Aust. J. Biol. Sci., 1979, **32**, 437–42

Comparison of Some Microfibrillar Proteins from Wool

Lyndsay M. Dowling, Keith H. Gough, Adam S. Inglis and Lindsay G. Sparrow

Division of Protein Chemistry, CSIRO, Parkville, Vic. 3052.

Abstract

The molecular properties of components 5 and 7c, two low-sulfur proteins from the microfibril of wool, have been compared. These proteins have almost identical molecular weights and very similar amino acid compositions. Comparison of the partial amino acid sequences determined to date show extensive regions of sequence homology. No such sequence homologies are observed when these components are compared with component 8c-1, another low-sulfur protein from wool.

Introduction

The low-sulfur proteins of wool (SCMKA) are the major constituents of that fraction of the reduced, carboxymethylated kerateine (SCMK) which is precipitated at pH 4·4 in the presence of salt. They are a mixture of similar proteins that are relatively rich in the amino acids aspartic acid, glutamic acid and leucine as compared to wool and have molecular weights in the range 40 000–60 000 (Crewther 1976). They have a significant content of α-helix in solution and appear to be derived from the microfibrils of the fibre (Jones 1975, 1976). Limited chymotryptic digestion of SCMKA yields protein fragments with enhanced α-helix content. From these fragments can be obtained two kinds of helical segments, type I and type II, of molecular weight approximately 12 000 (Crewther and Dowling 1971). The amino acid sequence of one unit from each of these groups has been determined (Crewther *et al.* 1978; Gough *et al.* 1978).

Electrophoretic studies have shown that there are eight protein species, in approximately equal amounts, in SCMKA. Based on their electrophoretic mobilities these have been designated components 5, 7a, 7b, 7c, 8a, 8b, 8c-1 and 8c-2 (Crewther *et al.* 1976). Marshall and Gillespie (1977) have speculated, from data obtained from polyacrylamide gel electrophoresis on extracts of various keratinous tissues from sheep, that component 5 is functionally equivalent to one or more of the components 7.

In this study we compare components 5, 7c and 8c-1 and show that components 5 and 7c are similar, probably homologous proteins, but are different from component 8c-1.

Experimental

Low-sulfur Proteins

Components 5, 7c and 8c-1 were prepared by the method of Crewther *et al.* (1976).

Cyanogen Bromide Cleavage

Cyanogen bromide digestion was carried out in 70% (v/v) formic acid (Gross and Witkop 1961). The reaction mixture was kept in the dark at room temperature for 16 h; it was then diluted 1:5 with water and lyophilized. The peptides were then fractionated by chromatography on Sephadex G100 (component 7c) or G75 (components 5 and 8c-1) in 8 M urea 0·01 M in tris, 0·05 M in ethanolamine, pH 7·4.

Tryptic Digestion

Protein was citraconylated by the method of Atassi and Habeeb (1972) and digested with TPCK-trypsin (Worthington) in 1% (w/v) ammonium bicarbonate pH 8·5 for 4 h at 37°C using an enzyme to protein ratio of 1:100. Peptides were then de-blocked by incubation at pH 3·5 for 4 h at 45°C and fractionated on columns of Sephadex G50 superfine (300 cm by 1 cm) coupled to G25 superfine (300 cm by 1 cm) eluting with 0·01 M ammonium bicarbonate buffer containing 10% (v/v) isopropanol (Corran and Waley 1974). Further purification was then achieved by ion-exchange chromatography or by high voltage paper electrophoresis in Varsol-cooled tanks (Ambler 1963). Some lysine residues apparently were not citraconylated as several peptides with C-terminal lysines were obtained.

Sequence Determination

Peptides were sequenced on an automatic protein sequenator (Edman and Begg 1967) and the phenylthiohydantoin amino acid derivatives identified by thin-layer chromatography (t.l.c.) (Inglis and Nicholls 1973). Manual peptide sequencing was performed using the dansyl-Edman technique (Hartley 1970). The dansyl amino acids were identified by t.l.c. on polyamide sheets (Woods and Wang 1967; Hartley 1970).

Amino Acid Analysis

1–2 mg of protein was hydrolysed *in vacuo* at 108°C for 24 h in 0·5 ml constant-boiling HCl containing 50 μl of 0·1 M phenol. For tryptophan determination protein was hydrolysed at 115°C for 24 h in *p*-toluenesulfonic acid containing 2 mg/ml tryptamine (Simpson *et al.* 1976). The hydrolysates were analysed on a Beckman model 120 C automatic amino acid analyser using a single column.

Sodium Dodecyl Sulfate–Polyacrylamide Gel Electrophoresis

This was carried out using the discontinuous system described by Laemmli (1970).

Results and Discussion

Molecular Weights and Amino Acid Analyses

The values for molecular weights of the individual low-sulfur protein components are still open to some doubt, however all the techniques used for molecular weight measurement [sodium dodecyl sulfate (SDS)–polyacrylamide gel electrophoresis, gel filtration, ultracentrifugation] suggest that components 5 and 7c are almost identical in molecular weight (56 000–58 000) (Crewther *et al.* 1976) and higher than component 8c-1 which is in the range 45 000–50 000 (Woods 1979).

The amino acid compositions of components 5, 7c and 8c-1 are shown in Table 1. It is evident that the compositions of components 5 and 7c are similar but different from that of component 8c-1 although all three components have compositions that are typical of wool low-sulfur proteins, being rich in aspartic acid, glutamic acid and leucine.

Sequence Studies

The strategy adopted for sequencing these three low-sulfur proteins involves a preliminary cleavage at methionine with cyanogen bromide, followed by subsequent

Table 1. Amino acid composition of components 5, 7c, and 8c-1 and of the large cyanogen bromide peptides [CNBr-1] derived from them

Values given as residues per 100 residues and are uncorrected

Amino acid	Component			CNBr-1 fragment		
	5	7c	8c-1	5	7c	8c-1
Ala	8·4	8·1	5·8	7·6	7·2	5·7
Arg	8·0	7·6	8·3	7·2	7·7	8·6
Asp	7·3	8·3	11·3	8·4	8·1	10·9
Glu	13·1	15·0	18·1	13·3	14·9	18·0
Gly	8·0	8·7	3·7	7·5	8·4	3·6
CMCys	4·5	6·1	5·8	4·2	5·3	5·5
His	1·1	0·5	0·6	1·3	0·4	0·7
Hse	—	—	—	0·3	0·4	—
Ile	4·0	4·1	3·7	3·9	4·4	3·7
Leu	8·2	9·3	12·3	9·4	9·3	12·5
Lys	4·8	4·7	2·9	5·2	5·2	2·9
Met	0·9	0·7	0·3	—	—	—
Phe	3·6	2·9	2·1	4·1	4·5	1·6
Pro	3·3	2·6	4·0	3·2	3·4	3·7
Ser	9·2	7·1	7·0	9·7	7·7	7·0
Thr	4·4	4·0	5·3	3·8	3·8	5·5
Trp	0·4	0·2	0·4	0·3	—	0·3
Tyr	3·7	3·0	2·8	4·1	3·1	3·4
Val	6·7	7·0	6·0	6·9	6·3	6·4

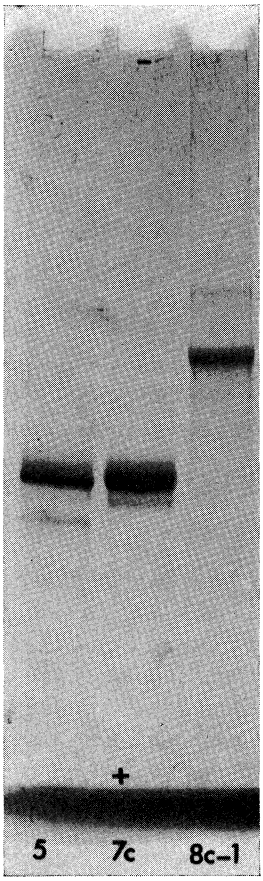

Fig. 1. SDS-polyacrylamide slab gel electrophoresis pattern of the large cyanogen bromide peptides (CNBr-1) of components 5, 7c and 8c-1 using 10% (w/v) acrylamide concentration.

isolation and sequencing of the resulting peptides. From each of the components a large peptide (CNBr-1) is obtained by cyanogen bromide cleavage and these have all been well characterized. Those from components 5 and 7c are very similar to each other in molecular weight (approximately 28 000–30 000) as shown by SDS-polyacrylamide gel electrophoresis (Fig. 1), have similar amino acid compositions (Table 1) and have blocked *N*-termini as have the components themselves thus establishing that CNBr-1 from components 5 and 7c are *N*-terminal. The blocked, *N*-terminal cyanogen bromide peptide of component 8c-1 (1 methionine residue per mole only) has a molecular weight of 7 000 (O'Donnell 1968) so that CNBr-1 from component 8c-1, which has a molecular weight of 38 000–40 000 (Fig. 1), and contains no homoserine, is the *C*-terminal peptide of component 8c-1.

Table 2. Homologous amino acid sequences from components 5 and 7c

No.	Component	Mol. wt[A]	Sequence
1	5	1 000	Ala-CMCys-Leu-Leu-Lys-Glu-Tyr-Gln-Glu-Val-Hse
	7c	1 000	Ala-CMCys-Leu-Leu-Lys-Glu-Tyr-Gln-Glu-Val-Hse
2	5	28 000	CMCys-Glu-Ile-Ser-Ala-Ala-Pro-Tyr-Arg
	7c	28 000	CMCys-CMCys-Ile-Thr-Ala-Ala-Pro-Tyr-Ala-Arg
3	5	28 000	Phe-Leu-Ala-Phe-Ile-Asx-Lys
	7c	28 000	Phe-Ala-Ala-Phe-Ile-Asp-Lys-Val-Arg
4	5	28 000	Gln-Gln-Ile-Lys
	7c	28 000	Gln-Gln-Ile-Lys
5	5	28 000	Asp-Val-Asp-CMCys-Ala-Tyr-Leu-Arg
	7c	28 000	Asp-Val-Asp-CMCys-Ala-Tyr-Val-Arg
6	5	28 000	Glu-Ala-Glu-His-Val-Glx-Ala-Asp-Ser-Gly-Arg
	7c	28 000	Glu-Ala-Glu-CMCys-Val-Glu-Ala-Asp-Ser-Gly-Arg
7	5	28 000	Lys-Tyr-Glu-Glu-Glu-Val-Ala-Leu-Arg
	7c	28 000	Lys-Tyr-Glu-Glu-Glu-Val-Ala-Leu-Arg
8	5	28 000	Ala-Thr-Ala-Glu-Asn-Glu-Phe-Val-Val-Leu-Lys
	7c	28 000	Ala-Thr-Ala-Glu-Asn-Glu-Phe-Val-Val-Leu-Lys
9[B]	5	3 000	Asp-CMCys-Val-Val-Ala-Glu-Ile-Lys-Ala-Gln-Tyr-Asp-Asp-Ile-Ala-Ser-Arg-Ser-Ala-Glu-Ala-Glu-Ser-Trp-Tyr-Arg-Ser-Lys-CMCys-Gln-Gln-Hse
	7c	16 000	Ala-Gln-Tyr-Asp-Asp-Ile-Ala-Ser-Arg-Ser-Ala-Glu-Ala-Glu-Ser-Trp-Tyr-Arg-Ser-Lys-CMCys-Gln-Glu-Ile

[A] Refers to the molecular weight of the cyanogen bromide fragments in which these sequences were found.

[B] Tryptic peptides were ordered and the sequences confirmed by sequenator analysis of the cyanogen bromide fragments.

From component 7c three additional cyanogen bromide peptides are obtained. Tentative molecular weights for these, based on amino acid analysis, are 16 000, 9 000 and 1 000. From component 5 a total of six cyanogen bromide peptides are expected but only two others, with molecular weights 3 000 and 1 000, have so far been purified. These have been completely sequenced as has the peptide of molecular weight 1 000 from component 7c. From the remaining cyanogen bromide peptides described a number of tryptic peptides have been isolated and sequenced.

Visual inspection of the partial sequences so far obtained for the three components shows that there are a number of sequence homologies between components 5 and 7c. These are shown in Table 2.

Of the nine peptide pairs shown some (Nos 1, 4 and 7) are identical while the others show several substitutions or deletions. Some of these substitutions are conservative, e.g. Ser for Thr in 2, Leu for Ala in 3, Leu for Val in 5, Val for Ala in 8 and Ile for Met in 9, but others are quite clearly not so, e.g. Glu for CMCys in 2 and His for CMCys in 6. The Ile for Met substitution in peptide pair 9 accounts for one of the extra methionines in component 5.

Peptides 5, 6, 7 and 8 are identical to or homologous with peptides found in a type II helical fragment obtained from limited chymotryptic digestion of SCMKA (Crewther *et al.* 1978). This strongly suggests that an entire type II fragment (mol. wt. 12 000) is present in CNBr-1 from both components 5 and 7c; this has been conclusively shown for component 7c by the isolation of all the tryptic peptides corresponding to peptides from a type II fragment. Peptides 2, 3 and 4 are also from these same CNBr peptides but must originate from outside the type II region. The remaining two homologous peptide pairs, 1 and 9, are from other cyanogen bromide peptides.

The order of the cyanogen bromide peptides in component 7c is 28 000, 16 000, 1 000, 9 000 (that for component 5 is not yet established) so there are peptides from component 5 that are identical to or homologous with peptides from widely different parts of component 7c. This suggests that not only are there discrete regions of these molecules that are homologous but also that their overall structure is very similar.

About 350 of the approximately 430 residues of component 8c-1 have been sequenced and no sequences have been found that are homologous with any from components 5 or 7c.

The similarities in amino acid composition, almost identical molecular weights and extensive sequence homologies described above all suggest that components 5 and 7c are homologous proteins and support the suggestion of Marshall and Gillespie (1977) that they are functionally equivalent. The complete amino acid sequences of these proteins will clarify this further and presumably will also indicate why the two components have quite different contents of α-helix in solution (26% for component 5 and 50% for component 7c—Crewther *et al.* 1976).

References

Ambler, R. P. (1963). The amino acid sequence of *Pseudomonas* cytochrome *c*-551. *Biochem. J.* **89**, 349–78.

Atassi, M. Z., and Habeeb, A. F. S. A. (1972). Reactions of proteins with citraconic anhydride. In 'Methods in Enzymology'. Vol. XXV. Part B. pp. 546–53. (Eds C. H. W. Hirs and S. N. Timasheff.) (Academic Press: New York.)

Corran, P. H., and Waley, S. G. (1974). The tryptic peptides of rabbit muscle triose phosphate isomerase. *Biochem. J.* **139**, 1–10.

Crewther, W. G. (1976). Primary structure and chemical properties of wool. Proc. 5th Int. Wool Text. Res. Conf., Aachen, 1975. (Ed. K. Ziegler.) Vol. I. pp. 1–100.

Crewther, W. G., and Dowling, L. M. (1971). The preparation and properties of large peptides from the helical regions of the low-sulphur proteins of wool. Applied Polymer Symposium No. 18 pp. 1–20.

Crewther, W. G., Dowling, L. M., Gough, K. H., Inglis, A. S., McKern, N. M., Sparrow, L. G., and Woods, E. F. (1976). The low-sulphur proteins of wool: studies on their classification, characterization, primary and secondary structure. Proc. 5th Int. Wool Text. Res. Conf., Aachen, 1975. (Ed. K. Ziegler.) Vol. 2. pp. 233–42.

Crewther, W. G., Inglis, A. S., and McKern, N. M. (1978). Amino acid sequences of α-helical segments from *S*-carboxymethyl-kerateine-A. II. Complete sequence of a type 1I segment. *Biochem J.* **173**, 365–71.

Edman, P., and Begg, G. (1967). A protein sequenator. *Eur. J. Biochem.* **1**, 80–91.
Gough, K. H., Inglis, A. S., and Crewther, W. G. (1978). Amino acid sequences of α-helical segments from *S*-carboxymethyl-kerateine-A. III. Complete sequence of a type I segment. *Biochem. J.* **173**, 373–85.
Gross, E., and Witkop, B. (1961). Selective cleavage of the methionyl peptide bonds in ribonuclease with cyanogen bromide. *J. Am. Chem. Soc.* **83**, 1510–1.
Hartley, B. S. (1970). Strategy and tactics in protein chemistry. *Biochem. J.* **119**, 805–22.
Inglis, A. S., and Nicholls, P. W. (1973). Identification of phenylthiohydantoins of amino acids by thin-layer chromatography. *J. Chromatogr.* **79**, 344–6.
Jones, L. N. (1975). The isolation and characterization of α-keratin microfibrils. *Biochim. Biophys. Acta* **412**, 91–8.
Jones, L. N. (1976). Studies on microfibrils from α-keratin. *Biochim. Biophys. Acta* **446**, 515–24.
Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**, 680–5.
Marshall, R. C., and Gillespie, J. M. (1977). The keratin proteins of wool, horn and hoof from sheep. *Aust. J. Biol. Sci.* **30**, 389–400.
O'Donnell, I. J. (1968). Studies on reduced wool. VIII. *N*-Acetyl peptides isolated from a major component. *Aust. J. Biol. Sci.* **21**, 385–93.
Simpson, R. J., Neuberger, M. R., and Liu, T-Y (1976). Complete amino acid analysis of proteins from a single hydrolysate. *J. Biol. Chem.* **251**, 1936–40.
Woods, E. F. (1979). Microfibrillar proteins of wool: partial specific volumes and molecular weights in denaturing solvents. *Aust. J. Biol. Sci.* **32**, 423–35.
Woods, K. R., and Wang, K. T. (1967). Separation of dansyl-amino acids by polyamide layer chromatography. *Biochim. Biophys. Acta* **133**, 369–70.

Manuscript received 29 March 1979, accepted 7 June 1979