von der Haar, F., & Gaertner, E. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 1378–1382.
von der Haar, F., & Cramer, F. (1976) *Biochemistry 15*, 4131–4138.
von der Haar, F., & Cramer, F. (1978) *Biochemistry 17*, 4509–4514.
von der Haar, F., Schlimme, E., Gomez-Guillen, M., & Cramer, F. (1971) *Eur. J. Biochem. 24*, 296–302.

Physical Properties of Collagen–Sodium Dodecyl Sulfate Complexes†

J. William Freytag,§ Milton E. Noelken, and Billy G. Hudson*

ABSTRACT: Sodium dodecyl sulfate (NaDodSO$_4$)–polyacrylamide gel electrophoresis and gel filtration chromatography of protein–NaDodSO$_4$ complexes are frequently used to characterize collagen-like polypeptide components in mixtures obtained from extracts of basement membranes. However, electrophoresis yields anomalously high apparent molecular weights for collagenous polypeptides when typical globular proteins are used as molecular weight standards, and the use of gel filtration chromatography for this purpose was suspect because Nozaki et al. [Nozaki, Y., Schechter, N. M., Reynolds, J. A., & Tanford, C. (1976) *Biochemistry 15*, 3884–3890] found that asymmetric particles, including NaDodSO$_4$–protein complexes, coeluted with native globular proteins of lower Stokes radius, when Sepharose 4B was used. To understand these effects and to improve the characterization of collagenous polypeptides, we investigated the secondary structure of NaDodSO$_4$–collagen complexes with the use of circular dichroism, measured the NaDodSO$_4$ content, studied the dependence of electrophoretic mobility on gel concentration, and extended work on gel filtration by use of a more porous gel, Sepharose CL-4B. We found that the anomalous behavior of collagen chains on NaDodSO$_4$–polyacrylamide gel electrophoresis is due in large part to treatment of data and that the method can be used to determine rather accurate values for the number of residues per polypeptide chain. Our gel filtration results indicated that reliable molecular weights can be obtained when Sepharose CL-4B is used. These methods can be applied equally well to collagenous and noncollagenous polypeptides.

Polyacrylamide gel electrophoresis in the presence of the anionic detergent NaDodSO$_4$[1] has proved to be an excellent method for estimating the molecular weights of proteins. It is especially useful for analysis of mixtures, such as those obtained by extraction of membranes, before the components can be separated and characterized by methods which have a stronger theoretical foundation, such as equilibrium sedimentation. However, collagen and related collagenous proteins chains, which are widely distributed in nature and are important components of many tissues, were found to behave anomalously on NaDodSO$_4$–polyacrylamide gel electrophoresis. Furthmayr & Timpl (1971), for example, found that the α_1-collagen chain, molecular weight 96 000, migrated at a slightly lower rate than the human serum albumin dimer, molecular weight 132 000, thus behaving as if it had a molecular weight 40% higher than the true value. A similar effect was noted for the α_2-collagen chain, and, in addition, the two collagen chains could be easily resolved from each other, although they presumably have almost identical molecular weights. Aggregation was an unlikely cause for these effects since a good linear relationship, in the molecular range 5000–96 000, was found between log molecular weight and electrophoretic mobility of α_1-collagen and four fragments obtained from it by cyanogen bromide cleavage. This would occur only if all the five species were either unaggregated or, rather unlikely, partially aggregated to a degree that varied in a unique, monotonic manner with molecular weight. Furthermore, a closely lying parallel line was obtained for the α_2-collagen chain and several of its cyanogen bromide fragments. Furthmayr & Timpl (1971) suggested that there was a fundamental conformation difference between NaDodSO$_4$ complexes of typical proteins and those of collagenous proteins, which have unusual amino acid compositions and sequences, and that this was responsible for the atypical behavior of the latter on NaDodSO$_4$–polyacrylamide gel electrophoresis. The subtlety of the conformation effect was indicated by the difference between the log molecular weight–mobility relationship of the α_1-collagen chain and its cyanogen bromide fragments and that of the α_2-collagen chain and its fragments. While this seemed to be a reasonable explanation, it was incomplete in the sense that other factors such as binding of NaDodSO$_4$ were not considered, and there were no independent conformation studies.

A second method commonly used to estimate molecular weights is gel filtration chromatography in the presence of NaDodSO$_4$. This has been used several times to characterize collagenous components of basement membranes [e.g., Grant et al. (1973) and Clark & Kefalides (1979)]. However, Nozaki et al. (1976) have shown that a fundamental assumption underlying this method, i.e., elution volume is determined solely by the Stokes radius, is invalid. NaDodSO$_4$–protein complexes and other asymmetric particles were shown to coelute from Sepharose 4B columns with native

† From the Department of Biochemistry, University of Kansas Medical Center, Kansas City, Kansas 66103, and the Department of Biochemistry, Oklahoma State University, Stillwater, Oklahoma 74074. *Received March 9, 1979.* This work was supported by National Institutes of Health Grant AM 18381.

§ Taken in part from the Ph.D. thesis of J.W.F., Department of Biochemistry, University of Kansas Medical Center, Kansas City, KS, 1978. Present address: Department of Biochemistry, Duke University Medical Center, Durham, NC 27710.

[1] Abbreviations used: NaDodSO$_4$, sodium dodecyl sulfate; BSA, bovine serum albumin; EDTA, ethylenediaminetetraacetic acid (tetrasodium salt); CD, circular dichroism; *N*, number of residues in a polypeptide chain.

0006-2960/79/0418-4761$01.00/0

globular proteins of lower Stokes radii. This effect, due perhaps to conformational factors, could lead to serious errors if not taken into account.

The purpose of this study was to extend the investigation of the anomalous $NaDodSO_4$–polyacrylamide gel electrophoresis behavior of collagen α_1 and α_2 chains and the gel filtration studies of Nozaki et al. (1976) with emphasis on collagen. Additional impetus was given to this research by the fact that some basement membrane polypeptides show similar $NaDodSO_4$–polyacrylamide gel electrophoresis behavior and possess some of the same chemical properties as collagen, i.e., high glycine and imino acid content, but appear to have markedly higher molecular weights than α_1 or α_2 chains of classical collagen (Kefalides, 1973; Hung et al., 1977, 1979; Freytag et al., 1978), the largest collagenous species considered by Furthmayr & Timpl (1971). Better understanding of the anomalous $NaDodSO_4$–polyacrylamide gel electrophoresis behavior of collagen chains or, alternatively, minimization of the conformation effect in gel filtration chromatography could lead to more accurate molecular weight estimates of collagenous components of mixtures, particularly those of basement membranes which have molecular weights greater than that of the α_1-collagen chain. Thus, we used circular dichroism and gel filtration chromatography to investigate the secondary structure and size, respectively, of the $NaDodSO_4$–collagen complex, studied the dependence of electrophoretic mobility on polyacrylamide gel concentration, and measured the amount of $NaDodSO_4$ bound in order to rationalize the anomalous electrophoretic mobility. In addition, we studied the gel filtration behavior of $NaDodSO_4$–collagen complexes, using a more porous gel than that used by Nozaki et al. (1976).

Materials and Methods

Materials. Acid-soluble calf skin collagen was purchased from Calbiochem; the α_1-collagen chain was isolated from it by the method of Piez et al. (1963). Rabbit muscle myosin, used as a standard for electrophoresis and gel filtration, was prepared by the method of Perry (1955). Other standard proteins were obtained from Sigma Chemical Co. and included *Escherichia coli* β-galactosidase, phosphorylase *a* (2× crystallized), BSA (fatty acid free), bovine catalase (2× crystallized), ovalbumin (Grade VI), and bovine pancreatic chymotrypsinogen (6× crystallized). Most supplies for electrophoresis, i.e., $NaDodSO_4$, acrylamide, *N,N'*-methylenebis(acrylamide), *N,N,N',N'*-tetramethylenediamine, ammonium persulfate, Coomassie brilliant blue R-250, bromophenol blue, etc., were purchased from Bio-Rad Laboratories; 2-mercaptoethanol (98%) was obtained from Eastman Organic Chemicals. $NaDodSO_4$ was twice recrystallized from boiling 95% ethanol (Emerson & Holtzer, 1967) for use in binding studies. Sulfur-35-labeled $NaDodSO_4$ (96.6 mCi/g) was purchased from Amersham/Searle. All other reagents were the highest purity commericially obtainable. Biofluor premixed scintillation cocktail was the product of New England Nuclear. Visking dialysis tubing (8/32) was purchased from Union Carbide. Sepharose CL-4B was purchased from Pharmacia Fine Chemicals.

$NaDodSO_4$–Polyacrylamide Gel Electrophoresis. This was performed essentially as described by Hudson & Spiro (1972) and Weber & Osborn (1969), using a continuous 0.1% $NaDodSO_4$–0.1 M sodium phosphate, pH 7.0, buffer system. Chymotrypsinogen was used instead of a tracking dye for calculations of the relative mobility, R_f, of proteins (migration distance divided by that of chymotrypsinogen) in order to minimize measurement errors due to diffusion. Samples were prepared for electrophoresis by incubation in 1% $NaDodSO_4$, 0.1 M sodium phosphate, pH 7.0, and 2% 2-mercaptoethanol overnight at 37 °C with constant shaking. Cylindrical gels (5.5 × 90 mm) of 4, 5, 6, 7, 8, and 9% acrylamide, with a constant ratio (37:1) of acrylamide to *N,N'*-methylenebis(acrylamide), were used, in order to construct Ferguson plots (log R_f vs. percent acrylamide; Ferguson, 1964). We obtained maximum accuracy in determining values of R_f by scanning the Coomassie brilliant blue stained gels in a Gilford Model 240 spectrophotometer equipped with a Model 42108 linear transport gel scanner and making measurements on the recorder traces and by applying all the proteins used as a complete mixture to individual gels. In the analysis of the electrophoresis results, the following molecular weights and number of residues were employed for the standard proteins: myosin, 212000 and 1690 (Darnall & Klotz, 1975; Huszar & Elzinga, 1971); *E. coli* β-galactosidase, 116000 and 1021 (Fowler & Zabin, 1977); phosphorylase *a*, 92500 and 797 (Seery et al., 1970; Sevilla & Fischer, 1969); bovine serum albumin, 66000 and 581 (Feldhoff & Peters, 1975); catalase, 57500 and 505 (Schroeder et al., 1969); ovalbumin, 43000 and 370 (Castellino & Barker, 1968; Neuberger & Marshall, 1966); chymotrypsinogen, 25700 and 245 (Smillie et al., 1968). Where possible, the molecular weight and number of residues were obtained from reported amino acid sequences and in other cases from reported physical measurements and mean residue weights obtained from amino acid compositions.

Determination of Stokes Radius by Gel Filtration Chromatography. The Stokes radii of collagen chain–$NaDodSO_4$ complexes were determined on Sepharose CL-4B (Pharmacia) packed in a column (1.2 × 90 cm) to a height of 80 cm. The column was equilibrated with 0.1% $NaDodSO_4$, 0.1 M sodium phosphate, pH 7.0, and 0.1% 2-mercaptoethanol. A sample, containing the proteins 2 mg of myosin, 2 mg of calf skin collagen, 1 mg of phosphorylase *a*, 1 mg of BSA, 1 mg of ovalbumin, 1 mg of chymotrypsinogen, and 1 mg of Dnp-alanine in 1.0 mL of 5% $NaDodSO_4$, 0.1 M sodium phosphate, pH 7.0, and 5% 2-mercaptoethanol, was prepared by incubation at 100 °C for 5 min, followed by incubation at 37 °C for another 3 h. The entire sample was applied to the Sepharose CL-4B column and eluted with 0.1% $NaDodSO_4$, 0.1 M sodium phosphate, pH 7.0, and 0.1% 2-mercaptoethanol at a flow rate of 4 mL/h. Fractions of 0.6 mL were collected. The elution of Dnp-alanine was monitored by measuring the absorbance at 410 nm. The elution volumes of the various proteins were determined by electrophoresing 100 μL of each fraction on 5% $NaDodSO_4$–polyacrylamide gels and then spectrophotometrically scanning the gels to determine the fraction containing the highest concentration of each given polypeptide. This method is particularly advantageous since it eliminates the error involved in making several different runs for the calibration alone, followed by a separate run for the unknown, and it is much faster and more convenient.

The chromatographic data were fitted with the empirical equation of Ackers (1967), $R_S = a_0 + b_0\text{erf}^{-1}(1 - K_d)$, where R_S is the Stokes radius of a $NaDodSO_4$–protein complex, K_d is its distribution coefficient, erf^{-1} is the inverse of the error function, and a_0 and b_0 are curve-fitting parameters. The Stokes radii of the standards were obtained from the reports of Reynolds & Tanford (1970) and Tanford et al. (1974). We used the Stokes radius related to the frictional coefficient as well as that related to intrinsic viscosity in calibrating the gel, since these are not identical and it is not known which more closely represents the theoretical radius which determines the partitioning of the solute. This is not an important matter in

this study since we are interested in comparing the hydrodynamic behavior of the NaDodSO$_4$–collagen α-chain complex with those of more typical proteins and each calibration method yields internally consistent results. The following are the standard proteins used and the values of Stokes radii from frictional coefficient and viscosity measurements: myosin, 146 Å (frictional coefficient) and 193 Å (viscosity); phosphorylase *a*, 87 and 107 Å; bovine serum albumin, 71 and 85 Å; ovalbumin, 53 and 60 Å; chymotrypsinogen, 38 and 41 Å.

Binding of NaDodSO$_4$. Binding was measured by equilibrium dialysis (Steinhardt & Reynolds, 1969) and gel filtration chromatography (Tanford et al., 1974). With the dialysis procedure, the binding of NaDodSO$_4$ to calf skin collagen, BSA, and ovalbumin was determined under two sets of conditions, one in which 5 mL of protein solution (1 mg/mL) was incubated in 1% NaDodSO$_4$, 0.26 ionic strength sodium phosphate, pH 7.2, 1% 2-mercaptoethanol, and 0.02% sodium azide at 37 °C overnight prior to exhaustive dialysis at room temperature against 400 mL of that solvent. The second set of conditions was identical except that 0.1% NaDodSO$_4$ was used instead of 1%. Dialysis tubing was treated before using by boiling for 1 h in several changes of 1 mM EDTA. The NaDodSO$_4$ in both protein solution and diffusate was prepared with a specific activity (^{35}S) of 50 μCi/g of NaDodSO$_4$. Four weeks were required to reach equilibrium; bacterial growth was inhibited by the sodium azide present. After equilibrium was attained, the NaDodSO$_4$ concentration of the inner and outer solutions was determined by liquid scintillation spectrometry. Aliquots of a 0.5-mL sample diluted with 10 mL of Biofluor scintillation fluid were counted in a Packard Tri-Carb liquid scintillation spectrometer, Model 3320. Conversion of observed counts to concentration units was made by comparison with standard radioactively labeled detergent solutions of the same specific activity. Corrections for quenching were made by adding aliquots of unlabeled NaDodSO$_4$ at the same concentrations as the radiolabeled solutions so all standard calibration points contained the identical amounts of NaDodSO$_4$; the only difference was the amount of labeled NaDodSO$_4$. No corrections for the Donnan effect were required. Protein concentrations inside the bag were determined by the method of Hartree (1972) using each protein for its respective calibration curve. For the purpose of this paper, detergent binding was most conveniently expressed on a weight basis rather than a molar basis, i.e., as $^{\delta}$NaDodSO$_4$ to designate grams of detergent per gram of protein. This was calculated by using the equation

$$^{\delta}\text{NaDodSO}_4 = \frac{C_{\text{NaDodSO}_4\text{(inside)}} - C_{\text{NaDodSO}_4\text{(outside)}}}{C_{\text{protein}}}$$

where $C_{\text{NaDodSO}_4\text{(inside)}}$ is the concentration of NaDodSO$_4$ (grams per milliliter) inside the dialysis bag and C_{protein} is the concentration of protein (grams per milliliter) inside the bag.

To rule out possible artifacts related to the extremely slow rate of equilibrium across the dialysis membrane at NaDodSO$_4$ concentrations above the critical micelle concentration, we also measured binding of NaDodSO$_4$ to collagen by gel chromatography. Calf skin collagen (5 mg) was incubated with 2.5 mL of 1% NaDodSO$_4$ and 0.26 ionic strength sodium phosphate (50 μCi of ^{35}S per g of NaDodSO$_4$) at 100 °C for 30 min. The entire sample was applied to a Sepharose CL-4B gel filtration column (1.2 × 90 cm) which had been equilibrated with 0.1% NaDodSO$_4$ and 0.26 ionic strength sodium phosphate, pH 7.2 (50 μCi/g of NaDodSO$_4$). The column was eluted with equilibration buffer at 4 mL/h and 1.5-mL fractions were collected. The column effluent was monitored

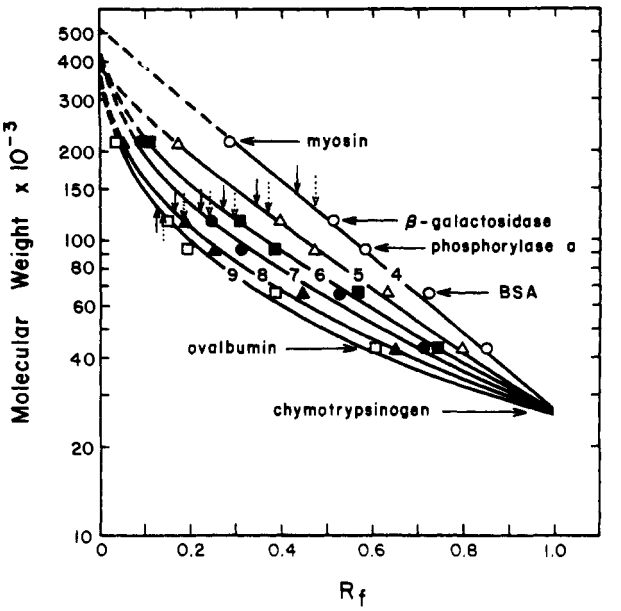

FIGURE 1: NaDodSO$_4$–polyacrylamide gel electrophoresis of standard proteins and collagen chains in 4, 5, 6, 7, 8, and 9% acrylamide gels. The solid arrow represents the α_1-collagen chain, and the broken arrow represents the α_2-collagen chain. The other symbols represent standard proteins as indicated.

for (^{35}S)NaDodSO$_4$ by counting 300-μL aliquots of each fraction in 10 mL of Biofluor in a scintillation counter. Since collagen has an extremely low absorbance at 280 nm, the position of each polypeptide chain was determined by subjecting 100 μL of each column fraction to NaDodSO$_4$–polyacrylamide gel electrophoresis. The concentration of NaDodSO$_4$ was determined for the peak fraction containing the α chains and for several fractions which were free of protein and NaDodSO$_4$ micelles. Protein concentration was also determined for the α-chain fraction by the method of Hartree (1972), using collagen as a standard. The binding was calculated by

$$^{\delta}\text{NaDodSO}_4 = \frac{C_{\text{NaDodSO}_4\text{(protein)}} - C_{\text{NaDodSO}_4\text{(base line)}}}{C_{\text{(protein)}}}$$

where $C_{\text{NaDodSO}_4\text{(protein)}}$ is the concentration (grams per milliliter) of NaDodSO$_4$ in the protein peak, $N_{\text{NaDodSO}_4\text{(base line)}}$ is the concentration (grams per milliliter) of NaDodSO$_4$ along the base line, and $C_{\text{(protein)}}$ is the concentration (grams per milliliter) of collagen in the protein peak.

Circular Dichroism. Spectra in the wavelength region 185–275 nm were obtained at room temperature in a Cary 61 CD recording spectropolarimeter. Protein concentrations (100–200 mg/L) were determined by the method of Hartree (1972). Samples were clarified with Millipore filters. A 1-mm cylindrical cell was used, and spectra were obtained by scanning twice at 0.1 nm/s with a 3-s time constant. The results were expressed in terms of [Θ], the mean residue ellipticity. A mean residue weight of 91.6 (Piez & Gross, 1960) was used in the calculations.

Results

NaDodSO$_4$–Polyacrylamide Gel Electrophoresis. The curves of log molecular weight vs. R_f for the standard proteins and collagen at several gel concentrations are plotted in Figure 1. A linear relationship was found at 4% gel concentration (correlation coefficient of magnitude greater than 0.99), but at higher gel concentrations the calibration curves could be fitted better with a nonlinear empirical relationship:

$$\log M_r = a_0 + a_1 R_f^{1/2} + a_2 R_f$$

where M_r is molecular weight and a_0, a_1, and a_2 are curve-fitting parameters which have values that depend upon gel concentration. Correlation coefficients of magnitude greater than 0.99 were obtained for the 5, 6, 7, 8, and 9% gel calibration curves. The curvature at 5, 6, and 7% gel concentrations was not apparent in many earlier studies reported in the literature largely because a widely used molecular weight for β-galactosidase of 130 000 (Ullmann et al., 1968; Weber & Osborn, 1969) is substantially higher than the value 116 000 recently found by Fowler & Zabin (1977), who determined the complete amino acid sequence. The values of R_f for α_1 and α_2 chains of collagen are indicated by arrows in Figure 1. The apparent molecular weights found for α_1 and α_2 chains had values of 133 000 ± 5000 (average difference) and 122 000 ± 3000 and had no significant dependence on gel concentration. Both of these values are higher than the value of 98 000 ± 5000 which was determined for the α_1 chain by sedimentation equilibrium analysis (Lewis & Piez, 1964) and the value 96 400 which is the product of the mean residue weight of 91.63 and 1052, the number of residues per chain (Fietzek & Kühn, 1976). Presumably, the α_2 chain, which has a very similar amino acid composition, has essentially the same molecular weight. These results are in accord with those of Furthmayr & Timpl (1971), who found similarly high apparent molecular weights when 5% gels were used. We also found that the β chains had apparent molecular weights greater than 230 000 compared with a real value of 192 000 and that they could be resolved by polyacrylamide gel electrophoresis.

We then evaluated the electrophoresis data by means of Ferguson plots (Ferguson, 1964) which are based on the equation

$$\log R_f = \log Y_0 - K_R T$$

where Y_0 is the free electrophoretic mobility and is, in theory, a function of molecular charge, size, and shape, T is the percentage of acrylamide, and K_R is the retardation coefficient, which is theoretically a function of size and shape. The existence of a common value of Y_0 for standard and unknown proteins and a good correlation between Y_0, K_R, and molecular weight are requirements for the determination of molecular weight at a single gel concentration. These conditions are not met in the case of the α-collagen chains, as they appear to deviate from the standard proteins used in $NaDodSO_4$–polyacrylamide gel electrophoresis. However, Frank & Rodbard (1975) pointed out that Y_0 and K_R can be independent variables [see also Rodbard & Chrambach (1971)] in that a particle may need to undergo shape changes in order to penetrate pores, while in free solution it may adapt a unique conformation that minimizes free energy in an electrical field. Thus, it is possible that a protein would behave anomalously at individual gel concentrations but still obey the same relationship of K_R to molecular weights as other proteins. We then determined values of K_R for the α chains and standard proteins using the values of R_f shown in Figure 1. A weighted linear regression analysis (log R_f vs. T) was made (Frank & Rodbard, 1975) in order to take into account the higher accuracy of larger values of R_f. The correlation coefficients of the straight lines obtained were 0.99 or greater, with the exception of that for catalase, for which a value of 0.97 was obtained. Values of molecular weights, M_r, are plotted vs. K_R in Figure 2. Following the procedure of Frank & Rodbard (1975), we fit the data for the standard proteins (open circles) to an empirical parabolic relationship

$$10^{-4} M_r = a_0 + a_1 100 K_R + a_2 (100 K_R)^2$$

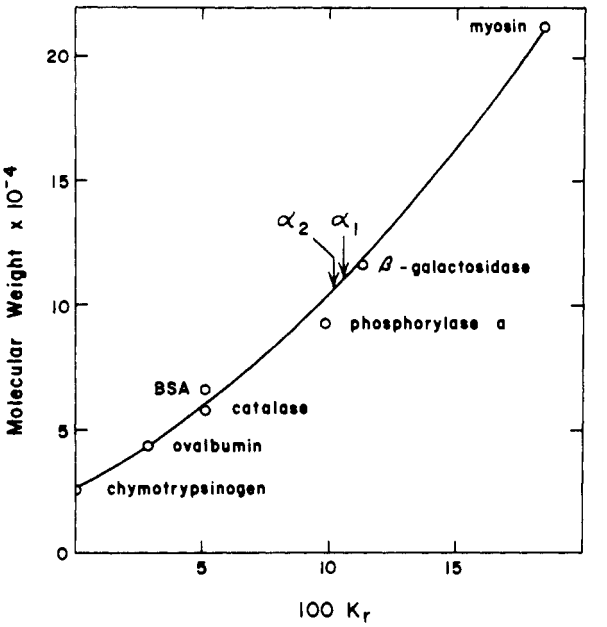

FIGURE 2: Values of the retardation coefficient, K_R, obtained from $NaDodSO_4$–polyacrylamide gel electrophoresis for several standard $NaDodSO_4$–protein complexes (see text), represented by open circles; $NaDodSO_4$–α_1-collagen chain and $NaDodSO_4$–α_2 chain complexes are represented by arrows.

where M_r is molecular weight and the values of a_0, a_1, and a_2 are 2.657, 0.510, and 0.027, respectively; the correlation coefficient was 0.996. Since Frank and Rodbard used a tracking dye migration distance as the basis for calculation of R_f (and thus K_R), our curve-fitting parameters differ from theirs.

The collagen chains have values of K_R that correspond to molecular weights of 110 000 and 106 000 for α_1 and α_2 chains, respectively. Thus, improved results were obtained by this method of analysis. The apparent molecular weights are closer to the true value of 96 000, and the α chains have more nearly identical apparent molecular weights when determined from K_R values rather than R_f values at single gel concentrations.

Gel Filtration Chromatography. The results are illustrated in Figure 3. A linear relationship was found between Stokes radius and $\text{erf}^{-1}(1 - K_d)$ for $NaDodSO_4$–protein complexes corresponding to protein molecular weights of 25 700–212 000, when either viscosity or frictional coefficient related Stokes radii for the standard proteins were used to construct the standard curve, shown in Figure 3A. A linear relationship was also found between log molecular weight and K_d (Figure 3B) over the molecular weight range 43 000 (ovalbumin) to 212 000 (myosin); the correlation coefficient was 0.999. In contrast to their behavior in $NaDodSO_4$–polyacrylamide gel electrophoresis, the α chains of collagen were not resolved by gel filtration chromatography. Furthermore, the Stokes radius of the $NaDodSO_4$ complex had a value expected for a protein of 96 000 molecular weight, a value slightly larger than that found for phosphorylase *a*, molecular weight 92 500 (Figure 3A). This similarity in elution behavior to $NaDodSO_4$ complexes derived from more typical proteins is further demonstrated in Figure 3B where the α-collagen chain is seen to have a value of K_d expected for a 96 000 molecular weight (formerly) globular protein of typical amino acid composition. Likewise, the β-collagen chain, which is derived from two α-collagen chains by cross-linking near the chain ends (Fietzek & Kühn, 1976), behaves on gel filtration as if it had an apparent molecular weight of 181 000, slightly less than the real value of about 192 000. This is consistent with overlap

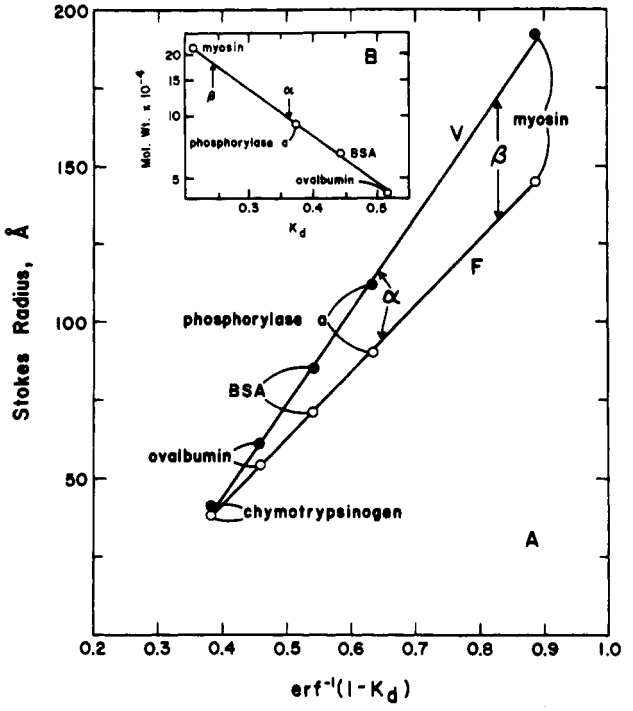

FIGURE 3: Gel filtration chromatography of NaDodSO$_4$–protein complexes. In (A), the open circles represent Stokes radii of standard NaDodSO$_4$–protein complexes obtained from viscosity measurements; the filled circles represent those obtained from frictional coefficients. The results for NaDodSO$_4$ complexes of the α- and β-collagen chains are indicated by arrows. In (B), standard NaDodSO$_4$–protein complexes are indicated by open circles and collagen chains are indicated by arrows.

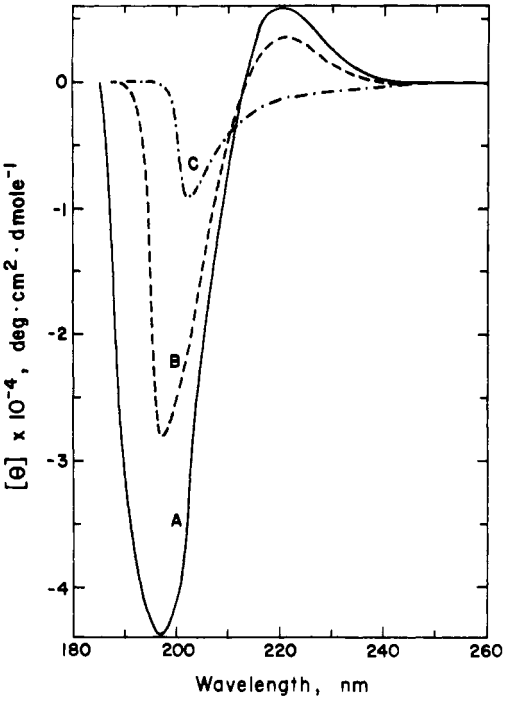

FIGURE 4: Circular dichroism spectra of calf skin collagen in various solvents. (A) 0.05% acetic acid; (B) 0.1 M sodium phosphate and 0.01% NaDodSO$_4$, pH 6.0; (C) 0.1 M sodium phosphate and 0.1% NaDodSO$_4$, pH 6.0. All spectra were measured at room temperature, approximately 25 °C.

of the chains due to the cross-links occurring between regions on each chain a bit distant from the end, so that the effective number of residues that determines hydrodynamic properties is less than twice that of an α chain.

Circular Dichroism. The results are shown in Figure 4. Curve A is that for native tropocollagen in 0.5% acetic acid. A major change in the CD spectrum was found when the protein was dissolved in 0.01% NaDodSO$_4$ and 0.1 M sodium phosphate, pH 6.0 (Figure 4, curve B), and further change occurred when the NaDodSO$_4$ concentration was increased to 0.1% (Figure 4, curve C). No additional change was found, however, when the NaDodSO$_4$ concentration was increased to 1.0%. Our binding studies indicate that maximal binding of detergent took place in 0.1% NaDodSO$_4$, consistent with a detergent-induced conformation change which was incomplete in 0.01% NaDodSO$_4$, and complete at detergent concentrations equal to or greater than 0.1%. Similar results were obtained for purified a_1-collagen which had first been allowed to fold into a triple helix before addition of NaDodSO$_4$. Analysis of the CD spectrum by use of the procedure of Chen et al. (1974) yielded an apparent α-helix content of only 6% at saturating levels of NaDodSO$_4$. This low value is consistent with the high content of proline and hydroxyproline residues and their rather uniform distribution along each polypeptide chain. The Chen et al. (1974) curve-fitting procedure also indicated apparent values of 10% for β structure and 84% for random coil, but the exact significance of the latter value is uncertain because the reference spectra were derived from spectra of globular protein models and the term "random coil" refers to that portion of a globular protein which possesses neither α helix nor β structure. Furthermore, the model proteins contained relatively small amounts of random coil; thus, the reference spectrum for that structure is somewhat inaccurate. In addition, a high proline and hydroxyproline content and the related hindrance to rotation about these residues may result in a CD spectrum which is unique to the NaDodSO$_4$–collagen complex. Madison (1969), for example, showed that a diamide, acetyl-L-proline dimethylamide, has optical properties very similar to those of poly(L-proline) and collagen, indicating the importance of short-range interactions and the effects of restricted rotation. Thus, there is no a priori reason to interpret the CD results to mean that collagen chains form extremely flexible random coils in aqueous NaDodSO$_4$, such as those formed by proteins in concentrated guanidine hydrochloride.

Table I: Binding of Sodium Dodecyl Sulfate to Collagen

protein	method of determination	concn of unbound NaDodSO$_4$ (mM)	ionic strength (Γ/2)	g of NaDodSO$_4$ bound per g of protein
collagen	equil dialysis	3.47	0.26	1.4 ± 0.2
collagen	equil dialysis	34.7	0.26	1.3
BSA	equil dialysis	3.47	0.26	1.3 ± 0.1
BSA	equil dialysis	34.7	0.26	1.3
ovalbumin	equil dialysis	3.47	0.26	1.2 ± 0.1
ovalbumin	equil dialysis	34.7	0.26	1.2
α chains	gel chromatography	3.47	0.26	1.1 ± 0.2

Binding of NaDodSO$_4$ to Collagen. The binding of NaDodSO$_4$ to collagen, under those conditions required for attaining a saturation level of 1.4 g of NaDodSO$_4$ per g of protein for most water-soluble proteins, was measured by two different methods. The results obtained by equilibrium dialysis at both 0.1 and 1.0% NaDodSO$_4$ showed collagen, as well as two control proteins, BSA and ovalbumin, to bind the nominal 1.4 g of NaDodSO$_4$ per g of protein (Table I). At these conditions, both concentrations of NaDodSO$_4$ were well above its critical micelle concentration, and therefore the dialysis was permitted to proceed for 4 weeks to assure the attainment of

equilibrium. However, when the concentration of $NaDodSO_4$ inside the dialysis bag (1% $NaDodSO_4$), at the beginning of the experiment, was greater than the concentration of $NaDodSO_4$ in the dialysate (0.1% $NaDodSO_4$), equilibrium was not reached even after 5 weeks of dialysis (data omitted). Since this sort of problem could introduce serious error, the binding determination was also conducted using a gel chromatography technique. In this method, column chromatography was utilized for the separation of the protein–detergent complex from the excess free and micelle $NaDodSO_4$. The results of this experiment are also shown in Table I. While the results are less accurate, they confirm the equilibrium dialysis data. Thus, collagen binds at least 110% and most probably 140% of its weight in $NaDodSO_4$, so that the abnormal migration on $NaDodSO_4$–polyacrylamide gel electrophoresis is probably not due to atypical binding of detergent.

Discussion

We have found in this study that the α chains of collagen, which behave anomalously on $NaDodSO_4$–polyacrylamide gel electrophoresis, bind the same amount of $NaDodSO_4$ on a weight basis as proteins behaving normally and that, perhaps surprisingly, the $NaDodSO_4$ complexes of α chains and β chains have virtually the same gel filtration behavior expected for $NaDodSO_4$ complexes of (formerly) globular protein counterparts of the same molecular weight provided that Sepharose CL-4B is used. Our circular dichroism results indicate that the $NaDodSO_4$–collagen complex is essentially devoid of α helix, while many, but not all, $NaDodSO_4$–protein complexes have considerable α-helix content. The collagen α chains have falsely high apparent molecular weights of 133000 ± 5000 and 122000 ± 3000 for α_1-collagen and α_2-collagen, respectively, regardless of gel concentration, when determined by direct comparison of their electrophoretic mobilities with those of $NaDodSO_4$ complexes of (formerly) globular proteins. However, use of the retardation coefficient, K_R, a conformational parameter obtained from plots of log mobility against gel concentration, led to molecular weight values of 110000 and 106000 for α_1-collagen and α_2-collagen, respectively, which, while still too high, are closer to the true value of 96000.

Success at rationalizing these results is limited by the current state of the theory of transport of macromolecules through gel networks. While considerable advances have been made, the theory does not have a basis as firm as that which underlies transport through fluids and led, for example, to important results in sedimentation, diffusion, and viscous flow. Nevertheless, guidelines are provided which are of great value in obtaining empirical relationships between either electrophoretic mobility or distribution coefficients and molecular properties of biopolymers. Turning first to the anomalously high apparent molecular weights of collagen chains as determined by $NaDodSO_4$–polyacrylamide gel electrophoresis, the explanation, offered by Furthmayr & Timpl (1971), that the causative factor is a local conformation effect related to the unusual amino acid sequence and composition is plausible. However, there is an alternative explanation. Collagen contains 33 mol % glycine and about 10 mol % each of proline and hydroxyproline. This results in an unusually low mean residue weight of 91.6 for calf skin collagen. On the other hand, the standard proteins used in our electrophoresis experiments have mean residue weights in the range 113.6–116.6 with an average value of 115.0 ± 1.2 (average difference). Thus, collagen has 26% more residues per unit molecular weight than the standard proteins, a value quite close to the percentage by which the apparent molecular weights are too high. Thus, it appears that the electrophoretic mobility is better correlated with the number of residues, N, in a polypeptide chain than with the molecular weight. With this in mind, the calibration curves shown in Figure 1 were recast as log N vs. R_f. An example of such a revised calibration curve is shown in Figure 5. From this analysis, the α_1-collagen chain has an apparent number of residues equal to 1156, and the α_2-collagen chain has 1060, compared to the value 1052 obtained from the amino acid sequence (Fietzek & Kühn, 1976). Thus, it appears that polyacrylamide gel electrophoresis of $NaDodSO_4$–protein complexes is better suited to determine the degree of polymerization rather than molecular weight and that the "anomalous" behavior of collagen chains could be, at least in part, a result of data handling.

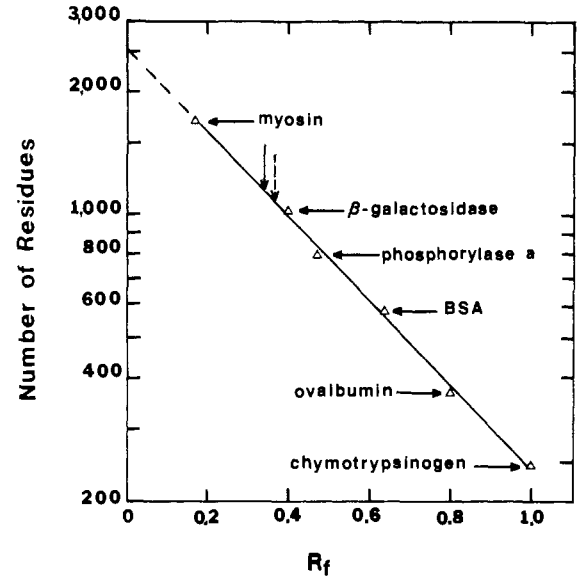

FIGURE 5: Replot of the data in Figure 1 for 5% acrylamide gels using the number of residues per chain rather than molecular weight. The best fit to the calibration curve is a straight line (log N = 7.844 − 2.366R_f); the correlation coefficient is −0.999.

The use of the number of residues in hydrodynamic theory has, in general, a firmer basis than the use of molecular weights, especially when comparing heteropolymers such as proteins, and it has been used successfully in studies on reduced unfolded proteins [e.g., Tanford (1968); Mann & Fish (1972), and Mann & Fass (1973)]. In spite of its apparently successful use, however, in interpreting our electrophoresis results, we should note that collagen contains large amounts of imino acid residues characterized by an unusual geometry. In addition, there is restricted rotation of polypeptide chains about imino acid residues. Thus, there was no overriding a priori reason to expect such an excellent correlation between number of residues and electrophoretic mobility, and the relationship should be regarded as empirical, especially since it does not help explain the observations of Furthmayr & Timpl (1971) that the α_1 chain of collagen and its cyanogen bromide fragments require a different calibration curve than the α_2 chain of collagen and its cyanogen bromide fragments. This in no way, of course, diminishes the value of polyacrylamide gel electrophoresis in characterizing collagenous polypeptides, since an accurate value for the number of residues has the same importance as an accurate molecular weight.

Keeping in mind the goal of using $NaDodSO_4$–polyacrylamide gel electrophoresis to determine molecular weights of collagen chains, we used the method of Frank & Rodbard (1975), utilizing the retardation coefficient, K_R, to interpret our electrophoresis results, since they had found improved

molecular weight estimates for rhodopsin by its use. We obtained molecular weights for α_1- and α_2-collagen that were only 10–15% higher than the real value, while those obtained by comparison of R_f values with those of standard protein–$NaDodSO_4$ complexes were 39 and 27% higher for α_1- and α_2-collagen, respectively. However, the values obtained by Frank and Rodbard's procedure may still have high errors associated with them because of error magnification when K_R values are obtained from the slopes of R_f vs. T curves, subsequent construction of a molecular weight vs. K_R curve, and some possible error in molecular weights of some of the standards used. Frank and Rodbard used this method for rhodopsin and found a molecular weight of 32 700 ± 5000, and we would thus expect an error of perhaps ±15 000 for each of the collagen chains. In addition, the accuracy of determining the number of residues per chain for α_1- and α_2-collagen was decreased; values for N of 942 and 909 were obtained from a plot of N vs. K_R. Thus, we turned to gel filtration chromatography in hopes of finding a more accurate method of determining molecular weights of collagenous chains in mixtures.

Effects of conformation on rate of migration may occur in gel filtration chromatography and seriously affect molecular weight estimations. Nozaki et al. (1976) found that $NaDodSO_4$–protein complexes and the asymmetric proteins, myosin and fibrinogen, had elution volumes on filtration through Sepharose 4B that were lower than those of native globular proteins with identical Stokes radii, when the Stokes radius was greater than 40 Å. Thus, a different calibration curve would be obtained for asymmetric particles than for spheres. Nozaki et al. (1976) found a linear relationship between Stokes radius and $\text{erf}^{-1}(1 - K_d)$ for globular proteins but, interestingly, a parabolic relationship for the asymmetric particles. This suggests a residual conformation effect within the latter series of particles, i.e., perhaps end-on insertion into the gel, which could produce a larger effect on elution volume for larger, more asymmetric particles than for smaller, more symmetrical ones. We found, however, that this factor is no longer important when Sepharose CL-4B, a cross-linked Sepharose with higher exclusion limits, was used. The Stokes radius–$\text{erf}^{-1}(1 - K_d)$ relationship was linear for $NaDodSO_4$–protein complexes as small as that of chymotrypsinogen to that as large as the $NaDodSO_4$–myosin complex.

Gel filtration chromatography in the presence of $NaDodSO_4$ has already been used to determine approximate molecular weights of collagenous chains from basement membranes (Grant et al., 1973; Clark & Kefalides, 1979), but using α- and β-collagen chains as standards and 6% agarose (Bio-Gel A) as the filtration medium. Our results support the validity of the method and indicate that demonstrably accurate results can be obtained if additional standard protein–$NaDodSO_4$ complexes are used and a linear calibration curve, i.e., log molecular weight vs. K_d, is obtained. The use of our gel filtration procedure may be preferable for procollagen-like chains, which contain noncollagenous regions, since it is not based on the assumption that these should behave like authentic collagen chains.

It is of interest, and perhaps fortuitous, that a collagen–$NaDodSO_4$ complex would have the same K_d value as a $NaDodSO_4$ complex of a (formerly) globular protein of the same molecular weight rather than one with the same number of residues. This further demonstrates the complexity of transport processes through pores of gels and the empirical nature of related equations. The gel filtration results yield a Stokes radius for the $NaDodSO_4$–α chain complex of collagen slightly greater than that found for the detergent complex of phosphorylase *a*, a protein which has 797 residues, compared with 1052 for the α_1-collagen chain and presumably very close to 1052 for the α_2 chain. Similarly, the β chain–$NaDodSO_4$ complexes behave as complexes of proteins with 1600 residues rather than 2100. Substantially higher Stokes radii would have been observed if the number of residues was the sole or major determinant of the size and gross shape of the complexes. There are, however, two other factors which affect the hydrodynamic properties of protein–$NaDodSO_4$ complexes: the amount of $NaDodSO_4$ bound and the detailed conformation of the polypeptide backbone. The amount of $NaDodSO_4$ bound is essentially constant among the proteins considered in our study, including collagen, and need not be considered further. The last factor is dependent on rotation about single bonds connected to α-carbon atoms. With respect to backbone conformation, Mattice et al. (1976) measured the CD spectra of $NaDodSO_4$ complexes of 15 proteins and found that they could be interpreted in terms of contributions from α-helical and randomly coiled regions. Values of the reduced mean residue ellipticity at 222 nm ranged from −1400 to −15 000 (deg cm^2)/dmol, consistent with a range for apparent α-helix content from 5 to 50%. In spite of this variation in apparent α-helix content, most of the complexes, including those which behave normally on $NaDodSO_4$–polyacrylamide gel electrophoresis, appear to have a common gross shape, as indicated by a regular variation in hydrodynamic properties with particle mass. Thus, α-helix content per se is not a requirement for normal behavior on gel filtration of $NaDodSO_4$–polyacrylamide gel electrophoresis. Mattice et al. presented a model for $NaDodSO_4$–protein complexes according to which they were somewhat flexible, with the overall shape of a random coil. Two other proposed models differ from that of Mattice et al. in some details but agree with it in the incorporation of flexibility and the lack of need of α-helical folding. These are the necklace model of Shirahama et al. (1974) and the deformable ellipsoid model of Wright et al. (1975), which, according to the authors, would have the same hydrodynamic properties as the necklace model. It is clear that the conformation varies in detail among the (formerly) globular protein complexes used as standards, and the fact that they behave as a family with respect to gross shape indicates some statistical averaging process involving the three factors cited above. Likewise, some unknown combination of these factors causes collagen–$NaDodSO_4$ complexes to have comparable overall dimensions as their counterparts derived from typical globular proteins of the same molecular weight. Similar explanations would result from consideration of the other "flexible" models.

Collagen chains, with their unusual repeating sequence $[(\text{Gly-X-Y})_n$, where X is often Pro and Y is frequently Hyp or Pro] and high proline and hydroxyproline content, might be expected to always have an overall shape which differs significantly from that of more typical proteins. This, however, is not supported by reports in the literature on the hydrodynamic properties of denatured collagen and related polypeptides. While collagen forms fibrous structures in the native state, this is primarily a result of association of chains to form triple helices. Heat-denatured collagen, for example, adapts the flexible gelatin conformation (Harrington & von Hippel, 1961). Furthermore, the collagen-related polypeptide poly(L-proline) II occurs in solution as a flexible coil in which there are rigid helical segments which can rotate freely, as units, about disordered parts of the chain (Carver & Blout, 1967).

Likewise, β-casein B, which contains 17 mol % proline (Grosclaude et al., 1972), is a random coil in dilute salt solution at low temperatures, as shown by the fact that it has the same intrinsic viscosity under those conditions as in 6 M guanidine hydrochloride and has the expected intrinsic viscosity for a random coil (Noelken & Reibstein, 1968). Thus, the primary structure of collagen chains may cause them to form flexible structures under certain conditions. This could allow formation of $NaDodSO_4$ complexes similar in hydrodynamic properties to those derived from globular proteins of the same mass.

In conclusion, our results resolve some problems in methodology and provide an empirical basis for determining the number of residues in polypeptide chains by polyacrylamide gel electrophoresis in the presence of $NaDodSO_4$ and the molecular weight by gel filtration of protein–$NaDodSO_4$ complexes. These methods can be applied with equal validity to collagen chains and procollagen-like polypeptides, as well as noncollagenous polypeptides.

References

Ackers, G. K. (1967) *J. Biol. Chem. 242*, 3026–3034.

Carver, J. P., & Blout, E. R. (1967) in *Treatise on Collagen* (Ramachandran, G. N., Ed.) Chapter 9, Academic Press, New York.

Castellino, F. J., & Barker, R. (1968) *Biochemistry 7*, 2207–2217.

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

Clark, C. C., & Kefalides, N. A. (1979) in *Cellular and Biochemical Aspects in Diabetic Retinopathy* (Regnault, F., & Duhalt, J., Eds.) p 121, North-Holland Publishing Co., Amsterdam.

Darnall, D. W., & Klotz, I. M. (1975) *Arch. Biochem. Biophys. 166*, 651–682.

Emerson, M. F., & Holtzer, A. (1967) *J. Phys. Chem. 71*, 1898–1907.

Feldhoff, R. C., & Peters, T., Jr. (1975) *Biochemistry 14*, 4508–4514.

Ferguson, K. A. (1964) *Metabolism 13*, 985–1002.

Fietzek, P. P., & Kühn, K. (1976) *Int. Rev. Connect. Tissue Res. 7*, 1–60.

Fowler, A. V., & Zabin, I. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 1507–1510.

Frank, R. N., & Rodbard, D. (1975) *Arch. Biochem. Biophys. 171*, 1–13.

Freytag, J. W., Dalrymple, P. N., Maguire, M. H., Strickland, D. K., Carraway, K. L., & Hudson, B. G. (1978) *J. Biol. Chem. 253*, 9069–9074.

Furthmayr, H., & Timpl, R. (1971) *Anal. Biochem. 41*, 510–516.

Grant, M. E., Schofield, J. D., Kefalides, N. A., & Prockop, D. J. (1973) *J. Biol. Chem. 248*, 7432–7437.

Grosclaude, F., Mahe, M. F., Mercier, J. C., & Ribadeau-Dumas, B. (1972) *Eur. J. Biochem. 26*, 328–337.

Harrington, W. F., & von Hippel, P. H. (1961) *Adv. Protein Chem. 16*, 1–138.

Hartree, E. F. (1972) *Anal. Biochem. 48*, 422–427.

Hudson, B. G., & Spiro, R. G. (1972) *J. Biol. Chem. 247*, 4229–4238.

Hung, C. H., Ohno, M., Freytag, J. W., & Hudson, B. G. (1977) *J. Biol. Chem. 252*, 3995–4001.

Hung, C. H., Noelken, M. E., & Hudson, B. G. (1979) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 38*, 816.

Huszar, G., & Elzinga, M. (1971) *Biochemistry 10*, 229–236.

Kefalides, N. A. (1973) *Int. Rev. Connect. Tissue Res. 6*, 63–104.

Lewis, M. S., & Piez, K. A. (1964) *Biochemistry 3*, 1126–1131.

Madison, V. (1969) Ph.D. Thesis, University of Oregon, Eugene, OR.

Mann, K. G., & Fish, W. W. (1972) *Methods Enzymol. 26*, 28.

Mann, K. G., & Fass, D. N. (1973) *Adv. Chem. Ser. No. 125*, 310.

Mattice, W. L., Riser, J. M., & Clark, D. S. (1976) *Biochemistry 15*, 4264–4272.

Neuberger, A., & Marshall, R. D. (1966) in *Glycoproteins: Their Composition, Structure, and Function* (Gottschalk, A., Ed.) p 300, Elsevier, London.

Noelken, M., & Reibstein, M. (1968) *Arch. Biochem. Biophys. 123*, 397–402.

Nozaki, Y., Schechter, N. M., Reynolds, J. A., & Tanford, C. (1976) *Biochemistry 15*, 3884–3890.

Perry, S. V. (1955) *Methods Enzymol. 2*, 582–588.

Piez, K. A., & Gross, J. (1960) *J. Biol. Chem. 235*, 995–998.

Piez, K. A., Eigner, E. A., & Lewis, M. S. (1963) *Biochemistry 2*, 58–66.

Reynolds, J. A., & Tanford, C. (1970) *J. Biol. Chem. 245*, 5161–5165.

Rodbard, D., & Chrambach, A. (1971) *Anal. Biochem. 40*, 95–134.

Schroeder, W. A., Shelton, J. R., Shelton, J. B., Roberson, J. B., & Apell, G. (1969) *Arch. Biochem. Biophys. 131*, 653–655.

Seery, V. L., Fischer, E. H., & Teller, D. C. (1970) *Biochemistry 9*, 3591–3598.

Sevilla, C. L., & Fischer, E. H. (1969) *Biochemistry 8*, 2161–2171.

Shirahama, K., Tsujii, K., & Takagi, T. (1974) *J. Biochem. (Tokyo) 75*, 309–319.

Smillie, L. B., Furka, A., Nagabhushan, N., Stevenson, K. J., & Parkes, C. O. (1968) *Nature (London) 218*, 343–346.

Steinhardt, J., & Reynolds, J. A. (1969) *Multiple Equilibria in Proteins*, Academic Press, New York.

Tanford, C. (1968) *Adv. Protein Chem. 23*, 121–232.

Tanford, C., Nozaki, Y., Reynolds, J. A., & Makino, S. (1974) *Biochemistry 13*, 2369–2376.

Ullmann, A., Goldberg, M. E., Perrin, D., & Monod, J. (1968) *Biochemistry 7*, 261–265.

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

Wright, A. K., Thompson, M. R., & Miller, R. L. (1975) *Biochemistry 14*, 3224–3228.