THE CHAIN WEIGHT OF WOOL KERATIN*

by

W. R. MIDDLEBROOK**

Department of Textile Industries, University of Leeds (England)

INTRODUCTION

It is well known that the wool fibre is constructed from at least two types of cell. The cuticle is made from a large number of flat cells called scales which overlap each other and cover the exterior of the fibre. The interior or cortex consists of many long spindle-shaped cells, cemented together and totalling more than 98% by weight of the wool. Hence they are responsible for most of the chemical and physical properties, except those due to surface effects. Unlike the cuticular cells, the cortical cells exhibit a high degree of molecular orientation along the fibre axis. Recently, GRALÉN, LINDBERG AND PHILIP[1] discovered a thin membrane which is believed to cover the scales and to amount to not more than 0.1% of the total weight of the fibres. In coarse wool fibres, a medulla is found in the centre, consisting of amorphous material with large air spaces. However, the wool used in the present work was non-medullated.

Accepting the conventional view that keratin is an association of a large number of protein chains, cross-linked by a series of cystine sulphur linkages[2,3], it is possible for each chain to contain several half residues of cystine, linked with an equal number of half residues in many different chains, producing an extensive three-dimensional network. In this case, the use of the term 'molecular weight' in the usual sense is meaningless, as the size of the cross-linked polymer is only limited by the dimensions of the individual piece of keratin. Evenso, this indefinite cross-linking of many chains might link together groups of chains, which are repeated throughout the polymer. Rather than referring to 'molecular weight', we might refer to 'repeat unit weight'. On the other hand, this cross-linking might be confined to a limited number of protein chains, producing a molecule of high molecular weight, constructed of chains of smaller 'molecular weight' or rather 'average chain weight'. It is this latter value which is of greater interest, and to that end the following work has been carried out.

In view of the heterogeneous nature of the wool fibre, it might be thought that the determination of the molecular weight or repeat unit weight was well-nigh impossible. It should be realised, however, that the material from which the various parts of the fibre are constructed may be relatively homogeneous in the chemical sense. Again, lack of uniformity in chemical analysis of different samples of wool does not necessarily mean that they differ fundamentally from each other, because it is known that part of the cystine of wool can be converted into cysteic acid, owing to the action of sunlight and

* Presented as a paper at the 1st International Congress of Biochemistry (Cambridge, 1949).
** Present address: Department of Biomolecular Structure, University of Leeds.

References p. 562.

atmospheric oxygen. Similarly, the cystine can be converted into lanthionine by alkali[4], without main chain hydrolysis.

It must be emphasised here that the estimation of average chain weight is valid whether all the chains are chemically identical and of the same weight, or are of dissimilar constitution and have a statistical distribution of chain weights about a mean value. However, assuming that wool keratin is a relatively pure protein, a value for the molecular weight or repeat unit weight has also been obtained.

EXPERIMENTAL

A sample of 1-fluoro-2:4-dinitrobenzene (FDNB) was prepared from the corresponding chloro-compound by the method of GOTTLIEB[5]. Both these compounds were used to prepare 2:4-dinitrophenyl (DNP) derivatives of the amino-acids[6,7,8].

Chromatographic Technique

The silica gel was prepared from a sodium silicate of low iron content (C. 140. I.C.I.)[9,10]. After standing overnight in a large volume of dilute hydrochloric acid, the gel was transferred to a large Büchner funnel and washed with distilled water until the filtrate was free from chloride. The gel was dried for four days at 110°.

The columns were prepared from glass tubing of 1 cm internal diameter cut into 25 cm lengths. One end of these tubes was slightly constricted by rotating it for a short time in a bunsen flame. A disc of filter paper was inserted into the wide end of the tube and carefully pushed to the other end so as to lodge firmly against the constriction. Originally these discs were supported by perforated silver supports[11], but they were found to be unnecessary. The paper discs, which were cut from filter paper with the appropriate size of cork borer, were boiled in dilute hydrochloric acid to remove any metallic contamination, washed several times with distilled water, and then dried. So that they would not retain small amounts of DNP-acids, the discs were damped with 2 *N* HCl and squeezed between filter paper immediately before use.

In order to obtain good separation on the columns, it was essential to have them evenly packed. Dry silica gel (1 g) was weighed into a small beaker and 0.5 ml of the appropriate immobile phase added. After thoroughly mixing by grinding with the flat end of a specimen tube (the gel remaining a friable powder), a small volume of the mobile phase was added and the whole poured into the glass tube. The gel remaining in the beaker was poured into the tube by the addition of more solvent. A flat-ended glass rod was introduced down the full length of the glass tube, rotated rapidly and at the same time gradually withdrawn, so as to describe a flat spiral motion. The solvent was allowed to drain out and the gel was firmly packed down by means of a plunger consisting of a piece of 'cotton-wool' fixed on to one end of a length of glass tubing. If a 2 g or 3 g column were desired, then the process was repeated by the addition of not more than 1 g of silica at a time, otherwise an uneven column often resulted.

As most of the solvents moved very slowly down the column, they were accelerated by the application of a positive air pressure at the top of the column by means of a 'hand-bellows' fitted with a valve. Care had, however, to be exercised always to maintain a 'head' of solvent above the gel surface, so that air was not forced into the silica, thereby ruining the column. It was ascertained that the increased rate of flow had no perceptible effect on the *R* values. Before a column was ready for use it was 'consolidated' by the addition of a small amount of solvent and the application of air pressure until the column ceased to contract in length. This could be observed easily because a thin ring of gel was always left on the glass surface at the original height of the column. This ring was finally pressed on to the gel surface by the 'cotton-wool' plunger.

A small sample of the appropriate DNP-acid was dissolved in a minimal amount of solvent, then transferred to the top of the column by means of a pipette, the tip of which had been bent slightly sideways, so that the jet of solution was projected on to the side of the glass tube, rather than on to the surface of the gel. The procedure followed, was that devised by GORDON, MARTIN AND SYNGE for acetamido-acids[11], and modified for DNP-acids by SANGER[7]. As the *R* values varied with the temperature, sometimes it was found advisable to record the solvent temperature and to run the standards at the same time as the unknowns on duplicate columns, or immediately afterwards on the same column. For example, with DNP-glycine on an unbuffered chloroform column, *R* at 15° = 0.17, *R* at 20° = 0.24. Since the *R* values on unbuffered columns varied considerably, according to the amount of DNP-acid placed on the column, the amounts added were arbitrarily divided into three grades, lightly loaded (L.L.), medium loaded (M.L.), and heavily loaded (H.L.), by visual

References p. 562.

examination. Standard DNP-acids were loaded similarly to the unknowns. For example, with DNP-leucine on an unbuffered chloroform column, R (L.L.) = 0.59 and R (H.L.) = 0.87. Overloading of the columns was avoided, as the bands tend to 'tail' into one another.

Determination of R Values

(i) Unbuffered columns.

As most of the solvents used, and R values obtained, were similar to those of SANGER[7], the details are not given here, with the exception of the R values on chloroform columns, as these differed somewhat from SANGER's.

The chloroform was washed with a dilute solution of sodium bicarbonate, followed by four lots of distilled water and stored in a dark bottle, otherwise the presence of small amounts of polar material caused the DNP-acids to move fast.

TABLE I

Stationary phase — Water: Mobile phase — Chloroform

Acid	R values		
	L.L.	M.L.	H.L.
Dinitroaniline		Fast	
DNP-leucine	0.59 – 0.58*	0.62	0.87 – 0.78*
DNP-valine		0.64 – 0.54*	0.79 – 0.73*
DNP-phenylalanine		0.61 – 0.49 (20°)*	0.80 – 0.74*
		0.48 – 0.45 (15°)*	
DNP-proline		0.57	
DNP-alanine		0.43 (20°)	0.58 (20°)
		0.35 (15°)	
DNP-methionine		0.35	
DNP-glycine		0.24 (20°)	
		0.20 (17°)	
		0.17 (15°)	
Bis DNP-tyrosine		0.41	
Bis DNP-lysine		0.27	
DNP-tryptophan		0.18	
DNP-threonine		0.05	0.10
Bis DNP-cystine		0.00	

* When two values are given, the first refers to measurements taken near the top of the column, and the second to those taken near the bottom.

(ii) Buffered columns.

Considerable use was made of buffered columns in conjunction with several solvents[12]. When ether was used, it was first freed from peroxides by shaking with an acidified solution of ferrous ammonium sulphate, followed by several changes of distilled water, and stored in a dark bottle

A p_H 6.0 sodium phosphate buffer was prepared by dissolving 14.32 g $Na_2HPO_4.12H_2O$ and 9.36 g $NaH_2PO_4.2H_2O$ in water and making up to 100 ml. For every 1 g silica gel used, 0.5 ml buffer solution was added. Care was taken not to dilute this solution, as this resulted in a considerable change in p_H.

A p_H 7.2 buffer was prepared by dissolving 14.32 g $Na_2HPO_4.12H_2O$ and 1.56 g $NaH_2PO_4.2H_2O$ in water and making up to 100 ml. For every 1 g silica gel used, 0.5 ml this solution was added.

A p_H 10.7 buffer was prepared by dissolving 7.16 g $Na_2HPO_4.12H_2O$ and 7.60 g

References p. 562.

TABLE II

Stationary phase: p_H 6.0 buffer
Mobile phase:
5% propanol/cyclohexane (a)
33% ether/ligroin (b)

Acid	*R* values	
	a	b
DNP-leucine	—	0.32
DNP-valine	0.30	0.21
DNP-phenylalanine	0.25	0.11
DNP-methionine	0.15	—

TABLE III

Stationary phase: p_H 6.0 buffer
Mobile phase:
1% butanol/chloroform (a), ether (b)

Acid	*R* values	
	a	b
Dinitroaniline	1.3	fast
DNP-leucine	0.90	fast
DNP-valine	0.64	fast
DNP-phenylalanine	0.50	fast
DNP-methionine	0.65	fast
DNP-tryptophan	0.39	fast
DNP-proline	—	fast
DNP-alanine	0.22	0.95
Bis DNP-tyrosine	0.70	0.90
Bis DNP-lysine	0.60	0.65
DNP-glycine	0.03	0.51
DNP-threonine	—	0.32
DNP-serine	—	0.16
DNP-hydroxyproline	—	0.16
DNP-glutamic acid	—	0.07
DNP-aspartic acid	—	0.02
Bis DNP-cystine	0.00	0.02

TABLE IV

Stationary phase: p_H 7.2 buffer
Mobile phase: ether

Acid	*R* values
Dinitroaniline	fast
Dinitrophenol	fast
DNP-leucine	fast
DNP-valine	fast
DNP-phenylalanine	fast
DNP-methionine	0.65
DNP-proline	0.35
DNP-alanine	0.33
DNP-tryptophan	0.90
Bis DNP-tyrosine	0.75
Bis DNP-lysine	0.55
DNP-glycine	0.16

TABLE V

Stationary phase: p_H 10.7 buffer
Mobile phase: ether

Acid	*R* values
Dinitroaniline	fast
Dinitrophenol	0.62
DNP-leucine	0.70
DNP-valine	0.50
DNP-phenylalanine	0.54
DNP-methionine	0.32
DNP-tryptophan	0.44
DNP-alanine	0.12
Bis DNP-tyrosine	0.41
Bis DNP-lysine	0.16

$Na_3PO_4.12H_2O$ in water and making up to 100 ml. As the buffering capacity was poor and the silica gel was believed to contain strongly adsorbed acid which could not be removed on washing, the effective p_H of the aqueous phase was probably considerably less than 10.7.

Treatment of Wool

Several locks of wool, from which 1½" at the tip ends were removed and rejected, were taken from a Lincoln Hog fleece still in the grease. The wool was extracted with ether in a Soxhlet apparatus for approximately 24 hours, followed by a similar extraction with ethanol. After washing in several changes of distilled water to free the wool from sand and other insoluble matter, as well as water-soluble material that was also present, the washing was continued with *N*/1000 HCl until the p_H no longer altered on indefinite immersion. Washing was repeated in distilled water (p_H 4.5) until the p_H remained at this value on indefinite immersion. The wool was blotted between filter papers and dried in the laboratory at room temperature, then stored in a humidity room at 65% relative humidity and 22.2°, after which, the moisture content was determined on samples by

References p. 562.

CHIBNALL's method[13], except that the oven temperature was 105° instead of 102°. By this method it was possible to determine the dry weight of the bulk of the wool without actually subjecting it to a high temperature.

Samples of wool, each 0.2 g, were thoroughly wetted out in a solution of 0.2 g sodium bicarbonate in 5 ml water and 0.2 ml (0.35 g) FDNB in 10 ml ethanol was added. The samples were thoroughly mixed in stoppered flasks and the reaction was allowed to proceed for varying lengths of time. The samples were then washed in several changes of water, followed by ethanol, and finally left overnight in a litre of distilled water. After drying, the increases in weight were determined. On treating wool samples with the above solutions containing no FDNB, it was noticed that their p_H was 8.4 owing to the presence of a small amount of sodium carbonate in the bicarbonate. In the first series the p_H was 7.0, as presumably the carbonate was decomposed by the HF liberated. On carrying out this treatment at 22.2°, the reaction was not complete after 90 hours, as judged by the increase in weight of the samples. A second series was treated at 40°, and as there was no further increase in weight after 48 hours, this was selected as the most suitable condition for treatment of a 20 g lot of wool used for most of the experimental work.

As the treatment at 40° and p_H 7.0 took place under conditions similar to those in a physiological environment, there was little danger of main chain hydrolysis. To meet any possible objections, however, one sample of wool was treated for 48 hours at 40° and another for 120 hours at 22.2°. Both samples were hydrolysed in 5.7 *N* HCl and (by a method to be explained later) their ether-soluble fractions were separated and estimated colorimetrically. The colour intensities were within 5% of each other, therefore the concentrations of the DNP-acids in both cases were assumed to be the same, owing to an equal number of end-groups. This would not be the case if main chain hydrolysis took place.

Extent of Reaction

Before attempting to determine the end-groups, it was necessary to ascertain that the FDNB had penetrated the wool fibres and reacted with all available reactive groups. The extent of reaction of the ε-amino-groups of the lysine side-chains was taken to be an indication of the penetration. Accordingly, a 0.6 g sample of FDNB treated wool was hydrolysed in 5.7 *N* HCl under reflux for 24 hours, extracted by five lots of ether and the residue made up to 100 ml. A 1 ml portion was then evaporated in vacuo, the residue was taken up in a minimal amount of 30% butanol/chloroform and run on a silica gel column with water as the immobile phase. A strong band $R = 0.18$ was obtained, and on running this substance with 66% methylethylketone/ether on a silica gel column with water as the immobile phase, a band was obtained $R = 0.27$, corresponding to an authentic sample of εN-DNP-lysine. The substance was made up to volume in *N* HCl and estimated colorimetrically. A sample of untreated wool was hydrolysed for 24 hours and the lysine content determined[14] by a modification of PORTER AND SANGER's copper complex technique for the preparation of εN-DNP-lysine[8]. By this method, it was found that 97.2% of the εamino-groups of the lysine in the wool had reacted with the FDNB, and thus it was reasonable to assume that it had diffused right into the wool fibres and combined with all available amino-groups.

Identification of the End-Groups

This was accomplished by separation and identification of the DNP-acids liberated

References p. 562.

after hydrolysis for 24 hours in 5.7 *N* HCl under reflux. Several methods of separation were used, but the following was considered to be the most satisfactory. The hydrolysate from a 1 g sample of treated wool was five times extracted with an equal volume of ether. After evaporation of the ether, a yellow residue was obtained and referred to, as the ether-soluble fraction. The ether-insoluble fraction was only found to contain εN-DNP-lysine as well as free amino-acids. The ether-soluble fraction was run on a 2 g p_H 6.0 column using aqueous ether as the solvent. A strong fast band was obtained, followed by a band $R = 0.9$ and these were collected together. Further bands $R = 0.51$, 0.32, 0.16, 0.07, 0.02 were obtained, and identified by running on at least three different columns with various solvents in parallel with authentic samples of DNP-acids. The bands were found to be due to DNP-glycine, -threonine, -serine, -glutamic acid, -aspartic acid.

The mixture containing the two fast bands was run on a p_H 7.2 column with chloroform as the mobile solvent. A fast band was collected and subsequently shown to consist of dinitroaniline, dinitrophenol and a decomposition product. The remaining coloured material at the top of the column was developed by the use of aqueous ether as the mobile solvent. A fast band was obtained and shown to be DNP-valine, by running on a p_H 10.7 column with ether. This band was followed on the p_H 7.2 column by a faint band $R = 0.50$ due to a decomposition product. A slower band $R = 0.33$ was found to be DNP-alanine after running on several other columns.

Thus valine, alanine, glycine, threonine, serine, glutamic acid, and aspartic acid were identified as the amino-acids having free α-amino groups.

Standardisation Curves

Standard solutions of the synthetic DNP-acids were prepared by dissolving them in 1% sodium bicarbonate solution (0.5–15 mg DNP-acid/100 ml solution). These were used to calibrate a photoelectric absorptiometer, using Ilford filters No. 601.

Breakdown of DNP-Acids on Hydrolysis

Weighed samples of treated wool (0.5 g) were hydrolysed for 24 hours under reflux in 50 ml 5.7 *N* HCl and the DNP-acids separated and estimated quantitatively. Further samples were hydrolysed for 36 hours and 48 hours, and again the DNP-acids were estimated. By this means, rates of decomposition were determined under the actual conditions of hydrolysis of the wool used for the estimation of the end-groups. However, SANGER[7] has pointed out that the rates of decomposition of DNP-acids in peptide form may be different from those of the free DNP-acids. The amounts of DNP-acids present after 24 hours were taken as standards and the amounts present after a further 12 and 24 hours' hydrolysis were expressed as percentages of the standards. This method was considered to be the most satisfactory, as it was subsequently found that the stability of the DNP-acids was considerably less in the absence of the hydrolysis products of wool*. This was most pronounced for DNP-threonine and DNP-serine, as also found by DESNUELLE[15], and is probably due to catalytic decomposition by traces of heavy metals. As traces of DNP-valylpeptides were still present after 24 hours' hydrolysis, the value for DNP-valine was rejected and redetermined by the following method.

A sample of treated wool was hydrolysed for 24 hours, the DNP-valine being separated and estimated quantitatively. Approximately 1 mg DNP-valine was weighed

* In contrast to lysozyme (A. R. THOMPSON, *Nature*, 168 (1951) 390).

References p. 562.

out, as well as 0.3 g treated wool and hydrolysed under reflux with 5.7 *N* HCl for 24 hours. The DNP-valine was separated chromatographically and estimated quantitatively. From this value was subtracted the amount of DNP-valine due to the end-groups, known to be present in the solution after 24 hours' hydrolysis, as a result of the first estimation. The fact that DNP-valyl peptides are still present after 24 hours' hydrolysis, would not invalidate the method.

As the stability of DNP-glycine was considerably less than that of the other DNP-acids, the following method of estimation was used. A sample of treated wool was hydrolysed under reflux in 5.7 *N* HCl for 48 hours, after which, it was confirmed by chromatography that all the DNP-glycine had decomposed. A weighed sample of DNP-glycine was introduced and the hydrolysis continued for a further 4 hours, after which the DNP-glycine was separated and estimated. This was repeated with another sample after 8 hours' hydrolysis.

TABLE VI

Time of hydrolysis in hours

DNP-acid	4	8	12	24
	% of DNP-acid remaining			
Valine	*90*	*82*	*74*	57
Alanine	*90*	*81*	73	55
Glycine	56	36		
Threonine	*96*	*92*	89	81
Serine	*94*	*89*	84	71
Glutamic acid	*90*	*82*	74	56
Aspartic acid	*91*	*82*	75	58

The italicized values are interpolated.

Quantitative Estimation of End-Groups

Samples of treated wool weighing approximately 1 g were hydrolysed for various lengths of time, under reflux in 5.7 *N* HCl and the DNP-acids separated by the method previously described. The DNP-valine, DNP-alanine and DNP-glycine were then run separately on unbuffered columns with chloroform, DNP-threonine, DNP-serine and DNP-glutamic acid with 1% butanol/chloroform, and DNP-aspartic acid with 3% butanol/chloroform. The DNP-acids were made up to 10 ml with 1% sodium bicarbonate solution and estimated colorimetrically. Peptides of all the DNP-acids, with the exception of DNP-serine, were found to be present after 4 and 8 hours' hydrolysis, as additional quantities of DNP-acids were separated on a further 12 hours' hydrolysis. Although the presence of the peptides robs the values of any real significance, they are given for purposes of comparison with later values, after correcting for decomposition.

As the minimal molecular weight or repeat unit weight was not already known, it was necessary to fix an arbitrary standard. This was obtained from the minimal weight of original wool calculated to contain two equivalents of end-amino serine. It might be thought more appropriate to have taken one equivalent of end-amino aspartic acid, but as DNP-serine is easily hydrolysed to the free DNP-acid from peptides, is easily separated and undergoes little decomposition under these conditions, it was felt to be a more suitable standard than DNP-aspartic acid, which is also found in a resistant peptide form.

References p. 562.

TABLE VII

Repeat Unit Weight

$1.59 \cdot 10^6$ $1.64 \cdot 10^6$

Time of hydrolysis in hours

DNP-acid	4	8
	Moles of DNP-acid/Repeat Unit Weight	
Valine	1.9 +	2.4 +
Alanine	2.2 +	2.1 +
Glycine	6.2 +	8.7
Threonine	7.3 +	7.1 +
Serine	2.0	2.0
Glutamic acid	1.2 +	1.4 +
Aspartic acid	1.0 +	0.9 +

+ denotes the presence of peptides

A further four samples of treated wool were hydrolysed for 24 hours and separated as previously.

TABLE VIII

Repeat Unit Weight

DNP-acid	$1.62 \cdot 10^6$	$1.57 \cdot 10^6$	$1.62 \cdot 10^6$	$1.62 \cdot 10^6$
	Moles of DNP-acid/Repeat Unit Weight			
Valine	4.1 +	3.5 +	4.1 +	3.5 +
Alanine	2.1	1.8	1.9	2.1
Glycine	—	—	—	—
Threonine	7.9	7.7	7.6	7.5
Serine	2.0	2.0	2.0	2.0
Glutamic acid	1.7	1.8	2.2	2.3
Aspartic acid	1.0	1.2	1.0	1.0

+ denotes the presence of peptides

The values for DNP-glycine were rejected as the correction factor for decomposition during hydrolysis, was too great.

A further method was developed to keep the correction factor for decomposition as low as possible. A sample of treated wool was hydrolysed for 4 hours, the free DNP-acids were separated and estimated, and all peptides were returned to the original hydrolysate, which was evaporated to dryness over phosphorus pentoxide in a vacuum desiccator containing a small beaker of sodium hydroxide to absorb hydrochloric acid. The residue was redissolved in 5.7 *N* HCl and rehydrolysed for a further 8 hours. The separation was again repeated and corrections made for breakdown of the DNP-acids.

The DNP-glutamyl and -aspartyl peptides were lost.

It is probable that small amounts of DNP-peptides were present after the second hydrolysis, as these were detected on 12 hours' hydrolysis of a sample of treated wool.

In an early attempt to avoid a large correction factor for the decomposition of DNP-glycine, a sample of treated wool was hydrolysed under reflux in 5.7 *N* HCl for

References p. 562.

TABLE IX

Repeat Unit Weight
$1.56 \cdot 10^6$
Total Time of Hydrolysis in Hours

DNP-acid	4	12	
	Moles of DNP-acid/Repeat Unit Weight		
	From DNP-acids A	From DNP-peptides B	A + B
Valine	1.9 +	0.8 +	2.7 +
Alanine	1.6 +	0.5	2.1
Glycine	6.1 +	1.8	7.9
Threonine	6.2 +	0.5	6.7
Serine	2.0	0.0	2.0
Glutamic acid	1.3 +	—	—
Aspartic acid	0.8 +	—	—

+ denotes the presence of peptides

2 hours. The DNP-glycine was separated chromatographically and estimated colorimetrically. As the DNP-serine was not estimated, a repeat unit weight of $1.60 \cdot 10^6$ was taken as the standard for calculating moles of DNP-glycine. On further hydrolysis of the remaining material, it was evident that considerable amounts of DNP-glycyl peptides were present after 2 hours' hydrolysis.

Moles of DNP-glycine/Repeat Unit Weight = 4.2.

The values obtained from previous experiments for DNP-glycine have been plotted against the time of hydrolysis, to illustrate the rate of liberation of DNP-glycine from the protein and DNP-glycyl peptides (Fig. 1). A similar curve has been drawn for valine (Fig. 2).

As already stated, the values for valine may be a little low, owing to the presence of peptides. PORTER AND SANGER[8] found in globulin that 10% of DNP-valine was still in peptide form after 16 hours' hydrolysis.

The Effect of Urea on the Chain Length of Wool Keratin

Untreated wool was cut into short lengths and immersed for 24 hours at 40° in

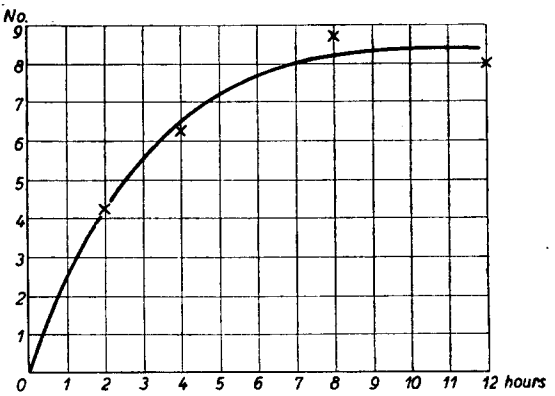

Fig. 1. Rate of liberation of DNP-glycine on hydrolysis of treated wool. Moles of DNP-glycine per repeat unit weight, plotted against time of hydrolysis in boiling 5.7 *N* HCl

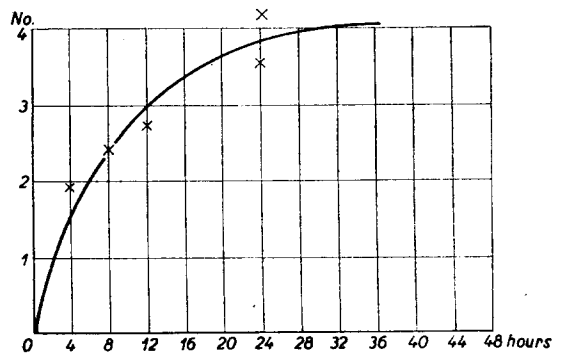

Fig. 2. Rate of liberation of DNP-valine on hydrolysis of treated wool. Moles of DNP-valine per repeat unit weight, plotted against time of hydrolysis in boiling 5.7 *N* HCl

References p. 562.

TABLE X

Acid	No. of free α-Amino Groups per Repeat Unit				Mean*	Nearest Whole No.
Valine	4.1	3.5	4.1	3.5	3.8	4
Alanine	2.1	1.8	1.9	2.1	2.0	2
Glycine	8.7	7.9			8.3	8
Threonine	7.9	7.7	7.6	7.5 (6.7)	7.7 (7.5)	8
Serine	**2.0**	**2.0**	**2.0**	**2.0**	**2.0**	2
Glutamic acid	1.7	1.8	2.2	2.3	2.0	2
Aspartic acid	1.0	1.2	1.0	1.0	1.0	1
				Total =	26.8	27

Repeat Unit Weight
$(1.62\ 1.59\ 1.62\ 1.57)\cdot 10^6 = 1.60\cdot 10^6$
Average Chain Weight = 60,000

* These values were given in the paper delivered to the 1st, International Congress of Biochemistry 1949, but the values in the Abstract are incorrect.

a saturated solution of urea containing 1% (W/V) sodium bisulphite at p_H 5[16]. The solution was filtered from the insoluble residue and precipitated by the addition of acetic acid. This precipitate was washed free from sodium bisulphite and urea, then dried over phosphorus pentoxide. The insoluble residue was again treated with a similar solution, only a trace of soluble material being removed. The residue was immersed in dilute acetic acid, then thoroughly washed with water. Both these products gave negative nitroprusside reactions for thiol groups. Samples were treated by FDNB and the end-groups determined.

All the end-groups which were present in the original wool were found in both fractions in the same ratio, with the exception of threonine, which had halved in value. The repeat unit weights were also approximately half the original value. This could be explained if the treatment with urea and bisulphite had halved the chains in length, with the liberation of identical end-groups, the threonine-ended chains remaining unchanged.

DISCUSSION

If wool keratin substantially consists of a single protein, or stoichiometric association of proteins, and the chains are associated in a regular order, then the most satisfactory arrangement is found to be a hexagonal close-packed system with the chain ending in aspartic acid at the centre (Fig. 3) surrounded by six chains ending in alanine, serine, glutamic acid, alanine, serine, glutamic acid. In turn, this series is surrounded by twelve chains ending in valine, glycine, glycine, repeated four times. For reasons to be mentioned later, the chains ending in threonine have not been included. It is not suggested that this is the only possible arrangement; for instance, the four chains ending in valine could be exchanged for two ending in alanine and two in serine, but at least it has the merit of being highly symmetrical. In Fig. 3, the free-amino ends of the chains are represented as occurring in the same plane, but this is not really necessary, because the chains may be related to each other much as fibres in a yarn, with the difference that the chains are firmly linked to each other by means of cystine disulphide linkages.

References p. 562.

For the sake of simplicity, in this discussion the end-groups are assumed to be in one plane. Again, the chains may be aligned alternately in opposite directions, so that there are approximately an equal number of α-amino and α-carboxyl groups at each end of the repeat unit, represented in the diagram by a positive and negative sign.

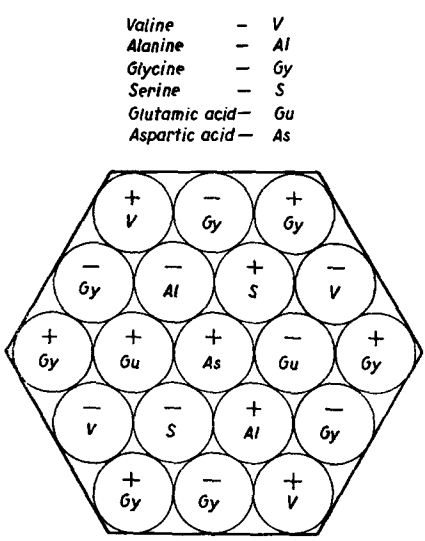

Fig. 3. A cross-section, illustrating the close-packing of the protein chains. The end-groups are shown together with a + and — sign, which signifies that the chains are arranged in opposite directions

The role of the chains ending in threonine was not immediately apparent, but it was felt that it might be somewhat different from the others, in view of the behaviour of the threonine chains in urea/sodium bisulphite solution. If these chains were packed completely round each group of 19 chains as shown in Fig. 4, so as to be shared by neighbouring groups, then the ratio of chains ending in aspartic acid to those ending in threonine would extend from 18:1 with one group of 19 chains, to 8:1 for an infinite number of groups of 19 chains. Rough calculations show that the ratio would be 8.2:1 for the microfibrils of wool, assuming that these are completely covered by a layer of threonine-ended chains. The chains at the corners are shared by three hexagons each, and those along the sides by two hexagons (Fig. 4).

$$6 \times {}^1/_3 + 12 \times {}^1/_2 = 8$$

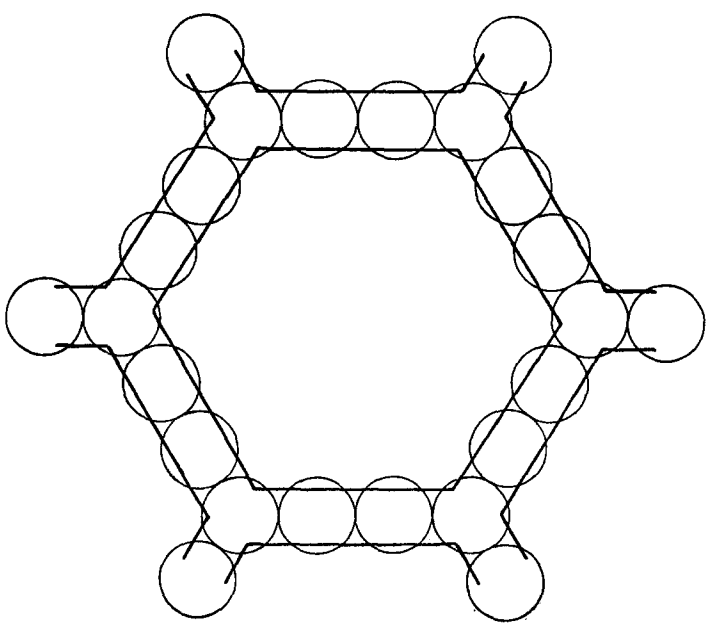

Fig. 4. Illustrating the role of the threonine-ended chains, which act as a cementing material between the groups of chains shown in Fig. 3

References p. 562.

Assuming then, that the repeat unit consists of a close-packed arrangement of twenty-seven protein chains each having a chain weight of approximately 60,000, then from the X-ray work of ASTBURY AND WOODS[17], approximate dimensions of the repeat unit can be calculated. An average residue weight of 107 for keratin enables the number of residues per chain to be calculated:

$$\frac{60{,}000}{107} = 560 \text{ residues.}$$

In α-keratin three residues occupy a length of 5.1 A, so that the total length of a chain would be:

$$560 \times \frac{5.1}{3} = 950 \text{ A.}$$

This would also be the length of the repeat unit.

In the regions of high cystine concentration, however, the length of the disulphide cross-linkage would severely restrict the folding of the chains, so that in these regions the chains might be in the β-form. The proportion of the β-form must be comparatively small, if an unduly strong β-photograph is not to be superimposed on the usual α-photograph obtained from normal wool keratin. However, these photographs contain fairly strong haloes which might obscure a weak β-diagram, but not the strong spots due to the α-fold. A simple calculation shows that the overall length of these chains and repeat unit will be approximately 1200 A, if the proportion of the chain in the β-form is between $^1/_4$ and $^1/_8$ of the total number of residues.

Total number of residues = 560.

$^3/_4$ in α-form = 710 A	$^7/_8$ in α-form = 910 A
$^1/_4$ in β-form = 480 A	$^1/_8$ in β-form = 240 A
Total length = 1190 A	Total length = 1150 A

Repeat unit length = 1200 A (approximately).

ASTBURY[18] has shown that the average distance between the chains of α-keratin in the side-chain direction is 9.8 A; but no value was given for the backbone spacing. If, therefore, the cross-section of a protein chain is assumed to be about 10 A, then it is possible to calculate the approximate diameter of a repeat unit. Since the repeat unit is seven chains in thickness, the overall diameter will be 70 A.

MERCER AND REES[19] treated wool for several days in a solution of trypsin at p_H 8 and 40°. Owing to the slow hydrolysis by the enzyme, the fibres broke down, and the resulting material was examined under the electron microscope. Two main types of material were found (i) amorphous material, consisting of a mass of uniform particles about 100 A in diameter, (ii) fibrous material, consisting of a series of particles of about 100 A in diameter and joined together like a string of beads. Mercer regarded these particles as wool molecules, with a molecular weight of about 300,000, but it should be emphasised that the method of preparation might have caused considerable breakdown of the original molecule, with particles as end-products. Furthermore, the 100 A particle may also be an artifact since the method of preparation might have caused the material to supercontract[20] to 50% of its original length. If this were the case, then the funda-

References p. 562.

mental unit (rather than the molecule) might have an overall length of about 200 A and a diameter of 80 A. This at least would explain the long spacing of 198 A along the fibre axis, found by BEAR[21] and MACARTHUR[22]. However, as will be shown, neither view seriously conflicts with the suggested molecular structure.

BERNAL[23], using X-ray diffraction methods, determined the length of a cystine residue, and PETERS[24] calculated the length of an S–S linkage from carbon atom to carbon atom, obtaining a value of 8.1 A. As this value is considerably less than the side-chain spacing (9.8 A) the chains must necessarily be brought closer to each other to accommodate the large amount of cystine known to be present in wool keratin.

By hydrolysing insulin (a protein with a high cystine content) with chymotrypsin, BUTLER, PHILLIPS AND STEPHEN[25] showed that the cystine was concentrated in a central "core" of about one third the total weight of the molecule. LINDLEY[26] treated wool with cetylsulphonic acid at 65° for several days, obtaining a resistant residue of high cystine and proline content. STEINHARDT AND FUGITT[27] had previously shown that cetylsulphonic acid combined stoichiometrically with the basic side chains of wool keratin, and under suitable conditions catalysed the hydrolysis of amide, and to a lesser extent, peptide linkages. Similarly, SANGER[28] found that $-SO_3H$ groups in protein chains catalysed the hydrolysis of peptide linkages in the same chains. PARTRIDGE AND DAVIS[29] have also suggested that the free carboxyl groups of aspartic and glutamic acids promote the hydrolysis of adjacent peptide bonds by weak acids. As these systems are a source of protons, their resulting high concentration probably causes the hydrolysis of peptide linkages in the immediate neighbourhood. Therefore, under weakly acid conditions, peptide linkages of basic amino-acids are more likely to be hydrolysed than those of the neutral amino-acids in the presence of cetylsulphonic acid. If cystine were uniformly distributed throughout the protein chains, it is difficult to see how treatment of wool with cetylsulphonic acid would result in an insoluble residue of high cystine content. Thus Lindley's treatment probably severed peptide bonds in the regions of high concentration of basic amino-acid residues, leaving insoluble "cores" of high cystine content (for details see Fig. 5, which is discussed later).

Assuming that there are regions of high cystine concentration at frequent intervals along the repeat unit, and that within these regions all the chains are linked by cystine cross-linkages to all neighbouring chains, then since the repeat unit is seven chains in thickness, the cross-section would be approximately $7 \times 8.1 = 57$ A in these regions. With five of these regions, the repeat unit would be divided into six "beads" of greater width, each with an approximate length of $1200:6 = 200$ A or with eleven narrow regions, the length of each "bead" would be $1200:12 = 100$ A. For purposes of illustration, the repeat unit with six "beads" has been chosen. Since the value of 10 A for the side-chain spacing is an average value, it would be somewhat greater than 10 A inside the "beads", making the overall cross-section of the "beads" considerably greater than 70 A as shown in Fig. 5.

COLEMAN AND HOWITT[30] pointed out that the axis of a peptide chain would change direction at a proline residue. As LINDLEY[26] found a large amount of proline associated with cystine in the insoluble residue, and CONSDEN AND GORDON[31] suggested that "proline residues occur near to, if not next to cystine residues", it is possible that they occur in the outer layers of chains at the junction of the wide portions of the repeat unit with the narrow portions, where the change in direction of the outer chains is considerable.

References p. 562.

It will be noticed that the end amino-acids have small side-chains compared with the other residues present in wool keratin, and as a result their cross-sectional area will be smaller than the average. Together with the fact that CONSDEN AND GORDON[31] were only able to identify cystine dipeptides of the neutral and acidic amino-acids, in which the amino-group of the cystine was combined, it is probable that the end-groups are immediately followed by a zone of high cystine content. With the exception of the threonine-ended chains, treatment of wool keratin in urea/sodium bisulphite solution divided the chains into two equal portions with the liberation of similar end-groups. It is suggested therefore, that all the end amino-acid residues are present at the centre of the regions of high cystine content, and that one section of the latter is broken in urea/bisulphite, converting these residues into end-amino groups with the liberation of chains having a molecular weight of 30,000.

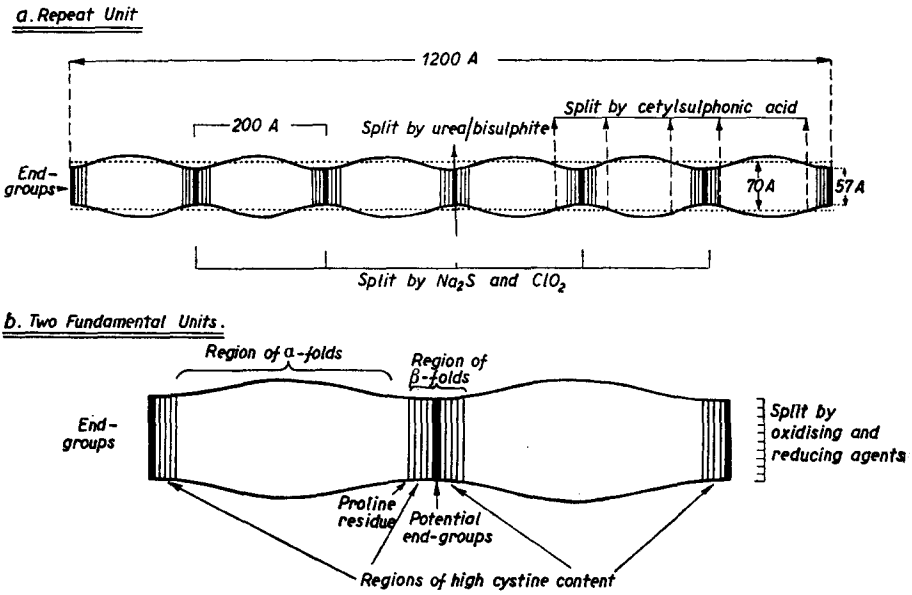

Fig. 5

PARTRIDGE AND DAVIS[29] hydrolysed proteins with dilute acetic or oxalic acid at 100° and showed that polypeptides of high molecular weight together with neutral and acidic amino-acids were obtained, with no trace of basic amino-acids. Conversely, peptide linkages incorporating these acids might be regarded as a source of weakness in a protein chain.

DAS[32] oxidised wool keratin with chlorine dioxide to break the disulphide linkages, and dissolved this material in cupriethylene diamine. He estimated the average molecular weight to be approximately 10,000, and isolated a fraction with a molecular weight of about 30,000. GRALÉN AND OLOFSSON[33] dissolved keratin in alkaline (p_H 11) sodium sulphide and obtained a molecular weight of about 10,000 and an axial ratio of 1:20. This solution was unstable as further degradation took place on standing. It is suggested that these methods of dispersion, which are relatively more severe than urea/bisulphite treatment, have again broken the chains into small sub-units with liberation of similar end-groups, identical with the short chains obtained from a single "bead". Assuming

References p. 562.

the cross-section of a single chain to be 10 A, then these values are in fairly good agreement. Length of chain = 200 A. Chain weight = 60,000 : 6 = 10,000. This is tantamount to assuming that each molecule or repeat unit of wool keratin is made up of several identical "beads" or sub-units, and that the size of the repeat unit is limited by the increased strain placed upon the weakest peptide bond in a given environment. A chain of weight 60,000, when stabilised by association with other chains through hydrogen bonds and disulphide linkages, is strong enough to survive the conditions likely to be encountered while on the sheep, but dissolution in urea/bisulphite solution at p_H 5 and 40° breaks most of the stabilising linkages and the chain is subject to greater flexional strain; consequently, irreversible breakage takes place, resulting in two equal units, each with a chain weight of 30,000. With alkaline sodium sulphide, the chains are again freed from each other, but are also subject to alkaline hydrolysis which would attack the weakest peptide bonds first. Chlorine dioxide would convert the cystine to cysteic acid and the proximity of $-SO_3H$ groups might still further weaken the peptide bonds of neutral and acidic amino-acid residues[28].

It is interesting to compare the hypothetical structure of wool keratin with that of the globular protein edestin. SANGER[34] has shown that the edestin molecule consists of seven chains, one ending in leucine and six ending in glycine. These may be associated in a hexagonal close-packed system with the leucine chain at the centre, surrounded by the six glycine chains. SVEDBERG AND PEDERSEN[35] found the molecular weight to be 309,000 from sedimentation velocity and diffusion measurements, and BURK AND GREENBERG[36] observed an apparent molecular weight of 49,000 in 6.7 *M* urea. The latter value is quite close to the average chain weight, 309,000 : 7 = 44,000, if it is assumed that urea dissociates the molecule into its individual chains. On acidification of edestin, its physical properties change and it is converted into edestan. ADAIR AND ADAIR[37] estimated the molecular weight of this material and obtained a value of 17,000. BAILEY[38] showed that this material was monodisperse and resulted from an irreversible breakdown of the sub-molecule into three equal units. This is very similar to the breakdown believed to take place when wool keratin is treated with sodium sulphide. Since, however, edestin is believed to contain cystine as well as cysteine, it is not clear how urea can dissociate the molecule into individual chains, in the absence of a reducing agent, unless the disulphide linkages are of an intrachain form. It should be recorded that edestin has the characteristic side-chain spacing of 10 A and it was the first globular protein known to be converted into a fibrous form on denaturation.

Lastly, an examination of some other proteins of the KMEF group, suggests that this type of association of end-groups and chains may be characteristic of KMEF protein complexes contained in certain biological materials.

Thanks are due to Prof. J. B. SPEAKMAN for his hospitality and continued interest, and to the International Wool Secretariat for financial assistance.

SUMMARY

1. A method is described for the separation of DNP-amino-acids, using buffered silica gel columns.

2. By treatment with FDNB, and separation of the DNP-amino-acids after hydrolysis, wool is shown to contain seven different amino-acid residues having free α-amino groups *i.e.* valine, alanine, glycine, threonine, serine, glutamic acid and aspartic acid.

3. The end-groups have been quantitatively estimated and the average chain weight determined.

4. A molecular model is proposed which is in accordance with these values.

References p. 562.

RÉSUMÉ

1. L'auteur décrit une méthode pour la séparation des DNP-acides aminés par moyen des colonnes de gel de silice tamponné.

2. Le traitement par le FDNB et la séparation des DNP-acides aminés après hydrolyse, montre que la laine contient sept restes d'acides aminés différents qui possèdent des groupements α-amino libres: valine, alanine, glycine, thréonine, sérine, acide glutamique et acide aspartique.

3. Les groupements terminaux ont été déterminés quantitativement et le poids moyen des chaînes évalué.

4. L'auteur propose un modèle moléculaire qui est en accord avec ces valeurs.

ZUSAMMENFASSUNG

1. Eine Methode zur Trennung von DNP-Aminosäuren mit Hilfe von gepufferten Silicagelsäuren wurde beschrieben.

2. Durch Behandeln mit FDNP und Trennung der gebildeten DNP-Aminosäuren nach Hydrolyse wurde gezeigt, dass die Wolle sieben verschiedene Aminosäure-Reste mit freien α-Aminogruppen enthält, nämlich Valin, Alanin, Glycin, Threonin, Serin, Glutaminsäure und Asparaginsäure.

3. Die Endgruppen wurden quantitativ bestimmt und das durchschnittliche Kettengewicht berechnet.

4. Ein Molekül-Modell wird vorgeschlagen, welches mit diesen Werten in Übereinstimmung ist.

REFERENCES

[1] N. GRALÉN, J. LINDBERG AND B. PHILIP, *Nature*, 162 (1948) 459.
[2] W. T. ASTBURY AND A. STREET, *Phil. Trans. A*, 230 (1931) 75.
[3] J. B. SPEAKMAN, *J. Soc. Dyers Colourists*, Jubilee No., (1934) 34.
[4] R. CONSDEN, A. H. GORDON AND A. J. P. MARTIN, *Biochem. J.*, 40 (1946) 580.
[5] H. B. GOTTLIEB, *J. Am. Chem. Soc.*, 58 (1936) 532.
[6] E. ABDERHALDEN AND P. BLUMBERG, *Z. physiol. Chem.*, 65 (1910) 318.
[7] F. SANGER, *Biochem. J.*, 39 (1945) 507.
[8] R. R. PORTER AND F. SANGER, *Biochem. J.*, 42 (1948) 287.
[9] A. J. P MARTIN AND R. L. M. SYNGE, *Biochem. J.*, 35 (1941) 1358.
[10] G. R. TRISTRAM, *Biochem. J.*, 40 (1946) 721.
[11] A. H. GORDON, A. J. P. MARTIN AND R. L. M. SYNGE, *Biochem. J.*, 37 (1943) 79.
[12] W. R. MIDDLEBROOK, *Nature*, 164 (1949) 501.
[13] A. C. CHIBNALL, M. W. REES AND E. F. WILLIAMS, *Biochem. J.*, 37 (1943) 354.
[14] W. R. MIDDLEBROOK, *Nature*, 164 (1949) 321.
[15] P. DESNUELLE, (1949) Private communication.
[16] C. B. JONES AND D. K. MECHAM, *Arch. Biochem.*, 3 (1943) 193.
[17] W. T. ASTBURY AND H. J. WOODS, *Phil. Trans. A*, 232 (1931) 333.
[18] W. T. ASTBURY, *J. Chem. Soc.*, (1942) 337.
[19] E. H. MERCER AND A. L. G. REES, *Australian J. Exptl Biol. Med. Sci.*, 24 (1946) 175.
[20] J. B. SPEAKMAN, *J. Soc. Dyers Colourists*, 52 (1936) 335.
[21] R. S. BEAR, *J. Am. Chem. Soc.*, 65 (1943) 1784.
[22] I. MACARTHUR, *Nature*, 152 (1943) 38.
[23] J. D. BERNAL, *Z. Krist.*, 78 (1931) 363.
[24] L. PETERS, Ph. D. Thesis, (1950) University of Leeds.
[25] J. A. V. BUTLER, D. M. P. PHILLIPS AND J. M. L. STEPHEN, *Nature*, 162 (1948) 418.
[26] H. LINDLEY, *Nature*, 160 (1947) 190.
[27] J. STEINHARDT AND C. H. FUGITT, *J. Research Nat. Bur. Standards*, 29 (1942) 315.
[28] F. SANGER, *Biochem. J.*, 45 (1949) 563.
[29] S. M. PARTRIDGE AND H. F. DAVIS, *Nature*, 165 (1950) 62.
[30] D. COLEMAN AND F. O. HOWITT, *Fibrous Proteins*, (1946) Soc. Dyers and Col., Bradford, 144.
[31] R. CONSDEN AND A. H. GORDON, *Biochem. J.*, 46 (1950) 8.
[32] D. B. DAS, Ph. D. Thesis, (1948) University of Leeds.
[33] N. GRALÉN AND B. OLOFSSON, (1947) 11th. International Congress of Pure and Applied Chemistry.
[34] F. SANGER, (1948) *Biochem. Soc. Symposium* No. 3, 29.
[35] T. SVEDBERG AND K. O. PEDERSEN, *The Ultracentrifuge*, (1940) Clarendon Press, Oxford.
[36] N. F. BURK AND D. M. GREENBERG, *J. Biol. Chem.*, 87 (1930) 197.
[37] G. S. ADAIR AND M. E. ADAIR, *Biochem. J.*, 28 (1934) 199.
[38] K. BAILEY, *Biochem. J.*, 36 (1942) 140.

Received April 11th, 1951