Journal of Chromatography, 324 (1985) 65–73
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROM. 17 533

REVERSED-PHASE HIGH-PERFORMANCE LIQUID CHROMATOGRAPHY OF SOME ULTRA-HETEROGENEOUS AND COVALENTLY-MODIFIED PROTEINS FROM HUMAN HAIR

HAYEL M. SAID*, A. E. NEWSOM, BARBARA L. TIPPINS and ROGER A. MATHEWS
Redken Laboratories, Inc., Biochemistry Department, 6625 Variel Avenue, Canoga Park, CA 91303 (U.S.A.)
(Received January 3rd, 1985)

SUMMARY

A new procedure for the fractionation of the heterogeneous cystine-rich proteins from human hair, utilizing reversed-phase high-performance liquid chromatography, is described. Of these proteins 27 fractions have been collected and analyzed for amino acid composition. There seems to be little correlation between the elution order and the hydrophobicity of the fraction constituents except for the late-eluting fractions. Based on the elution profiles and amino acid contents, these fractions appear to fall into four families. The effects of alkyl chain length, flow-rate and gradient slope, as well as various additives to the organic modifier on the separation have also been investigated. A low flow-rate (0.4 ml/min) and a shallow gradient were essential for the separation of these proteins as was the use of short alkyl chain (C_4) or medium alkyl chain (C_8) columns. However, with the C_4 column reproducibility and recovery were excellent.

INTRODUCTION

Hard keratins (hair, wool, nail, claw, hoof, etc.) are proteins which are extremely insoluble due to their highly cross-linked structure[1,2]. Breakage of the disulfide links is a necessary step for their partial solubilization. The most widely studied member of these keratins is wool in which three groups of proteins have been identified[3–6]: (i) the cystine-poor or low-sulfur group which comprises the organized fibrillar components, (ii) the cystine-rich or high-sulfur group which is thought to make up the matrix that cements the fibrils together, (iii) the high tyrosine-glycine proteins whose function is not established yet and are absent from human hair.

The limited solubility of the keratin proteins, particularly the cystine-rich group, has hampered satisfactory separation and characterization of these species by conventional means. Indeed, separations based upon DEAE-cellulose chromatography, gel electrophoresis, ultracentrifugation or differential solubility fractionation have given poor results to date[7]. Therefore, the cystine-rich proteins from human hair which are even harder to solubilize have not been successfully separated or

0021-9673/85/$03.30

characterized. A few recent studies[8,9] have utilized two-dimensional polyacrylamide gel electrophoresis and obtained some improved resolution. These procedures however, were not adequate for quantitation due to the smearing effects encountered, and not suitable for analytical studies due to the recovery limitations intrinsic to the procedure.

High-performance liquid chromatography (HPLC) has been successfully employed in the fractionation of proteins from various sources[10–16]. We report here the successful hydrophobic separation of the cystine-rich proteins of human hair into 27 components, each of which was collected and analyzed for amino acid composition, and the effects of various hydrophobic supports and modifiers that influenced this separation.

MATERIALS AND METHODS

Keratin protein extraction

Unless otherwise indicated, normal hair from a six year old female child was used in these studies. The hair was thoroughly washed with light petroleum (b.p. 35–80°C), ethanol and water, and dried overnight at room temperature. Soluble keratin proteins were extracted by a procedure modified from Baden *et al.*[17] and Harrap and Gillespie[18]. Coarsely-cut hair was pulverized in a steel mortar and pestle precooled in liquid nitrogen, weighed and placed in a sealable flask containing nitrogen-flushed reducing buffer (buffer–hair, 100:1). The buffer was made up of 0.2 *M* Tris (pH 9.5), 8 *M* urea and 0.03 *M* dithiothreitol. After incubation at 50°C with intermittant agitation for 1 h the suspension was homogenized for 3 min at full speed in ice in a polytron homogenizer (Brinkmann, Westbury, NY, U.S.A.), reincubated at 50°C for 1 h and left overnight at room temperature under nitrogen blanket. Precipitation of the non-solubilized material was achieved by centrifugation for 30 min at 28,000 *g* and 4°C in a Beckman J-21B centrifuge (Beckman, Palo Alto, CA, U.S.A.). The pellet was washed three times, lyophylized, and stored at −20°C. The supernatant, containing the solubilized keratin proteins was treated with iodoacetic acid (2.5 g/g hair dissolved in 0.5 *M* sodium hydroxide) to alkylate the free sulfhydryl groups. The iodoacetic acid was added slowly in the dark while the pH was maintained constant at 8.0 by using a Metrohm Herisau automatic titrator (Brinkmann Instruments). The resultant S-carboxymethylated proteins were dialyzed against running tap water overnight followed by deionized-distilled water for 24 h at 4°C. The cystine-rich proteins were isolated by the standard zinc ion precipitation method[19]. The dialysis bags were placed in 0.03 *M* zinc acetate solutions for 72 h at 4°C with one medium change and centrifuged as mentioned before. The supernatant, now containing the cystine-rich proteins, was dialyzed for 24 h with two medium changes, against 0.02 *M* sodium citrate to remove the residual zinc, then against deionized-distilled water, lyophylized, and stored at −20°C. The pellet, comprising the cystone-poor proteins, was suspended in water and treated similarly. The yield of the cystine-rich proteins was 16% and that of the cystine-poor was 25% of the original hair.

Reversed-phase high-performance liquid chromatography (RP-HPLC)

Fractionation of the cystine-rich proteins was performed on a Waters Assoc.

(Milford, MA, U.S.A.) HPLC system consisting of two M6000A pumps, M720 system controller solvent programmer, WISP (Model 710B) autoinjector, 441 UV detector and 740 data module. An LKB 7000 UltroRac fraction collector (LKB, Bromma, Sweden) was controlled by the system controller. Reversed-phase columns of varying alkyl chain lengths included a radial compression module RCM-100 equipped with a C_{18} Radial-Pak cartridge (Waters Assoc.); a Protesil 300 Octyl (C_8 column, Whatman, Clifton, NJ, U.S.A.); a Hi-Pore RP-304 (C_4, BioRad, Richmond, CA, U.S.A.); and a C_1 methyl column (IBM Instruments, Danbury CT, U.S.A.). Acetonitrile, 1-propanol and methanol, all HPLC grade, were purchased from Burdick & Jackson (Muskegon, MI, U.S.A.), Trifluoroacetic acid (TFA) (sequanol grade) was from Pierce (Rockford, IL, U.S.A.). Ultrapure water was obtained by passing deionized-distilled water through a Barnstead (Boston, MA, U.S.A.) 4-module Nanopure II purification system. Gradients and conditions are as described in the legends.

The cystine-rich proteins were dissolved in the primary solvent prior to injection. Fraction collection was achieved by a timed-events table programmed into the system controller. Fractions were pooled, lyophylized and stored at −20°C. Protein recovery was determined by amino acid analysis of the injected sample collected prior to and after passage through the column. Periodic cleaning of the columns was achieved by running repeated cycles overnight of a shallow gradient (4 h) oscillating between 40% acetonitrile and 80% methanol at a low flow-rate. It was critical in all of these studies to keep the pH of the various solvents away from the range 3–4 to avoid precipitation of these proteins.

Amino acid analysis

Fractions as well as whole cystine-rich proteins, and whole hair were hydrolyzed in vacuo in sealed ampules containing 2 ml 6 *M* hydrochloric acid at 110°C for 24 h (48 h for whole hair). Hydrochloric acid was removed by a SpeedVac concentrator (Savant Instruments, Hicksville, NY, U.S.A.). Hydrolysates were redissolved in 1.0 ml of an amino acid dilution buffer (Beckman) and analyzed for amino acid composition using either a Beckman 121MB automated amino acid analyzer or a Beckman 6300 high-performance analyzer.

RESULTS

Hydrophobic separation of the cystine-rich proteins

Effect of alkyl chain length. Long chain (C_{18}) and ultra-short chain length (C_1) columns yielded very poor separations of the cystine-rich proteins from human hair. Retention of these proteins by the column support was very high and recovery very low on the C_{18} column while the opposite was observed on the C_1 column (data not shown). The best separation was achieved on a short chain (C_4) column where at least 27 components of these proteins were resolved and collected (Fig. 1). The profile obtained in Fig. 1 was reproduced over the two hundred cycles used to collect the different components, and was used to identify numerically the peaks. Retention times and areas of these peaks are presented in Table I.

The medium chain length (C_8) column which was originally employed in most of the primary studies yielded acceptable resolution. Retention was expectedly

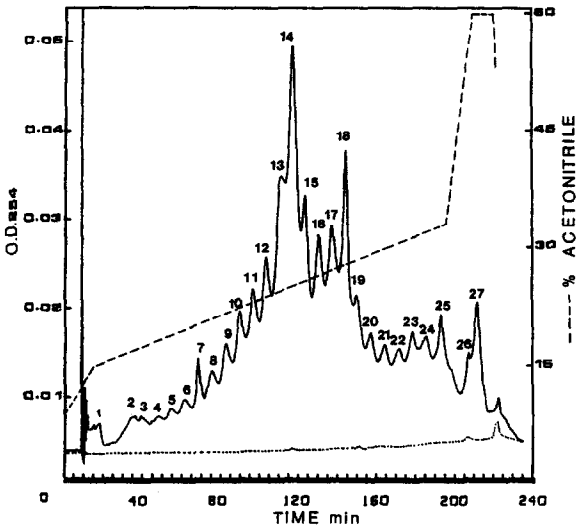


Fig. 1. RP-HPLC profile of the cystine-rich proteins of human hair. Samples (0.5–1.0 mg) were dissolved in solvent A (0.02% TFA) and injected onto a Bio-Rad Hi-Pore RP-304 column (C_4). Initial conditions: 85% A, 15% B (solvent B = 60% acetonitrile, 0.02% TFA). Flow-rate: 0.4 ml/min. The gradient (dashed line) was linearly increased from 15% B to 25% B in 15 min, to 55% B in 180 min and to 100% B in 20 min. Detection was at 254 nm. Dotted line is a tracing from a blank run. Numericals refer to the fractions in their order of elution.

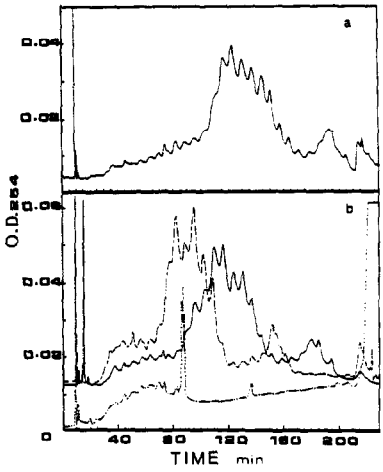
stronger and a steeper gradient was required to obtain a profile similar to that of the C_4 column (Fig. 2a). Recovery, however, was slightly lower than that of the C_4 column (90% *vs.* 95%).

Effect of salt, sodium dodecyl sulfate (SDS) and urea. In order to obtain more insight into the structure of the cystine-rich proteins, several additions to the organic

TABLE I

RETENTION TIMES AND PERCENT AREAS OF THE 27 PEAKS RESULTING FROM THE RP-HPLC OF HUMAN HAIR CYSTINE-RICH PROTEINS ON A C_4 COLUMN

Peak No.	*Retention time*	*Area (%)*	*Peak No.*	*Retention time*	*Area (%)*
1	16.4	1.2	15	122.8	6.8
2	35.8	0.1	16	129.9	6.8
3	38.9	0.1	17	136.7	7.3
4	47.6	0.2	18	143.6	8.3
5	54.88	0.5	19	149.6	3.6
6	61.6	1.0	20	156.6	2.7
7	68.3	1.4	21	163.9	2.3
8	75.4	1.9	22	171.2	2.0
9	82.4	3.0	23	177.9	2.6
10	89.4	3.1	24	185.0	2.5
11	96.4	4.9	25	192.8	3.5
12	103.0	5.4	26	207.0	1.0
13	110.3	9.5	27	211.2	3.3
14	116.4	15.0			



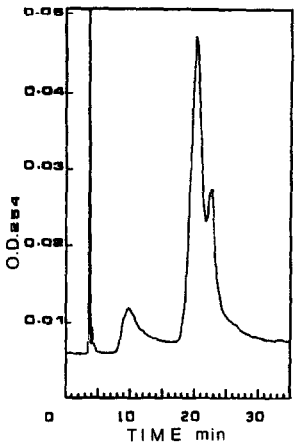


Fig. 2. RP-HPLC of the cystine-rich proteins on a Whatman Protesil 300 C_8 column. (a) Same conditions as in Fig. 1 except that the gradient was increased from 15% B to 29% B in 15 min, to 60% B in 180 min and to 100% B in 15 min. (b) Same as in (a) except that solvent B contained either 0.2 *M* sodium chloride (solid line), or 4 *M* urea (dashed line), or 0.4% SDS (dotted line).

Fig. 3. Fast separation of the cystine-rich proteins on a C_8 column. An amount of 0.2 mg sample was injected. Initial conditions: 80% A and 20% B (both A and B as in Fig. 1). Flow-rate: 1 ml/min. The gradient was linearly increased to 40% B in 1 min where it was held for 10 min, and increased to 90% B in 10 min and held for an additional 5 min.

modifier were made without changing its overall concentration. Addition of 0.2 *M* sodium chloride did not result in major changes in the elution profile of these proteins from the C_8 column (Fig. 2b), indicating that hydrophobic interactions are the major factors involved in the separation process. There was however a change in the relative amount of the late-eluting peaks which may not rule out some adsorption effects. No such changes were observed with the less hydrophobic short-chain C_4 column (data not shown). When 4 *M* urea was introduced into the organic modifier, a shift in the

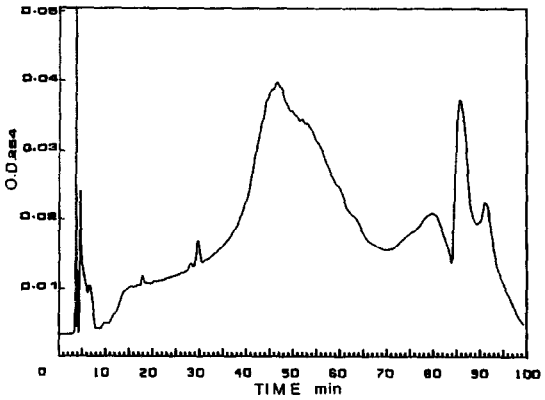


Fig. 4. Effect of gradient steepness on the separation of the human hair cystine-rich proteins on a C_8 column. Conditions were as in Fig. 2a except that the flow-rate was 1.0 ml/min and the time of stepping up the gradient was reduced by a factor of 2.5 to maintain a comparable gradient slope.

relative areas of the intermediate peaks was noticed (Fig. 2b) without changes in the retention times of these peaks. The ionic detergent SDS caused a shift in the opposite direction (Fig. 2b), resulting in increased retention and loss of resolution. Substitution of methanol or 1-propanol for acetonitrile as the organic modifier yielded unsatisfactory separations (data not shown).

Effect of flow-rate and gradient configuration. A low flow rate (0.4 ml/min) and a shallow gradient were essential for the separation of the cystine-rich proteins. At a flow-rate of 1 ml/min with a steep gradient, only three major groups of these proteins were detected (Fig. 3). When the flow-rate was kept at 1 ml/min and the gradient steepness (% change of solvent B/ml) was reduced (Fig. 4) to match that of Fig. 2a resolution was considerably poorer. This observation emphasizes the importance of an optimal interaction time between the protein molecules and the hydrophobic support, and stresses the need for selecting the proper flow-rate for different protein separations.

Amino acid analysis

Peaks shown in Fig. 1 were collected in 27 fractions and their amino acid content was determined and is graphically presented in Fig. 5. It should be noted in this regard that the fractions are not homogeneous as each was resolved into several

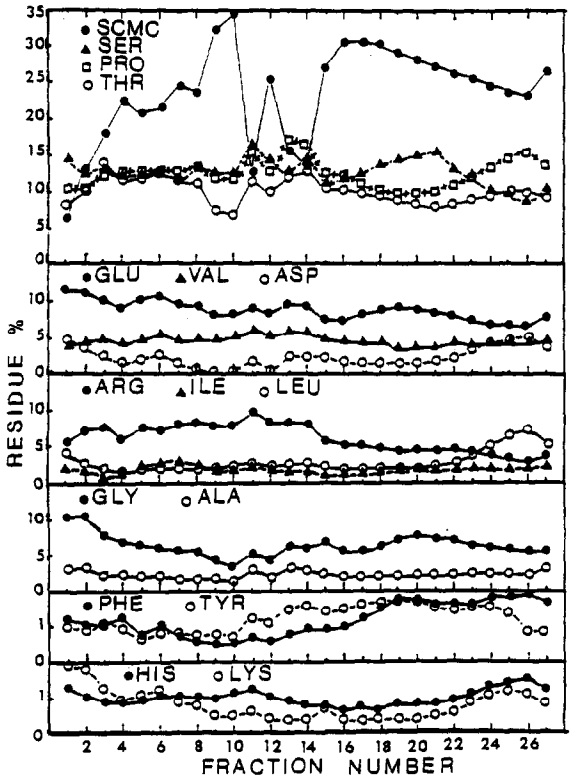


Fig. 5. Graphic representation of the amino acid composition of the 27 fractions from human hair cystine-rich proteins.

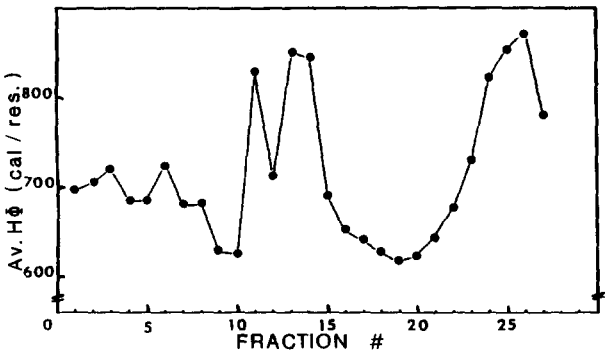


Fig. 6. Average hydrophobicity (expressed as cal/residue) of the cystine-rich protein fractions. Hydrophobicity values were calculated from estimates of the free energy change for transfer of amino acids from ethanol to water as suggested by Tanford[22] and Bigelow[21].

subcomponents by size-exclusion HPLC in 6 *M* urea[20]. Comparison of amino acid content, therefore, is not very helpful except in pointing out some general trends. First there is a large fluctuation in the cystine content of these fractions [represented by S-carboxymethyl cysteine (SCMC)] which ranged between a low-content of 10–15% and a high content of 30–35%. Second, four amino acids namely proline, serine, threonine and cystine–cysteine constitute on the average more than 55% of the whole amino acid population of these fractions. Third, fractions 15–27 may comprise two families of oligomeric proteins as evidenced by the inflection point at fraction 20. The amino acids cystine, serine, glutamic acid, arginine and glycine which vary inversely with the rest of the amino acids possibly make up the bulk of the amino acid residues of one of the families. Fractions 11–14 which are low in cystine content may fall into a third family and fractions 1–10 probably form yet a fourth whose members fall into the low-molecular-weight range (below 3000 daltons[20]). Fourth, there seems to be a lack of correlation between the elution order of the fractions and their hydrophobicities (Fig. 6) calculated according to the procedure of Bigelow[21] based on that of Tanford[22]. This however, is not surprising in view of the heterogeneity of these proteins and the fact that major structural changes have been introduced into these molecules by disrupting the disulfide linkages and therefore destabilizing hydrophobic bonding[21].

DISCUSSION

RP-HPLC has resulted in the succesful fractionation of human hair cystine-rich proteins into twenty-seven collectable fractions. This fractionation profile illustrated in Fig. 1 represents a marked improvement in the methodology of isolating these ultra-heterogeneous proteins which traditionally require a combination of chromatographic and electrophoretic procedures to obtain even partial resolution[23]. The advantages of this procedure stems from the powerful resolving capacity of HPLC supports. Many studies on hair and wool[24–27] have indicated that the electrophoretic separations of these proteins were modest. This weakness in the otherwise potent technique is due to charge load which is tagged onto these molecules during solubilization and subsequent covalent modification. Even in studies where two-dimen-

sional gel electrophoresis was employed at low pH, the components which are very high in cystine (S-carboxymethyl derivative) were not resolved[8,9].

The fractions obtained in our study are not homogenous. The broad peaks observed allude to this fact. Size-exclusion HPLC studies of these fractions in 6 *M* urea have resulted in further fractionation (manuscript in preparation) with a wide range of molecular size distribution. The data presented here suggest that at least three families (Fig. 3) or possibly four (Figs. 2, 4 and 5) may make up the cystine-rich protein population of human hair. In comparison, five such families have been recognized in wool[28] and three in mouse hair[29] cystine-rich proteins.

The factors that influence the interaction of these proteins with each other selectively affect their interaction with the hydrophobic support. Urea for example, resulted in a change in the relative amounts of some peaks (Fig. 2b) without changing retention time, whereas sodium dodecyl sulfate increased retention of most of the fractions. Whether this effect is caused by an increase in the polarity of the mobile phase or a decrease in that of the fractions is not yet determined. In this regard, sodium chloride did not have a pronounced effect.

The lack of correlation throughout the separation of the elution order with that of hydrophobicity is not unique to these proteins. Barford *et al.*[30] encountered a similar situation with standard proteins, and other investigators[31] found that the correlation only holds for small peptides (15 residues or less). For the first third of the fractions which are mostly low-molecular-weight fragments, retention may correlate with the increasing arginyl content (Fig. 5) which at low pH would interact more strongly with the stationary phase[11,32]. The last third of the fractions seems to correlate well with hydrophobicity measurements (Fig. 6), and an interesting trait here is the increasing content of the hydrophobic amino acid leucine and decreasing values for SCMC.

O'Hare *et al.*[33] experienced lower protein recoveries with decreasing flow-rates and low rates of gradient change. We have not encountered such problems with the cystine-rich proteins since low flow-rate and the shallow gradient employed in the present studies enhanced resolution and did not affect recovery. Allowing time for the different distribution equilibria of the various protein molecules to be reached between the hydrophobic support and the mobile phase, is a critical factor in any separation.

ACKNOWLEDGEMENTS

We wish to extend our appreciation to Paula and John Meehan for their continued support of research on keratinized tissues.

REFERENCES

1 R. D. B. Fraser, T. P. Macrae and G. F. Roger, in I. N. Kugelmass (Editor), *Keratins: Their Composition, Structure and Biosynthesis*, C. C. Thomas, Springfield, IL, 1972, p. 30–55.
2 J. M. Gillespie and R. C. Marshall, *Aust. J. Biol. Sci.*, 30 (1977) 401.
3 W. G. Crewther, *Proc. 5th Int. Wool Text. Res. Conf., Aachen 1975, I*, Deutsches Wollforschungsinstitut, 1976, pp. 1–101.
4 L. S. Swart, D. Parris and F. J. Joubert, *J. Chromatogr.*, 78 (1973) 363.
5 I. J. O'Donnell and E. O. P. Thompson, *Aust. J. Biol. Sci.*, 17 (1964) 973.

6 E. G. Bendit, in D. A. D. Parry and L. K. Creamer (Editors), *Fibrous Proteins: Scientific, Industrial, and Medical Aspects*, Vol. 2, Academic Press, NY, 1980, pp. 185–194.
7 J. M. Gillespie and J. M. Frenkel, *Comp. Biochem. Physiol.*, 47B (1974) 339.
8 R. C. Marshall, *J. Invest. Dermatol.*, 80 (1983) 519.
9 J. M. Gillespie and R. C. Marshall, *J. Invest. Dermatol.*, 80 (1983) 195.
10 E. C. Nice, M. W. Capp, N. Cooke and M. J. O'Hare, *J. Chromatogr.*, 218 (1981) 569.
11 W. S. Hancock and J. T. Sparrow, *J. Chromatogr.*, 205 (1981) 71.
12 M. T. Hearn, *Adv. Chromatogr.*, 20 (1982) 1.
13 F. E. Regnier, *Methods Enzymol.*, 91 (1983) 137–190.
14 L. R. Gurley, D. A. Prentice, J. G. Valdex and W. D. Spall, *J. Chromatogr.*, 266 (1983) 609.
15 A. R. Kerlavage, C. J. Weitzman, T. Hasan and B. S. Cooperman, *J. Chromatogr.*, 266 (1983) 225.
16 S. D. Power, M. A. Lochrie and R. O. Poyton, *J. Chromatogr.*, 266 (1983) 585.
17 H. P. Baden, L. D. Lee and J. Kubilus, *Amer. J. Hum. Genet.*, 27 (1975) 472.
18 B. S. Harrap and J. M. Gillespie, *Aust. J. Biol. Sci.*, 16 (1965) 542.
19 J. M. Gillespie, *Aust. J. Biol. Sci.*, 17 (1964) 282.
20 H. M. Said, A. E. Newsom, B. L. Tippins and R. A. Mathews, unpublished results.
21 C. Bigelow, *J. Theoret. Biol.*, 16 (1967) 187.
22 C. Tanford, *J. Amer. Chem. Soc.*, 84 (1962) 4240.
23 F. J. Joubert, P. J. de Jager and L. S. Swart, in W. G. Crewther (Editor),*Symposium on Fibrous Proteins, Australia, 1967*, Plenum Press, New York, 1968, pp. 343.
24 R. J. Pollett and P. D. Stonier, *Biochem. J.*, 122 (1971) 433.
25 L. S. Swart, F. J. Joubert, T. Haylett and P. J. de Jager, *Proc. 3rd Int. Wool Text. Res. Conf., Paris, I*, Deutsches Wollforschungsinstitut, 1965, p. 493.
26 H. Lindly, J. M. Gillespie and T. Haylett, in W. G. Crewther (Editor), *Symposium on Fibrous Proteins, Australia, 1967*, Plenum Press, New York, 1968, p. 353.
27 R. L. Darskus, *J. Chromatogr.*, 69 (1972) 341.
28 R. L. Darskus, J. M. Gillespie and H. Lindly, *Aust. J. Biol. Sci.*, 22 (1969) 1197.
29 R. C. Marshall and J. M. Gillespie, *Aust. J. Biol. Sci.*, 29 (1976) 11.
30 R. A. Barford, B. J. Sliwinski, A. C. Breyer and H. L. Rothbart, *J. Chromatogr.*, 235 (1982) 281.
31 M. J. O'Hare and E. C. Nice, *J. Chromatogr.*, 171 (1979) 209.
32 B. Grego, F. Lambrou and M. T. W. Hearn, *J. Chromatogr.*, 266 (1983) 89.
33 J. J. O'Hare, M. W. Capp, E. C. Nice, N. H. C. Cooke and B. G. Archer, *Anal. Biochem.*, 126 (1982) 215.