The Preparation and Properties of a Group of Proteins from the High-Sulphur Fraction of Wool

By H. LINDLEY and T. C. ELLEMAN

*Division of Protein Chemistry, Commonwealth Scientific and Industrial Research Organization,
Parkville (Melbourne), Vic. 3052, Australia*

(Received 17 January 1972)

A method is reported for the preparation of a group of three proteins from the *S*-carboxymethylated high-sulphur fraction of wool. These proteins have been partially characterized by their tryptic peptides. All have similar structural features and show an interesting homology within the group and some similarities of sequence with a different group of wool high-sulphur proteins. The evidence for the sequence of some of the peptides is given in a supplementary paper that has been deposited as Supplementary Publication 50008 at the National Lending Library for Science and Technology, Boston Spa, Yorks. LS23 7BQ, U.K., from whom copies can be obtained on the terms indicated in *Biochem. J.* (1972) 126, 5.

The high-sulphur proteins of wool, hair and other α -keratins are a group of cystine-rich proteins believed to occur in the non-filamentous matrix of the fibre. A major feature of these proteins is their high degree of heterogeneity, exceeded only by that found in the immunoglobulins. From this high-sulphur fraction of wool Gillespie (1963) isolated the *S*-carboxymethyl derivative of what he believed to be a homogeneous protein which he designated SCMCK-B2 (SCMK indicates *S*-carboxymethylkerateine). The criteria of homogeneity used were absence of lysine, histidine and methionine from the protein, a single peak in moving-boundary electrophoresis at several pH values and the identity of molecular weight values derived from ultracentrifuge and light-scattering methods. Subsequent papers from this laboratory have described work on the SCMCK-B2 fraction, including some investigation of the amino acid sequence which showed that some peptides could be isolated in near stoichiometric yields (Lindley *et al.*, 1968; Haylett & Lindley, 1968). However, as described by Lindley *et al.* (1971) compelling evidence was finally obtained that fraction SCMCK-B2, as isolated by Gillespie (1963), is a mixture of closely related proteins, and accordingly we have sought fresh techniques for the isolation of more homogeneous preparations.

Darskus *et al.* (1969) chromatographed the high-sulphur protein fraction of wool on DEAE-cellulose (pH 4.6, by using an ionic-strength gradient) and it was shown by tryptic peptide 'maps' that as much as 40% of the total high-sulphur protein of wool contains some peptides characteristic of SCMCK-B2 fraction (Gillespie *et al.*, 1968). Although little resolution of the high-sulphur protein fraction into individual components could be seen under these con-

ditions, the chromatographic behaviour of this 'SCMK-B2 like' fraction from the high-sulphur protein indicated a complex mixture of proteins. Real progress in the fractionation of this 'SCMK-B2 like' material only became possible with the finding that much better resolution was achieved at low pH values in 5.0M-urea on SE-Sephadex columns.

In the present paper we describe a technique which permits the isolation of three components of the SCMCK-B2 family from the high-sulphur protein fraction of wool. We designate them components SCMCK-B2A, SCMCK-B2B and SCMCK-B2C, and in this report refer to them simply as components A, B and C.

Materials and Methods

Enzymes

The trypsin (diphenylcarbamoyl chloride-treated, B-grade, lot no. 801619) and thermolysin were from Calbiochem, Los Angeles, Calif., U.S.A. Chymotrypsin and pepsin were obtained from Worthington Biochemical Corp., Freehold, N.J., U.S.A. Carboxypeptidases A and B (di-isopropylphosphorofluoridate-treated) were obtained from Sigma Chemical Co., St. Louis, Mo., U.S.A.

Preparation of proteins

S-Carboxymethyl kerateines were prepared from Lincoln wool (150g) by the method of Gillespie (1962). The high-sulphur protein fraction from wool was extracted at 0°C for 16h with 0.8M-potassium thioglycollate at pH 10.2. The extracted protein was precipitated by the addition of acetic acid to pH 5.0,

redissolved in 0.1M-potassium thioglycollate, and alkylated with iodoacetic acid. The pH was maintained throughout the alkylation between 8 and 9.0.

Chromatography of reduced and carboxymethylated high-sulphur protein (approx. 20g), was performed on a column (100cm×2.5cm) of SE-Sephadex C-25 by using a linear gradient from 2% (w/v) citric acid (2 litres) to 4% (w/v) tricine [*N*-tris(hydroxymethyl)methylglycine, 2 litres] both in 5.0M-urea previously purified by passage through a mixed-bed ion-exchange resin. Three peaks were obtained (Fig. 1) from each of which the major component was purified on a column (100cm×1.5cm) of DEAE-cellulose by using a linear gradient from 2.0M-urea (800ml) to 2.0M-urea–0.5M-formic acid (800ml). Column effluents were monitored at 280nm and also by discontinuous polyacrylamide-gel electrophoresis of selected fractions (Davis, 1964).

Molecular weights were determined in the ultracentrifuge by the high-speed equilibrium method of Yphantis (1964). Measurements were made at 44000rev./min across the whole cell in 0.1M-NaCl–0.05M-Na₂B₄O₇ buffer, pH9.0.

Enzymic digestion

Enzymes were used at an enzyme/substrate ratio of 1:100 (w/w). Tryptic and chymotryptic digestions were generally done for 4h at 37° in a 1% (w/v)

solution of NH₄HCO₃ at pH8.0. Thermolytic and carboxypeptidase digestions were done for 8h under the same conditions. Pepsin digestions were done in aq. 5% (v/v) formic acid for 8h.

Fractionation of peptides

Peptides from tryptic digestion of each protein fraction (approx. 200mg) were initially separated on a column (100cm×1.5cm) of DEAE-cellulose by using a linear gradient from 0.15% pyridine (1.6 litres) to 3% (v/v) pyridine–30% (v/v) acetic acid (1.6 litres), the peaks being detected by the Technicon AutoAnalyzer system of Catravas (1964). The *E*₂₈₀ indicated the position of tryptophan-containing peptides.

Further purification of the peptides was achieved by high-voltage paper electrophoresis as described by Ambler (1963). Peptides were detected on paper by staining with 0.25% ninhydrin–2% (w/v) collidine in ethanol, and further over staining with the chlorination method of Reindel & Hoppe (1954). Peptides were eluted from paper strips with aq. 1% (w/v) NH₃.

Amino acid analysis and sequence determination

Proteins and peptides were hydrolysed with 5.7M-HCl (0.3ml+1μl of mercaptoethanol) in sealed evacuated tubes at 105°C for 24h, or longer if necessary. Amino acid compositions were determined with a Beckman model 120C amino acid analyser.

The main method of amino acid sequence determination was the dansyl-Edman technique (Gray, 1967). Large peptides (>12 residues) were subjected to further enzymic digestion and the products, after purification by high-voltage paper electrophoresis, were subjected to the dansyl-Edman procedure. C-Terminal residues were released by either carboxypeptidase A or B and were identified by high-voltage paper electrophoresis at pH1.9. The presence of amide groups in peptides was deduced from the electrophoretic mobility of the peptide at pH6.5 (Offord, 1966) relative to aspartic acid (*m* = –1.0) or alternatively of the Dns-peptide relative to Dns-sulphonic acid (*m* = –1.0).

Identification of the N-terminal residue by mass spectrometry

The ninhydrin-negative amino-terminal chymotryptic heptapeptide was permethylated both by the procedure of Thomas (1969) and also by the procedure of Roepstorff *et al.* (1970) which has been designed to avoid problems due to the formation of sulphonium ions. The usual preliminary acetylation step of these procedures was omitted in this case. Mass spectra were obtained on an AEI model MS9 mass spectrometer.

Fig. 1. Elution profile of the high-sulphur protein fraction of wool from SE-Sephadex

A 100cm×2.5cm column was used at a flow rate of 2.5ml/min, with a linear gradient from 2% (w/v) citric acid–5.0M-urea (2 litres) to 4% (w/v) tricine–5.0M-urea (2 litres), and 20 ml fractions were collected. Components A, B and C are indicated.

Results

Components B and C are isolated from the final chromatography on DEAE-cellulose as single peaks (Fig. 2), and also give single bands on polyacrylamide-gel electrophoresis; component A however, yields a flat-topped elution profile, and shows the presence of two bands on polyacrylamide-gel electrophoresis. Each component appears homogeneous in the ultracentrifuge, the molecular weights of components B and C being estimated at approx. 18000 and that of component A at approx. 20000. Amino acid analyses based on these values are shown in Table 1.

Tryptic digests of all three components give relatively simple peptide mixtures as judged by DEAE-cellulose chromatography and paper electrophoresis; hence it was thought these preparations were sufficiently homogeneous for sequence studies.

Examination of Fig. 3 shows that the 'fingerprints' of the tryptic peptides from the three components are very similar. Sequence studies have shown that peptides 1a, 1b and 2 are identical in the three components (Fig. 4), whereas peptides 4 and 6a show

very similar amino acid compositions in each component (Table 2). [Peptides corresponding to 1a, 1b and 2 have been isolated from component SCMKB-2 (Haylett & Lindley, 1968) but a sequence Gly-Gln-Val-Gly-Ser-Gly-Ser-Ala-Ser-Val-Ser-Arg was found for the peptide corresponding to 1b. This is almost certainly incorrect, being derived from papain and carboxypeptidase digests, whereas the present sequence relies only on dansyl-Edman results and has been determined for each component.]

The release of *S*-carboxymethylcysteine from each component by carboxypeptidase A indicates that this residue is *C*-terminal in each component, and that peptide 6a constitutes the *C*-terminal sequence of each component since all other tryptic peptides have *C*-terminal arginine or tyrosine. This peptide differs slightly in all three components, A and C containing five residues more than B (Table 2). Sequence studies

Fig. 2. Elution profile of component SCMKB-2C from DEAE-cellulose

A column 100cm×1.5cm was used with a linear gradient from 2.0M-urea (800ml) to 2.0M-urea-0.5M-formic acid (800ml). —, Formic acid concentration; ---, E_{280} for the first fractionation; —, E_{280} for a second fractionation under the same conditions of chromatography; the arrow indicates the fractions combined for sequence studies.

Table 1. Amino acid composition of components SCMKB-2A, SCMKB-2B and SCMKB-2C

The results are expressed as mol of amino acid/mol of component. Abbreviations in Tables and Figures:

Cm

|
CmCys or Cys, *S*-carboxymethylcysteine;

Ac

|
Ac-Ala or Ala, *N*-acetylalanine.

Amino acid residue	Component		
	A	B	C
Arg	6.0	6.0	7.0
CmCys*	36.0	33.1	32.8
Asp	1.5	1.2	1.0
Thr*	16.4	14.0	15.2
Ser*	24.0	23.3	21.9
Glu	19.6	17.3	15.1
Pro	16.5	13.7	13.0
Gly	15.0	15.5	13.8
Ala	5.0	5.1	6.0
Val	5.6	6.7	5.5
Ile	7.6	6.9	4.8
Leu	3.1	3.1	3.2
Tyr	3.7	3.9	3.7
Phe	2.4	1.9	1.9
Trp†	1	1	1
<i>N</i> -Terminus	Ac-Ala	Ac-Ala	Ac-Ala
<i>C</i> -Terminus	CmCys	CmCys	CmCys
Mol.wt.	20000	18000	18000

(approx. value)

* Values obtained after 24 h hydrolysis and uncorrected for loss by destruction.

† From sequence studies.

Fig. 3. Elution profiles from DEAE-cellulose of tryptic digests of components SCMKB-2A, SCMKB-2B and SCMKB-2C of high-sulphur protein

DE-52 DEAE-cellulose was used in a 100cm×1.5cm column. The flow rate was 1.0ml/min, with a linear gradient from 0.15% pyridine to 3% (v/v) pyridine–30% (v/v) acetic acid. The mobilities (m) of the peptides (●) determined from electrophoretograms at pH6.5 relative to aspartic acid (–1.0) at pH6.5 are shown below the corresponding fractions. (a), (b) and (c) are the profiles of components A, B and C respectively.

on these peptides (see Supplementary Publication SUP 50008, Tables 1–3) show the additional residues to be located at the C-terminus (Fig. 5.)

Peptide 4 released tyrosine when incubated with carboxypeptidase A, indicating this residue to be C-terminal and suggesting some chymotrypsin-like activity in the diphenylcarbonyl chloride-treated trypsin sample. This has been confirmed in later studies with 1-chloro-4-phenyl-3-L-toluene-*p*-sulphamidobutan-2-one- (TPCK-) treated trypsin (Worthington) which failed to produce peptides 4 and 6a, but each component gave a single new fragment eluting later than peptide 6a from the DEAE-cellulose column. On digestion with diphenylcarbonyl chloride-treated trypsin these fragments produced peptides of identical amino acid composition, and *N*- and C-terminal end groups identical with those of peptides 4 and 6a of the corresponding protein.

The differences between the tryptic digests of components A, B and C can be rationalized largely by the presence of an arginine residue in component C (residue 25 of Fig. 6) that is absent in components A and B. Cleavage adjacent to this residue by trypsin results in replacement of peptide 7 of components A

Peptide 1a Thr-Arg

Peptide 1b Gly-Gln-Val-Gly-Ser-Ser-Gly-Ala-Val-Ser-Ser-Arg

Peptide 2 (Trp)-CmCys-Arg-Pro-Asp-CmCys-Arg

Fig. 4. Amino acid sequence of peptides 1a, 1b and 2

Sequence determination was by the dansyl-Edman technique, shown by $\rightarrow$; the C-terminal amino acids remaining were identified as the free amino acid, shown by $\Rightarrow$. Peptide 2 was Ehrlich-positive, and the *N*-terminal tryptophan has been placed in parentheses.

and B by peptides 6b and 3 of component C (Table 3), and peptide 8 of A and B, by peptides 6b and 5 of C. Peptides 5 and 8 originate by non-specific tryptic cleavage of the tyrosine-glycine bond in peptide 3 of component C and peptide 7 of components A and B, respectively (see Fig. 7 and Table 4).

The amino acid sequence of component C (Elleman, 1971) is shown in Fig. 6, with the cor-

responding tryptic peptides indicated below; a simplified structural plan is shown in Fig. 7 to clarify the relationship between the tryptic digests of the components. All members of the group are found to be built on the same general plan. All yield on chymotryptic digestion the fragment Ac-Ala-CmCys-CmCys-Ser-Thr-Ser-Phe (see Supplementary Publication SUP 50008, Table 4, for evidence of this sequence in component C), characteristic of the

SCMK-B2 family (Gillespie *et al.*, 1968). Confirmation of the *N*-acetylalanine residue of this peptide was obtained in the present work by mass spectrometric techniques. The spectra obtained were too complex to permit deduction of the complete sequence of the heptapeptide but gave strong peaks at *m/e* 43 and 128 that correspond to acetyl and acetyl-methylalanyl ions.

The *N*-terminal region of component C (Fig. 6)

Table 2. Amino acid composition of tryptic peptides 4 and 6a from components A, B and C

Results are expressed as mol of amino acid/mol of peptide.

Component ...	Peptide 4			Peptide 6a		
	A	B	C	A	B	C
Arg	—	—	—	2.1	1.8	1.7
CmCys	4.6	4.5	4.6	11.9	10.2	11.0
Asp	—	—	—	—	—	—
Thr	1.8	1.8	2.1	1.9	2.0	2.8
Ser*	3.0	3.9	3.8	3.8	2.8	3.9
Glu	2.1	2.1	2.2	6.0	5.0	5.8
Pro	4.0	3.2	3.1	6.0	5.0	6.0
Gly	1.2	1.1	1.0	1.0	1.3	1.2
Ala	—	—	—	2.8	2.6	3.0
Val	2.2	2.6	2.3	0.8	1.1	0.9
Ile	—	—	—	1.0	1.1	1.1
Leu	1.8	2.0	1.8	—	—	—
Tyr	0.9	1.0	1.0	1.3	1.3	1.4
Phe	—	—	—	0.4	—	—
<i>N</i> -Terminus	Val	Val	Val	Tyr	Tyr	Tyr
<i>C</i> -Terminus	Tyr	Tyr	Tyr	CmCys	CmCys	CmCys

* Corrected for loss by destruction (+11%).

Fig. 5. Comparison of amino acid sequences from the *C*-terminal ends of high-sulphur wool proteins

Homologous residues are underlined. Components IIIB₂ (Haylett & Swart, 1969), IIIB₃ and IIIB₄ (Haylett *et al.*, 1971) and SCMK-B2A were isolated from Merino wool, components SCMK-B2B and SCMK-B2C from Lincoln wool. The *C*-terminal sequence of component SCMK-B2A is derived from thermolytic fragments of peptide A6a, which have been aligned by homology with the sequences of peptides B6a and C6a.

Table 3. *Amino acid composition of tryptic peptides 6b and 3 from component C and peptides 7 from A and B*

Results are expressed as mol of amino acid/mol of peptide.

Amino acid residue	Peptide				
	C _{6b}	C ₃	C _{6b} +C ₃	B ₇	A ₇
Arg	1.0	1.0	2.0	1.0	1.0
CmCys	7.0	7.8	14.8	15.5	17.2
Asp	—	—	—	—	—
Thr	3.0	6.5	9.5	10.0	12.7
Ser*	4.6	10.9	15.5	18.4	18.6
Glu	—	7.0	7.0	10.1	12.0
Pro	1.2	2.2	3.4	4.9	7.3
Gly	3.0	8.6	11.6	13.4	12.6
Ala	2.0	1.0	3.0	2.2	2.0
Val	—	1.9	1.9	3.2	2.0
Ile	1.0	3.0	4.0	6.0	6.7
Leu	—	1.0	1.0	1.0	1.0
Tyr	—	1.0	1.0	0.9	0.9
Phe	1.7	—	1.7	1.9	2.0
N-Terminus	Nil	Ser		Nil	Nil
C-Terminus	Arg	Arg		Arg	Arg

* Corrected for loss by destruction (+11%).

Table 4. *Amino acid compositions of tryptic peptide 5 from component C and peptide 8 from A and B*

Results are expressed as mol of amino acid/mol of peptide.

Amino acid residue	Peptide		
	C ₅	B ₈	A ₈
Arg	—	—	—
CmCys	7.5	15.2	18.3
Asp	—	—	—
Thr	6.6	9.6	12.7
Ser*	6.5	13.3	14.3
Glu	6.0	9.8	10.7
Pro	2.3	5.0	7.5
Gly	6.0	10.1	10.3
Ala	—	1.3	1.3
Val	—	1.2	—
Ile	3.0	5.6	6.1
Leu	1.0	1.0	1.0
Tyr	1.0	0.9	1.0
Phe	—	2.0	2.0
N-Terminus	Ser	Nil	Nil
C-Terminus	Tyr	Tyr	Tyr

* Corrected for loss by destruction (+11%).

contains a partial triplication of a ten-residue sequence between residues 21 and 52. Amino acid analyses of component B, tryptic peptides 7 and 8 from component B (Table 3), and the fragments from

thermolytic digestion of these peptides (Fig. 8) indicate that component B differs from C in containing an extra repetition of this decapeptide. Similarly the amino acid composition and larger size of component A, and the amino acid composition of tryptic peptide 7 from component A suggest that this component may differ from B in containing yet a further repetition of this decapeptide.

Discussion

The isolation of pure components from solubilized wool proteins has proved extremely difficult. The only derivatives possessing reasonable solubility in aqueous solvents (keratases in which cystine residues have been converted into cysteic acid, *S*-carboxymethyl and *S*-aminoethyl derivatives) all have a high net charge and hence their isoelectric points tend to be at extremes of pH. Possibly because of this, fractionation procedures operating at neutral pH values (Crewther *et al.*, 1965) have not been very successful and we feel that our success is attributable to the use of chromatography on SE-Sephadex or DEAE-cellulose at pH values not far removed from the isoelectric points of the *S*-carboxymethyl derivatives of the high-sulphur proteins. As is evident from information in this and previous papers (Darskus *et al.*, 1969; Lindley *et al.*, 1971), the proteins are chemically very similar indeed, so that the problems of separation are necessarily difficult, and only near the isoelectric point do the differences in prop-

Fig. 6. Complete amino acid sequence of high-sulphur protein SCMKB2C (Elleman, 1971)

The peptides produced by cleavage with diphenyl-carbamoyl chloride-treated trypsin are shown. X indicates the position of a partial cleavage in peptide 3 which produces fragments 5 and 1b.

Fig. 7. Structural plan of the SCMKB2 family of the high-sulphur protein fraction from wool, showing the arrangement of the tryptic peptides

The numbers immediately below the boxes show the residue numbers for component SCMKB2C, and the other numbers are those of the peptides shown in Fig. 3.

erties become significant enough for satisfactory fractionation.

A criterion of purity of a wool high-sulphur protein is very difficult to obtain; no criterion of homogeneity other than complete determination of sequence seems to suffice. All proteins of this type isolated so far have the *N*-terminal sequence Ac-Ala-CmCys-CmCys, and *S*-carboxymethylcysteine is *C*-terminal. Within each group molecular size is approximately uniform, as is amino acid composition. The complete absence of lysine and histidine and the isolation of some peptides in theoretical yield (Haylett & Lindley, 1968) suggested that the SCMKB2 preparation of Gillespie (1963) was homogeneous, but this is now seen to be incorrect. Polyacrylamide-gel electrophoresis (7½% and 15% acrylamide) does not show the microheterogeneity demonstrable by chromatography on DEAE-cellulose with the formic acid gradient. This microheterogeneity is reflected in the amino acid composition of component A which has so far resisted further resolution by the techniques successfully applied to components B and C. Thus the experimentally determined number of residues per mol of aspartic acid and phenylalanine in analyses of component A are consistently and significantly different from whole numbers.

Our current view of the SCMKB2 family of proteins is that it comprises three distinct subfractions differing both in size and also in certain residues. Within each subfraction further microheterogeneity occurs. The situation is reminiscent of the variable section of the immunoglobulin chains, although it is less complex, and the mechanism by which the heterogeneity has arisen is most simply explained by a multiplicity of genes differing by partial gene duplications and point mutations.

The amount of this group of proteins in wool is of interest since it is clear that they represent a considerable portion of the high-sulphur protein fraction of wool. As isolated from the SE-Sephadex column, each of the fractions A, B and C, represents approx. 10% of the high-sulphur protein, although each fraction is microheterogeneous to some extent, so that final yields of pure protein from the DEAE-cellulose column are considerably decreased.

m	Composition	Sequence
-1.00	CmCys(2.0); Thr(0.9); Ser(1.6); Ala(1.0)	Ninhydrin-negative N-terminal fragment. Sequence not determined.
-0.40	CmCys(1.0); Pro(1.0); Gly(1.2); Phe(2.0)	$\begin{array}{c} \text{Cm} \\ \\ \text{Phe-Cys-Gly-Phe-Pro} \end{array}$
-0.49	CmCys(1.0); Ser(1.8); Ile(1.0)	$\begin{array}{c} \text{Cm} \\ \\ \text{Ile-Cys-Ser-Ser} \end{array}$
-0.71	CmCys(5.0); Thr(3.9); Ser(4.6); Glu(3.0); Pro(2.0); Gly(2.8); Val(0.8)	$\begin{array}{c} \text{Cm} \qquad \text{Cm} \qquad \text{Cm} \\ \qquad \qquad \\ \text{Val-Gly-Thr-Cys-Gly-Ser-Ser-Cys-Gly-Gln-Pro-(Thr, Cys,} \\ \text{Ser, Gln, Thr, Ser, Cys, Cys, Gln, Pro, Thr)-Ser} \end{array}$
-0.55	CmCys(1.6); Thr(1.0); Ser(0.9); Glu(2.0); Pro(1.0); Ile(1.0)	$\begin{array}{c} \text{Cm} \qquad \text{Cm} \\ \qquad \\ \text{Ile-Gln-Thr-Ser-Cys-Gln-Pro} \end{array}$
-0.58	CmCys(2.8); Thr(1.8); Ser(1.8); Glu(2.0); Pro(1.2); Ile(2.0)	$\begin{array}{c} \text{Ile-Ser-} \\ \\ \text{Ile-Gln-Thr-Ser-Cys-Gln-Pro-Thr-Cys} \end{array}$
-0.65	CmCys(2.1); Thr(2.1); Ser(2.1); Glu(2.2); Gly(4.6); Ile(1.1); Leu(1.0)	$\begin{array}{c} \text{Cm} \qquad \text{Cm} \qquad \text{Cm} \\ \qquad \qquad \\ \text{Leu-Gln-Thr-Ser-Gly-Cys-Glu-Thr-Gly-Cys-} \\ \text{Gly-Ile-Gly (Gly, Ser)} \end{array}$
0.00	Gly(1.0); Ile(1.0); Tyr(1.0)	$\begin{array}{c} \text{Ile-Gly-Tyr} \end{array}$

Fig. 8. Some thermolytic peptides obtained from tryptic peptide 8 of component B

Residues determined by the dansyl-Edman method are shown by $\rightarrow$; remaining C-terminal amino acids identified as the Dns derivative without hydrolysis are shown by $\leftarrow$. Residues released by carboxypeptidase A are shown by $\leftarrow$. The mobility (m) is relative to aspartic acid (-1.0).

An interesting homology is found when the C-terminal sequences of this group of proteins are compared with those of another group of lower-molecular-weight, high-sulphur proteins from wool studied by Haylett *et al.* (1971). The significant features are shown in Fig. 5. Obvious homology between these two groups is far less extensive in the N-terminal region, being limited to the sequence Ac-Ala-CmCys-CmCys and is absent from the remaining regions of the molecules.

We are deeply indebted to Dr. A. P. Damoglou and Dr. C. H. Moore for their help in the initial stages of this work, and also to Dr. Bhupesh C. Das for mass spectrometry of the N-terminal heptapeptide while one of us (H. L.) was a guest worker in Professor E. Lederer's laboratory at Gif-sur-Yvette, France.

References

- Ambler, R. P. (1963) *Biochem. J.* **89**, 349
Catravas, G. N. (1964) *Anal. Chem.* **66**, 1146
Crewther, W. G., Fraser, R. D. B., Lennox, F. G. & Lindley, H. (1965) *Advan. Protein Chem.* **20**, 191
Darskus, R. L., Gillespie, J. M. & Lindley, H. (1969) *Aust. J. Biol. Sci.* **22**, 1197
Davis, B. J. (1964) *Ann. N.Y. Acad. Sci.* **121**, 405
Elleman, T. C. (1971) *Nature New Biol. (London)* **234**, 148
Gillespie, J. M. (1962) *Aust. J. Biol. Sci.* **15**, 262
Gillespie, J. M. (1963) *Aust. J. Biol. Sci.* **16**, 241
Gillespie, J. M., Haylett, T. & Lindley, H. (1968) *Biochem. J.* **110**, 193
Gray, W. R. (1967) *Methods Enzymol.* **11**, 139
Haylett, T. & Lindley, H. (1968) *Biochem. J.* **110**, 198
Haylett, T. & Swart, L. S. (1969) *Text. Res. J.* **39**, 917
Haylett, T., Swart, L. S. & Parris, D. (1971) *Biochem. J.* **123**, 191
Lindley, H., Gillespie, J. M. & Haylett, T. (1968) *Symp. Fibrous Proteins, Int. Conf. 3rd*, 1967, p. 353
Lindley, H., Broad, A., Damoglou, A. P., Darskus, R. L., Elleman, T. C., Gillespie, J. M. & Moore, C. H. (1971) *Appl. Polym. Symp.* **18**, 21
Offord, R. E. (1966) *Nature (London)* **211**, 591
Reindel, F. & Hoppe, W. (1954) *Chem. Ber.* **87**, 1103
Roepstorff, P., Norris, K., Severinsen, S. & Brunfeldt, K. (1970) *FEBS Lett.* **9**, 235
Thomas, D. W. (1969) *FEBS Lett.* **5**, 63
Yphantis, D. A. (1964) *Biochemistry* **3**, 297