THE ISOLATION AND PROPERTIES OF SOME SOLUBLE PROTEINS FROM WOOL

VI. THE PHYSICOCHEMICAL PROPERTIES OF SCMKB2

By J. M. Gillespie* and B. S. Harrap*

[*Manuscript received August* 26, 1962]

Summary

The molecular weight of a high-sulphur protein (SCMKB2) from Merino wool has been determined by the Archibald technique and by light scattering, values of 22,600 and 22,100, respectively, being obtained. Optical rotatory dispersion data show that in aqueous solution the protein behaves as a random coil. This is consistent with the frictional properties of the molecule as deduced from its sedimentation coefficient and intrinsic viscosity. The protein appears to have one *N*-terminal arginyl end group; since it also contains two acyl groups per 22,000 molecular weight unit, the possibility of a multi-chain structure for the protein unit has been considered. However, no reduction in molecular weight could be achieved in any of several disaggregating solvents.

I. Introduction

In Part V of this series (Gillespie 1963) the preparation and fractionation from Merino wool of a reduced and carboxymethylated protein (SCMKB2) having a high sulphur content was described. This paper reports the physicochemical characterization of this protein with respect to molecular weight, shape, and conformation.

It has recently been shown (Gillespie, Harrap, and Inglis, unpublished data) that this protein contains two acyl groups per molecular weight of 22,000. Although the residues to which these groups are attached have not been positively identified they are not liberated at 37°C at pH < 10. On this basis they are more likely to be *N*-acyl than *O*-acyl substituents (Herriott 1935). Since amino acid analysis of SCMKB2 (Gillespie 1963) indicated the absence of lysine, *N*-acyl groups must either be *N*-terminal or attached to the guanidino side-chains of arginine residues. *N*-acetyl arginyl residues have never been reported in proteins although an intramolecular acylation of arginine by the carboxyl group of an aspartic acid residue has recently been postulated to explain certain mechanisms (Erlanger 1960). Acetylated *N*-terminal residues have been reported in several proteins and peptides (e.g. O'Donnell, Thompson, and Inglis 1962; Titani, Narita, and Okunuki 1962). If they occur in SCMKB2 the possibility of a multi-chain unit must be considered. The behaviour of the protein in solvents known to favour disaggregation has therefore been studied.

II. Materials and Methods

(*a*) *Preparation of SCMKB2*

SCMKB2 was prepared from Merino 64's top (Wintoc MW127) by the procedure described in Part V (Gillespie 1963). Except where otherwise stated, measure-

* Division of Protein Chemistry, C.S.I.R.O. Wool Research Laboratories, Parkville, Vic.

ments were performed in a buffer consisting of 0·025M NaH_2PO_4, 0·025M Na_2HPO_4, and 0·3N NaCl at pH 6·8. For most physicochemical measurements the protein solutions were dialysed to equilibrium against the appropriate solvent; in these procedures Visking 18/32 cellulose tubing, which had been previously extracted with boiling distilled water and ethanol, was used. In experiments with formic acid, the protein was first extensively dialysed against distilled water, freeze-dried, and then dissolved in the solvent.

(*b*) *Reagents*

All chemicals used were of A.R. grade; distilled water was used throughout.

(*c*) *Methods*

(i) *Optical Rotatory Dispersion Determinations.*—A spectropolarimeter, which was built up from a photoelectric polarimeter (W. F. Stanley and Son), a quartz prism monochromator (Carl Leiss), and a high-pressure mercury arc (Philips SP 500) light source, was used for these determinations. Measurements were made at the following wavelengths: 578, 546, 486, 436, 405, 365, and occasionally, 334 mμ. Solutions were contained in water-jacketed 1-dm cells at 25$\pm$0·02°C.

(ii) *Viscosities.*—Viscosities were measured at 25$\pm$0·02°C with simple Ostwald capillary viscometers with a time of flow for water of approximately 200 sec.

(iii) *Ultracentrifuge Studies.*—These were performed with a Spinco model E ultracentrifuge equipped with phase-plate diagonal Schlieren optics and R.T.I.C. unit. Most Archibald runs were made with double-sector cells in order to obtain a solvent base line under the same conditions as those obtaining in the solution. When formic acid was used this procedure was not possible as double-sector cells made of "Kel–F" (a plastic material resistant to this solvent) were not available. However, redistribution of salt in this solvent (of density 1·226 g cm^{-3}) is very small and it is considered that no great error is introduced by extrapolating the plateau region back to the meniscus to provide a base line. With the exception of formic acid solutions, concentrations were determined in terms of arbitrary area units with a double-sector synthetic boundary cell. With formic acid solutions, c_0 was calculated from a measurement of Δn and the instrument constants (cf. O'Donnell and Woods 1962).

(iv) *Light-scattering Measurements.*—These were made at 546 mμ with a Brice–Phoenix series 1000 photometer which was calibrated with an opal disk of known transmittance supplied by the manufacturer. Solutions were clarified by pressure-filtration of each dilution through a sintered-glass filter (Jena, porosity 5) directly into the light-scattering cell. Adsorption of protein on the filter was checked by measurement of optical densities of each dilution at 278 mμ and found to be negligible ($< 2\%$).

(v) *Refractive Index.*—Increments in refractive index were measured at 546 mμ with a Brice–Phoenix differential refractometer at 25$\pm$0·02°C.

(vi) *Phosphorus Estimations.*—For the estimation of phosphorus, the procedure of Allen (1940) was used. The protein was hydrolysed with perchloric acid and then ammonium molybdate and amidol reagents were added. The absorption of the colour so produced was measured at 700 mμ.

(vii) *Identification of* N-*Terminal Groups.*—To identify *N*-terminal groups, the protein was coupled with fluorodinitrobenzene and hydrolysed for 4 or 16 hr with 6N HCl under reflux. As described by Fraenkel-Conrat, Harris, and Levy (1955), the ether-extractable dinitrophenyl (DNP)-amino acids were chromatographed with a toluene–pyridine–2-chloroethanol–ammonia system, and the water-soluble DNP-amino acids with a tertiary amyl alcohol–phthalate system.

(viii) *Protein Concentrations.*—These were determined by the semi-micro-Kjeldahl procedure, the nitrogen content of SCMKB2 being taken as 14·8% (Gillespie 1963).

III. Results and Discussion

(*a*) *Optical Rotatory Dispersion*

The plot of specific rotation $[\alpha]$ *v.* wavelength (λ) of SCMKB2 in phosphate–chloride buffer is shown in Figure 1(*a*). The data were treated according to the Moffitt and Yang (1956) equation:

$$m' = \frac{3}{n^2+2} \cdot \frac{M}{100}[\alpha] = a_0\left(\frac{\lambda_0^2}{\lambda^2-\lambda_0^2}\right)+b_0\left(\frac{\lambda_0^2}{\lambda^2-\lambda_0^2}\right)^2,$$

where m' is the reduced residue rotation, obtained by correcting $[\alpha]$ for the refractive index of the solvent n and the mean residue weight M (taken as 115); λ_0 is the dispersion constant (taken as 212 mμ (Moffitt and Yang 1956)); and a_0 and b_0 are parameters, the latter providing a measure of the helical content of the protein (Urnes and Doty 1961). The value of b_0 calculated from the data of Figure 1(*a*) was effectively zero (< -10). This value is characteristic of polypeptides and proteins in the random-coil conformation.

(*b*) *Viscosity*

The reduced viscosity ($\eta_{sp.}/c$) was measured over the protein concentration range 0·2–0·8% in phosphate–chloride buffer. The values obtained are plotted in Figure 1(*b*) and extrapolate to an intrinsic viscosity $[\eta]$ of 0·21 dl/g.

(*c*) *Sedimentation Coefficient*

Rates of sedimentation were measured at several different protein concentrations in phosphate–chloride buffer at a centrifuge velocity of 59,780 r.p.m. The sedimentation coefficient was calculated from the rate of movement of the top of the peak in the Schlieren diagram. A plot of $1/S$ against c is shown in Figure 1(*c*). The sedimentation coefficient at zero concentration, corrected to the standard state of water at 20°C, $S^0_{20,w}$ was found to be 1·55 S.

(*d*) *Molecular Weights in Aqueous Solution*

(i) *Archibald Method.*—Molecular weights were measured by the Archibald (1947) approach-to-equilibrium method as modified by Klainer and Kegeles (1955).

Only the meniscus position was employed (denoted by the subscript m) and the equation

$$M_{\text{app.}} = \frac{RT}{(1-\bar{v}\rho)\omega^2} \cdot \frac{(\mathrm{d}c/\mathrm{d}x)_m}{x_m c_m},$$

was used, with the following correction for the change in concentration at the meniscus:

$$c_m = c_0 - \frac{1}{x_m^2}\int_{x_m}^{X} x^2 \cdot \frac{\mathrm{d}c}{\mathrm{d}x} \cdot \mathrm{d}x,$$

(cf. O'Donnell and Woods (1962) for the definition of symbols). Measurements were made at various speeds in the vicinity of 29,000 r.p.m. and at several protein concentrations. In Figure 2(*a*) a plot of $1/M_{\text{app.}}$ *v.* c_m, the concentration at the meniscus, gives a value for the molecular weight of 22,600.

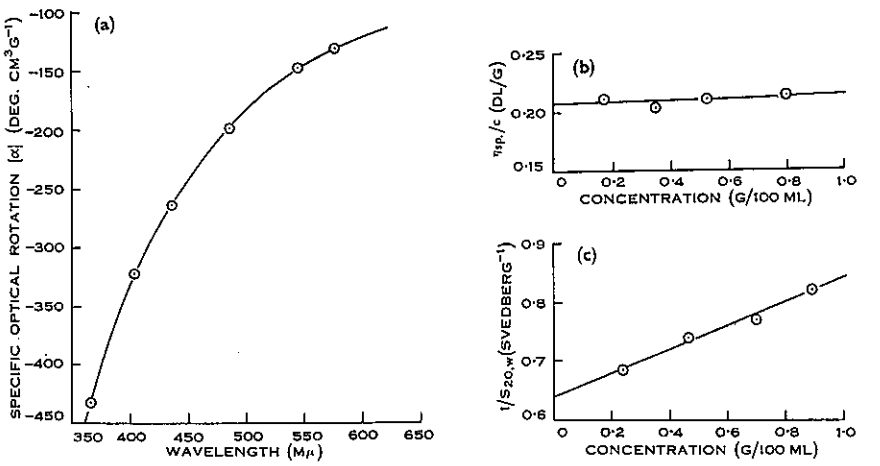

Fig. 1.—Specific optical rotation as a function of wavelength (*a*), and reduced viscosity (*b*) and reciprocal of sedimentation coefficient (*c*) as a function of concentration for SCMKB2 in phosphate–chloride buffer, pH 6·8.

(ii) *Light-Scattering Method.*—Light-scattering data have been treated according to the equation

$$1/M_{\text{app.}} = Kc/R_{90} = 1/M + 2Bc,$$

where c is the protein concentration (g/ml), K is the parameter relating the Rayleigh ratio R_{90} with molecular weight M, and B is a second virial coefficient (O'Donnell and Woods 1962). Measurements were made in phosphate–chloride buffer at several protein concentrations and the plot of $1/M_{\text{app.}}$ *v.* c is shown in Figure 2(*a*) together with the data obtained by the Archibald method. The dissymmetry of scattering of each concentration was essentially zero (less than 1·05). The extrapolated molecular weight was 22,100, in excellent agreement with the Archibald value. The

slopes of the two plots of $1/M_{app.}$ *v.* c are quite similar, as expected, but are much larger than is normally found for low molecular weight proteins in neutral solution, and indicates a high degree of intermolecular interaction. Since the protein is apparently in random-coil conformation this cannot be due to any marked asymmetry of the molecule. It is likely to be due to electrostatic interactions reflecting the high net negative charge on the molecule at pH 6·8. This has been acquired mainly as a result of the large number ($\simeq$ 1750 μmoles/g) of carboxymethyl groups introduced on alkylation of the reduced protein.

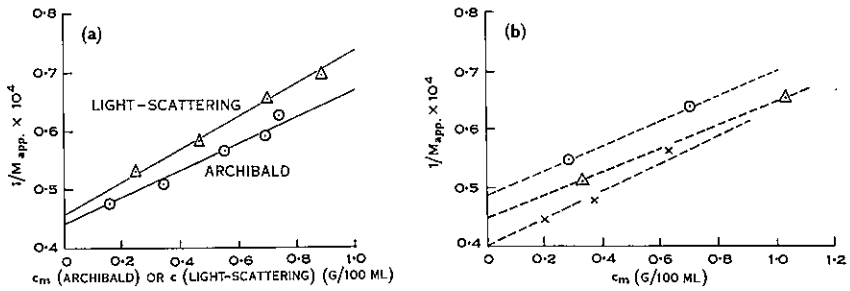

Fig. 2.—(*a*) Plots of $1/M_{app.}$ *v.* c_m (concentration at the meniscus) as determined by the Archibald technique and of $1/M_{app.}$ *v.* protein concentration c (light-scattering technique) for SCMKB2 in phosphate–chloride buffer, pH 6·8; (*b*) plots of $1/M_{app.}$ *v.* c_m as determined by the Archibald technique for SCMKB2 in 10M urea (0·2M with respect to KCl) at 35°C (⊙), in 2 : 1 v/v acetic acid–water (0·3M with respect to NaCl) at 25°C (△), and in formic acid (0·3M with respect to KCl) at 25°C (×).

(*e*) *End Groups*

Only traces of ether-soluble DNP-amino acids were detected. However, a major spot corresponding to DNP-arginine was found in the water-soluble fraction. In order to correct for operational losses a sample of DNP-arginine was used as a control for all operations from the hydrolysis on. Correcting for losses, a value of 1·06 mole DNP-arginine per mole (22,000) of protein was found, corresponding to one *N*-terminal arginine residue.

(*f*) *Molecular Weight Measurement in Disaggregating Solvents*

The possibility that the 22,000 molecular weight unit consists of two chains on the basis of two *N*-terminal acyl groups has already been considered (see Introduction). If we consider as well the *N*-terminal arginine this means that the protein may even consist of three polypeptide chains. These chains could be held together by non-covalent forces, as has been found to occur with several proteins such as insulin, hemoglobin, and β-lactoglobulin.

Attempts have been made to break possible non-covalent links by the use of solvents which disaggregate the reduced and alkylated low-sulphur wool proteins (Harrap 1963) and reduced and alkylated bovine serum albumin (Harrap and Woods, unpublished data). The solvents used were 10M urea which was 0·2M with respect to KCl, 2 : 1 (v/v) acetic acid–water, 0·3M with respect to NaCl, and formic acid

which was 0·3M with respect to KCl. The temperature of the first solvent was 35°C and the latter two 25°C. Since only evidence of disaggregation was being sought, measurements were not made at a large number of protein concentrations in each solvent. Clearly, if any major disaggregation occurred, the decrease in the extrapolated molecular weight would be apparent. The values of $1/M_{\text{app.}}$ *v.* c_m, as determined by the Archibald technique, in the three solvents are shown in Figure 2(*b*); quite clearly no significant disaggregation occurred under these conditions. The small differences between the values for molecular weight in the different solvents could be due to differences in the partial specific volume of the protein in different solvents.

(*g*) *Concluding Remarks*

From the hydrodynamic data given previously the shape factor β may be calculated from the Scheraga and Mandelkern (1953) equation:

$$\beta = \{S^0_{20,w}.[\eta]^{\frac{1}{3}}.N.\eta_0\}/M^{\frac{2}{3}}(1-\bar{v}\rho),$$

The value of β so obtained using data for the phosphate–chloride buffer was $2{\cdot}5\times10^6$. It is not possible from this value to distinguish between a random-coil polymer ($\beta = 2{\cdot}55\times10^6$ (Mandelkern and Flory 1952)) or a prolate ellipsoid with an axial ratio of about 14 : 1 (Scheraga and Mandelkern 1953). However, the optical rotatory dispersion data are more consistent with the random-coil conformation. This is also made more likely by the high content of proline in this protein (Gillespie 1963) which would inhibit development of a large degree of α-helix (Urnes and Doty 1961).

The agreement between the molecular weights as determined by the Archibald method and by light scattering suggests that the protein possesses a high degree of homogeneity with respect to molecular weight. Values obtained by light scattering are very sensitive to the presence of any high molecular weight contaminants.

The inability to reduce the molecular weight by the use of disaggregating solvents suggests that the protein molecule may consist of a single chain. If this is indeed the case, and since the lysine content is close to zero, the acyl groups must be attached to side-chain functional groups. The distinction between *O*-acyl and *N*-acyl groups made by Herriott (1935) is somewhat arbitrary and it is possible that we are here dealing with particularly stable *O*-acyl groups. As mentioned previously, another, somewhat unlikely, possibility is that the acyl groups are attached to the guanidino side-chains of arginine residues.

We have also considered that three chains may be held together by covalent cross-links. Of these, amide cross-links would seem to be ruled out by the absence of lysine residues. A determination of the phosphorus content of the protein gave a value of $< 0{\cdot}3$ mole phosphorus per mole protein, which is far below that necessary for phosphate diester cross-links to be significant. Ester cross-links between the β- or γ-carboxyls of aspartic or glutamic acids and the hydroxyl groups of serine, threonine, and tyrosine must also be considered. To test this possibility the protein was heated with hydroxylamine at 40°C for 2 hr to convert possible ester cross-links to the corresponding hydroxamic acids (Crewther and Dowling, personal com-

munication). The protein was then equilibrated against the phosphate–chloride buffer and a single Archibald run made. A value of 16,800 at a c_m of 0·4% was obtained which was not sufficiently different from that of the untreated protein to indicate significant splitting into separate chains.

IV. Acknowledgments

Thanks are due to Miss C. M. Tomlinson and Mr. D. J. A. Evans for capable technical assistance.

V. References

Allen, R. J. C. (1940).—*Biochem. J.* **34**: 858.

Archibald, W. J. (1947).—*J. Phys. Chem.* **51**: 1204.

Erlanger, B. I. (1960).—*Proc. Nat. Acad. Sci., Wash.* **46**: 1430.

Fraenkel-Conrat, H., Harris, J. I., and Levy, A. L. (1955).—"Methods of Biochemical Analysis." Vol. 2. (Interscience Publishers, Inc.: New York.)

Gillespie, J. M. (1963).—*Aust. J. Biol. Sci.* **16**: 241.

Harrap, B. S. (1963).—*Aust. J. Biol. Sci.* **16**: 231.

Herriott, R. M. (1935).—*J. Gen. Physiol.* **19**: 283.

Klainer, S. M., and Kegeles, G. (1955).—*J. Phys. Chem.* **59**: 952.

Mandelkern, L., and Flory, P. J. (1952).—*J. Chem. Phys.* **20**: 212.

Moffitt, W., and Yang, J. T. (1956).—*Proc. Nat. Acad. Sci., Wash.* **42**: 596.

O'Donnell, I. J., and Woods, E. F. (1962).—In "Modern Methods of Plant Analysis". (in press). (Springer-Verlag: Berlin.)

O'Donnell, I. J., Thompson, E. O. P., and Inglis, A. S. (1962).—*Aust. J. Biol. Sci.* **15**: 732.

Scheraga, H. A., and Mandelkern, L. (1953).—*J. Amer. Chem. Soc.* **75**: 179.

Titani, K., Narita, K., and Okunuki, K. (1962).—*J. Biochem., Japan* **51**: 350.

Urnes, P., and Doty, P. (1961).—*Advanc. Protein Chem.* **16**: 402.