. 242,* 2859–2871.

Wu, C. S. C., and Sayre, R. N. (1971), *J. Food Sci. 36,* 133–137.

Wu, C. S. C., and Yang, J. T. (1975), *Fed. Proc., Fed. Am. Soc. Exp. Biol. 34,* 538.

Wyckoff, H. W., Tsemoglou, D., Hanson, A. W., Knox, J. R., Lee, B., and Richards, F. M. (1970), *J. Biol. Chem. 245,* 305–328.

Yang, J. T., Chen, G. C., and Jirgensons, B. (1976), Handbook of Biochemistry and Molecular Biology, 3rd ed, Vol. IA, Fasman, G. D., Ed., Cleveland, Ohio, The Chemical Rubber Co., in press.

Young, D. M., Himmelfarb, S., and Harrington, W. F. (1965), *J. Biol. Chem. 240,* 2428–2436.