The amino acid sequence of component 7c, a type II intermediate-filament protein from wool

Lindsay G. SPARROW,* Christine P. ROBINSON, David T. W. McMAHON† and Michael R. RUBIRA‡
C.S.I.R.O. Division of Wool Technology, 343 Royal Parade, Parkville, Vic. 3052, Australia

Component 7c is one of the four homologous type II intermediate-filament proteins that, by association with the complementary type I proteins, form the microfibrils or intermediate filaments in wool. Component 7c was isolated as the S-carboxymethyl derivative from Merino wool and its amino acid sequence was determined by manual and automatic sequencing of peptides produced by chemical and enzymic cleavage reactions. It is an N-terminally blocked molecule of 491 residues and M_r (not including the blocking group) of 55600; the nature of the blocking group has not been determined. The predicted secondary structure shows that component 7c conforms to the now accepted pattern for intermediate-filament proteins in having a central rod-like region of approximately 310 residues of coiled-coil α -helix flanked by non-helical N- and C-terminal regions. The central region is divided by three non-coiled-coil linking segments into four helical segments 1A, 1B, 2A and 2B. The N- and C-terminal non-helical segments are 109 and 71 residues respectively and are rich in cysteine. Details of procedures used in determining the sequence of component 7c have been deposited as a Supplementary Publication SUP 50152 (65 pages) at the British Library Document Supply Centre, Boston Spa, Wetherby, West Yorkshire LS23 7BQ, U.K., from whom copies can be obtained on the terms indicated in Biochem. J. (1989) 257, 5. The information comprises: (1) details of chemical and enzymic methods used for cleavage of component 7c, peptides CN1, CN2 and CN3 and various other peptides, (2) details of the procedures used for the fractionation and purification of peptides from (1), including Figures showing the elution profiles from the chromatographic steps used, (3) details of methods used to determine the C-terminal sequence of peptide CN3, and (4) detailed evidence to justify a number of corrections to the previously published sequence.

INTRODUCTION

Wool, like other hard keratins, is made up of filamentous structures, the microfibrils, embedded in an amorphous matrix (Fraser *et al.*, 1972). The proteins that constitute the fibre are highly cross-linked by disulphide bonds, and are usually solubilized by reduction in disaggregating media followed by alkylation of the resultant thiols with iodoacetate. The resulting mixture of reduced carboxymethylated proteins can be readily fractionated into three classes: (i) proteins poorer in cysteine than the original fibre and derived from the microfibrils (Jones, 1976), (ii) proteins richer in cysteine than the original fibre and derived from the amorphous matrix, and (iii) proteins rich in glycine and tyrosine, also derived from the matrix (Crewther, 1976). Each of these classes is a complex mixture of proteins. A total of 20 matrix proteins has been purified and their amino acid sequences have been determined (Dopheide, 1973; Swart *et al.*, 1976; Marshall *et al.*, 1980).

Electrophoretic and chromatographic studies of the wool microfibrillar proteins, as their reduced carboxymethylated derivatives, show eight major species in the mixture; these can be divided into two families, each of four proteins. The two families are the type I (components

8a, 8b, 8c-1 and 8c-2) and type II (components 5, 7a, 7b and 7c) wool microfibrillar proteins (Crewther *et al.*, 1980). By ion-exchange chromatography in buffered 8 M-urea with the use of shallow salt gradients, components 5, 7c, 8a and 8c-1 can be isolated as relatively homogeneous proteins (Crewther *et al.*, 1976). The complete amino acid sequence of component 8c-1 has been determined (Dowling *et al.*, 1986), and a partial sequence of component 7c (Sparrow & Inglis, 1980) and some limited sequence data for component 5 (Dowling *et al.*, 1979) have been presented. Comparison of this sequence information with that available for various intermediate-filament proteins shows that the wool microfibrillar proteins show sequence similarity to intermediate-filament proteins and they are now classified as a sub-class of the cytokeratins (or epidermal intermediate-filament proteins) (Dowling *et al.*, 1983).

An earlier study (Sparrow & Inglis, 1980) established the strategy for obtaining the complete amino acid sequence of component 7c via cleavage at methionine residues with CNBr. Such cleavage yields five peptides, CN1, CN2, CN3, CN4 and CN5, consistent with component 7c itself containing four methionine residues. The complete sequences of peptides CN4 and CN5 have been published, together with partial sequences of peptides

Abbreviations used: h.v.p.e., high-voltage paper electrophoresis; PAGE, polyacrylamide-gel electrophoresis; IEF, isoelectric focusing; CmCys, S-carboxymethylcysteine.

* To whom correspondence should be addressed.

† Present address: Scot's Presbyterian Church, Godfrey Street, Boort, Vic. 3537, Australia.

‡ Present address: Ludwig Institute for Cancer Research, Parkville, Vic. 3050, Australia.

These sequence data have been submitted to the EMBL/GenBank Data Libraries.

CN1, CN2 and CN3 (Sparrow & Inglis, 1980). The complete sequences of peptides CN1, CN2 and CN3 are now presented, together with evidence establishing the order of these peptides in the component 7c molecule. Some amendments to the published partial sequences have been necessitated.

EXPERIMENTAL

Preparation of component 7c and the CNBr-derived peptides

Preparation of component 7c from fine non-Peppin Merino wool (designated MW168), its cleavage at methionine with CNBr and the fractionation of the resulting peptides were carried out as previously described (Sparrow & Inglis, 1980).

Acid cleavage of component 7c

Component 7c was subjected to acid cleavage by incubation in 70% (v/v) formic acid at 37 °C for 72 h. A portion of the resulting cleavage mixture was analysed directly in a spinning-cup Sequenator (Inglis *et al.*, 1981). The remainder was fractionated by gel filtration, and the small cleavage peptide, H1, was further purified by h.p.l.c.

Enzymic cleavages

Unless otherwise specified, all enzymic cleavages were carried out in 1% (w/v) NH_4HCO_3 at 37 °C. Cleavage with clostripain was carried out in sodium phosphate buffer, pH 7.1, containing 1 mM-dithiothreitol at room temperature.

Peptide fractionations

Enzyme digests were fractionated either by a single step of reverse-phase h.p.l.c. or by a number of steps involving combinations of gel-filtration chromatography [Sephadex G-50 (superfine grade) and/or Sephadex G-25 (superfine grade), Bio-Sil TSK 125], ion-exchange chromatography (Dowex 50 or anion-exchange h.p.l.c.), reverse-phase h.p.l.c. and h.v.p.e. Gel filtration on Sephadex G-50 (superfine grade) and/or Sephadex G-25 (superfine grade) was carried out on coupled columns each 1 cm $\times$ 150 cm; either two columns of Sephadex G-50 (superfine grade) were used or four columns, two each of Sephadex G-50 (superfine grade) and Sephadex G-25 (superfine grade). The eluent was 10 mM- NH_4HCO_3 containing 10% (v/v) propan-2-ol. Gel-filtration chromatography on Bio-Sil TSK125 was carried out in 10 mM-ammonium formate buffer, pH 7.0, containing 10% (v/v) acetonitrile.

H.p.l.c. was carried out in a Perkin-Elmer Series 4 liquid chromatograph equipped with an LC95 detector. Reverse-phase columns used were either Zorbax ODS (du Pont), Vydac 214TP or Vydac 218TP (The Separations Group, Hesperia, CA, U.S.A.). Two different buffer systems were used at a flow rate, unless otherwise specified, of 1 ml/min.

1. The trifluoroacetic acid system: buffer A was approx 0.1% (v/v) trifluoroacetic acid in water and buffer B was 0.1% (v/v) trifluoroacetic acid in 70% (v/v) acetonitrile; buffer A was adjusted with trifluoroacetic acid to the same absorbance at 214 nm as buffer B.

2. The ammonium formate system: buffer A was approx. 15 mM-ammonium formate buffer, pH 7.0, and buffer B was 25 mM-ammonium formate buffer, pH 7.0,

in 70% (v/v) acetonitrile; buffer A was adjusted with 1 M-ammonium formate solution to the same absorbance at 214 nm as buffer B.

Analytical procedures

Amino acid analyses were carried out on 1–5 nmol portions of peptide hydrolysed for 20–24 h at 108 °C in constant-boiling HCl containing 0.1% (w/v) phenol. Tryptophan determination was carried out after hydrolysis at 115 °C for 24 h in 4 M-methanesulphonic acid containing tryptamine. Hydrolysates were analysed either on a modified Beckman analyser or on a Millipore Waters amino acid analyser. Amino acid sequence analysis was carried out either manually as described by McKern *et al.* (1985) or in a spinning-cup Sequenator as described by Inglis *et al.* (1981). Identification of phenylthiohydantoin derivatives of amino acids was carried out by h.p.l.c. on a Zorbax ODS or Zorbax Gold column (du Pont) on a Hewlett-Packard 1084A liquid chromatograph.

C-Terminal analysis

Identification of C-terminal amino acids was carried out by the method of selective tritiation (Holcomb *et al.*, 1968). Selective tritiation was used to identify both the C-terminal amino acid (after complete acid hydrolysis of the labelled protein) and the C-terminal peptide released by thermolytic digestion of the labelled protein.

RESULTS

The complete amino acid sequence of component 7c is given in Fig. 1, which also shows the peptides used in establishing this sequence, including those used in determining the previously published partial sequence (Sparrow & Inglis, 1980). Each of these peptides was purified and for most of them both amino acid sequence and composition were determined.

The amino acid sequences of the small CNBr-cleavage peptides CN4 and CN5 have been published previously; for the peptides CN1, CN2 and CN3 the sequence determination is summarized below.

Peptide CN1 (residues 1–266)

Mild acid hydrolysis of peptide CN1 or component 7c cleaved the bond Asp-99–Pro-100. The smaller acid-cleavage peptide H1 (residues 1–99) was isolated from the cleavage mixture by gel-filtration chromatography followed by reverse-phase h.p.l.c. No sequence data were obtained by direct sequencing of this peptide as its N-terminus is blocked. Its amino acid sequence was obtained by using data from peptides derived by cleavage of peptide H1 with trypsin, chymotrypsin and thermolysin and of peptide CN1 with the *Staphylococcus aureus* V8 proteinase (an anomalous cleavage at the bond Gly-11–Ser-12), together with sequence data from peptides obtained by thrombin cleavage of component 7c. The nature of the N-terminal blocking group of component 7c has not been determined, nor has the order of the first two amino acid residues.

Sequence analysis of the mixture after acid cleavage of component 7c gave a unique sequence starting at Pro-100 and extending for 38 residues. The remaining sequence of peptide CN1 was obtained by using peptides derived from digestion with trypsin, *S. aureus* V8 proteinase and Endoproteinase Lys-C. The overlap at the

Fig. 1. Amino acid sequence of component 7c

Designation of peptides: —→, sequence only; —, sequence and composition; ----, composition only; ·····, previously published (Sparrow & Inglis, 1980) (no further experimental detail is given for peptides from the previous publications); T, tryptic; C, chymotryptic; Th, thermolytic; S, *S. aureus* V8 proteinase; E, Endoproteinase Lys-C; AP, *A. lyticus* proteinase; Clos, clostripain; AMP, *A. mellea* proteinase; Throm, thrombin; H, acid cleavage; W, oxidative cleavage at tryptophan. Peptides from these digests are numbered in order of their occurrence in the complete sequence.

acid-cleavage point Asp-99-Pro-100 was found in the sequence of the large thrombin-digest peptide starting at residue 65 (Sparrow & Inglis, 1980).

Peptide CN2 (residues 275–385)

This sequence was established from sequences of peptides derived by cleavage of peptide CN2 with trypsin, *S. aureus* V8 proteinase, clostripain and the lysine-specific proteinase from *Armillaria mellea*. The order of these peptides was determined by a Sequenator run on the intact peptide, which gave 43 residues of sequence, and a further Sequenator run on the peptide mixture obtained after cleavage of component 7c at the single tryptophan, residue 299.

Peptide CN3 (residues 397–491)

This peptide lacks homoserine and is thus the C-terminal CNBr-cleavage peptide. Its sequence was obtained from the sequences of peptides derived by cleavage with trypsin together with a Sequenator run from the N-terminal Asn-397. The long tryptic peptide starting at Gly-436 was sequenced through to the residue 483 with the Sequenator; it was further cleaved with clostripain to give the peptide Pro-450-Arg-490, from which all thermolytic peptides were subsequently isolated. The thermolytic peptides were ordered by the automated sequence determination from Gly-436, with the exception of the overlap to establish the bond Ser-487-Leu-488; this overlap was obtained from DNA sequence data for

the clone PSWK3 covering the C-terminal portion of the molecule from Glu-305 (K. Ward, personal communication). The C-terminus of peptide CN3 and therefore of component 7c was established from the following data: (i) free carboxymethylcysteine isolated from the tryptic digest of peptide CN3, (ii) tritiated carboxymethylcysteine found in an acid hydrolysate after C-terminal labelling, and (iii) the tritiated peptide Leu-Gly-Arg-CmCys isolated from a thermolytic digest of C-terminally labelled peptide CN3.

Overlaps to establish the order of the CNBr-cleavage peptides

The four methionine-containing peptides needed to order the CNBr-cleavage peptides were isolated as follows.

Sequence containing Met-266: this peptide, which ran from Val-252 to Arg-270, was obtained from the tryptic digest of peptide CN1; incomplete cleavage by CNBr at Met-266 resulted in the peptide CN1-CN5, which co-purified with CN1 on gel-filtration chromatography.

Sequence containing Met-274: a peptide Met-266-Lys-282 was obtained from a digest of component 7c with the *Achromobacter lyticus* proteinase.

Sequences containing Met-385 and Met-396: peptides Lys-382-Leu-389 and Lys-390-Ser-398 came from digestion of component 7c with *Armillaria mellea* proteinase.

The order of the CNBr-cleavage peptides deduced from these data is CN1-CN5-CN2-CN4-CN3.

Discrepancies between the sequence presented in Fig. 1 and the published partial sequence (Sparrow & Inglis, 1980) were found at positions 102, 199, 247, 264, 275, 280, 282, 335, 380 and 384. These are further discussed in Supplementary Publication SUP 50152.

DISCUSSION

Aspects of the sequence analysis

The sequence determination was both hindered and aided by some unusual cleavage points for certain enzymes. The *S. aureus* V8 proteinase unexpectedly cleaved the bonds Gly-11-Ser-12 (peptide S1), Ser-19-Ala-20 (results not shown), Gly-49-Ser-50 (peptide S2), Arg-64-Ser-65 (peptide S3), Gly-67-Tyr-68 (results not shown) and Ser-433-Ser-434 (results not shown), and Endoproteinase Lys-C gave unexpected cleavages at Ser-65-Phe-66 (results not shown) and Gly-67-Tyr-68 (results not shown). Such cleavages complicated the fractionation of the enzymic digests, but in some cases sequences giving useful overlaps were obtained. All but one of the cleavage anomalies reported here occurred in the *N*-terminal 70 residues, which also contains a cleavage site for the highly specific enzyme thrombin. It is perhaps significant that five enzymes used in this study (trypsin, thermolysin, chymotrypsin, thrombin and Endoproteinase Lys-C) all had cleavage points between Arg-64 and Tyr-68. Although the conformation of this segment of component 7c is not known, apparently it renders the segment particularly susceptible to enzyme cleavage. Other examples of anomalous cleavages have been reported for the *S. aureus* V8 proteinase after serine, histidine, isoleucine and valine (Berchtold *et al.*, 1982)

and for Endoproteinase Lys-C after arginine and phenylalanine (Aggarwal *et al.*, 1985).

Peptide fractionations at all stages were further complicated by low levels of chemical modification of various residues. Component 7c was isolated from the wool intermediate-filament protein complex by extended chromatography in buffers containing 8 M-urea, conditions that favour carbamylation of protein amino groups. Some peptide separations were also carried out in such buffers. Precautions taken to minimize carbamylation included the incorporation of ethanolamine (50 mM) where possible, as a cyanate scavenger, in urea-containing buffers. Despite such precautions, homocitrulline was occasionally found in peptide fractions, and such artifactual microheterogeneity gave rise to spurious peaks during peptide isolation. Modification of methionine residues causing resistance to cleavage by CNBr also contributed extra peaks to chromatograms of enzyme digests. Peptides were isolated that contained internal homoserine, possibly arising from carboxymethylmethionine (Elleman, 1977); such a peptide, T17, provided the overlap at Met-266. Acid-catalysed hydrolysis of the bond Asp-99-Pro-100 was used to prepare the blocked *N*-terminal segment (H1) of component 7c. Partial hydrolysis of this peptide bond also occurred during CNBr cleavage of component 7c, and the fragments of peptide CN1 generated in this way chromatographed with peptides CN2 and CN3; peptides originating from peptide CN1 were thus often isolated from digests of apparently pure peptides CN2 or CN3.

As shown in Fig. 1, component 7c, isolated by the procedure of Crewther *et al.* (1976) and used in this study, shows considerable microheterogeneity. Substitutions were found at residues 74 (Gly, CmCys and Ser), 80 (CmCys and Ser), 144 (Phe and Tyr), 232 (Val and

Table 1. Amino acid composition of component 7c and the peptides derived by cleavage with CNBr

Values in parentheses were calculated from the sequence.

Amino acid	Amino acid composition (mol of residue/mol)					
	Component 7c	Peptide CN1	Peptide CN2	Peptide CN3	Peptide CN4	Peptide CN5
CmCys	28.8 (32)	15.4 (16)	3.7 (4)	10.5 (11)	0.7 (1)	
Asp	41.9 (40)	21.7 (20)	10.5 (10)	6.0 (6)		3.9 (4)
Thr	20.4 (20)	10.2 (10)	5.1 (5)	4.9 (5)		
Ser	39.8 (40)	20.7 (22)	4.9 (5)	11.5 (12)		1.3 (1)
Glu	74.9 (73)	40.3 (39)	24.3 (24)	7.2 (7)	2.4 (3)	
Pro	12.9 (12)	7.5 (9)		3.3 (3)		
Gly	41.0 (41)	21.7 (22)	3.4 (3)	16.5 (16)		
Ala	40.6 (40)	18.6 (18)	15.5 (16)	5.3 (5)	1.0 (1)	
Val	34.3 (35)	18.2 (18)	6.9 (6)	9.5 (10)	0.9 (1)	
Met or Hse	3.3 (4)	0.7 (1)	1.0 (1)		Trace (1)	0.5 (1)
Ile	20.1 (21)	10.7 (12)	6.3 (7)	1.9 (2)		
Leu	45.1 (44)	24.9 (23)	8.2 (8)	10.3 (10)	1.8 (2)	1.0 (1)
Tyr	12.9 (12)	8.1 (8)	2.0 (2)	1.0 (1)	0.9 (1)	
Phe	13.0 (14)	12.3 (14)				
Lys	23.4 (24)	13.1 (13)	8.8 (9)	1.0 (1)	0.9 (1)	
His	2.0 (2)	0.8 (1)	1.0 (1)			
Arg	35.7 (36)	21.1 (20)	9.2 (9)	6.0 (6)		1.0 (1)
Trp	0.7 (1)		0.4 (1)			
Total	(491)	(266)	(111)	(95)	(11)	(8)

Ser), 276 (Asp, Asn and CmCys) and 284 (Gln and His). Substitutions at residues 276 and 284 may explain the observation that peptide CN2 appears as a number of equally spaced bands on PAGE in the presence of 8 M-urea though behaving as a single species in the presence of SDS (Sparrow & Inglis, 1980). In principle, the substitutions in peptide CN1 at residues 74 and 80 could lead to similar electrophoretic heterogeneity; that this was not observed may be due to the higher M_r of peptide CN1, but could be explained by the combination of substitutions at the two positions being charge-compensating.

The preparation of component 7c used in this study gives a single broad band on PAGE in the presence of SDS or 8 M-urea and on IEF in the presence of 8 M-urea. However, component 7c, as defined by PAGE of the total wool intermediate-filament protein complex, appears on IEF as a smear of pI 4.8–5.0, of which only the lower-pI part corresponds with the preparation used in this study (Marshall & Blagrove, 1979). Thus component 7c is even more microheterogeneous than shown in the present study.

Such microheterogeneity could arise either through polymorphism or through the existence of multiple alleles. For some epidermal keratins multiple alleles have been demonstrated (Rosenberg *et al.*, 1988; Tyner *et al.*, 1985; Trevor & Oshima, 1985), whereas for others (e.g. Krieg *et al.*, 1985) and for the type III intermediate filaments desmin and vimentin (Quax *et al.*, 1985) the genes have been shown to be unique. The genes encoding some of the matrix proteins of wool have also been shown to be unique (Rogers, 1985). For component 7c polymorphism cannot be excluded since the source of the protein used in this study was not wool from a single sheep.

Fig. 1 shows only those peptides needed to establish the amino acid sequence of component 7c unequivocally. In fact, many more peptides were isolated and characterized as a consequence of the large number of enzyme digests needed to complete the sequence determination.

The only part of the sequence not established is the *N*-terminal segment, where the order of the first two residues has not been determined, nor has the nature of the *N*-terminal blocking group (X in Fig. 1). Since the *N*-terminally blocked chymotryptic tripeptide X(Gly,CmCys)-Phe was not isolated, the sequence in this region is deduced from (a) the specificity of chymotrypsin, (b) the amino acid composition of the blocked thermolytic

pentapeptide ending at Thr-5, (c) the amino acid composition of the blocked tryptic peptide ending at Arg-13 and (d) the sequences of the chymotryptic peptide Ser-4-Phe-10 and the thermolytic peptide Val-6-Ala-14.

All other overlaps were of at least three residues with the exception of those covering the bonds Arg-167-Glu-168, Ser-487-Leu-488 and several in the region Tyr-35-Gly-43. Peptides in the region Tyr-35-Gly-43 were ordered by analogy with the corresponding section of the homologous protein from the wool intermediate-filament protein complex, component 5. This section includes the sequence Arg-CmCys-CmCys-Ile-Ser-Ala-Ala-Pro-Tyr-Arg-Gly-Val-Ser-CmCys-Tyr-Arg-Gly-Leu-Thr-Gly-Phe-Gly, which was established with a single peptide isolated from a digest of the analogous acid-cleavage peptide H1 of component 5 with the *A. lyticus* proteinase (L. G. Sparrow, unpublished work). The overlap at Arg-167-Glu-168 had previously been established through the homology of this region of component 7c with a helical fragment excised from the wool intermediate-filament protein complex by limited proteolysis (Sparrow & Inglis, 1980). No peptide including the bond Ser-487-Leu-488 was obtained, and this overlap was deduced from DNA sequence information, which also confirmed the amino acid sequence from Glu-305 to the C-terminus (K. Ward, personal communication).

Table 1 shows the amino acid composition determined for component 7c and its CNBr-cleavage peptides compared with values calculated from the sequence shown in Fig. 1. Given the size and microheterogeneity of com-

Fig. 2. Schematic representation of the conformation of component 7c

The scale is in amino acid residues. For an explanation of the symbols see the text. The position of the discontinuity in the heptad repeat is shown by the dashed lines across segment 2B.

Table 2. Comparison of the amino acid sequence of component 7c with consensus sequences for type I and type II intermediate-filament proteins

Each score records the number of identities found on comparison of the sequence within each segment of component 7c with that from the corresponding segment of the consensus sequences for epidermal type I and type II intermediate-filament proteins shown in Conway & Parry (1988). The numbers at the head of each column show the position of the segment in the component 7c sequence.

	Amino acid sequence comparison								
	H1 79–109	1A 110–144	L1 145–154	1B 155–255	L12 256–272	2A 273–291	L2 292–299	2B 300–420	H2 421–434
Type I	1/31	15/35	2/10	24/101	5/17	5/19	4/8	44/121	–
Type II	19/31	25/35	3/10	50/101	11/17	11/19	6/8	75/121	10/14

ponent 7c, the agreement between calculated and observed values is satisfactory for all residues. The M_r calculated from this sequence is 55 568 (not including the blocking group), in very good agreement with the values determined by low-speed sedimentation equilibrium (56 000) and SDS/PAGE (57 000) (Woods, 1979).

Analysis of the sequence

Analysis of the sequence of component 7c suggests that it fits readily into the pattern characteristic of intermediate filaments in having a central rod-like helical domain that is divided by three non-helical linking segments, L1, L12 and L2, into four helical tracts, 1A, 1B, 2A and 2B. Immediately adjacent to this central domain, on both the *N*- and the *C*-terminal sides are sections, H1 and H2 respectively, of the non-helical regions that are highly conserved in a number of intermediate-filament classes (Steinert & Parry, 1985). The structure of component 7c is shown schematically in Fig. 2.

Component 7c has been recognized as belonging to the type II class of intermediate-filament proteins on the basis of the similarity of the sequence of a 109-residue segment of peptide CN1 to that of a type II helical segment isolated after limited proteolysis of the wool intermediate-filament protein complex (Sparrow & Inglis, 1980). The similarity of the sequence of component 7c to those of other type II intermediate-filament proteins extends through the entire central rod-like region and beyond, as shown in Table 2, where the sequence of component 7c is compared with consensus sequences for type I and type II intermediate-filament proteins (Conway & Parry, 1988). In this comparison, only the H regions of the *N*- and *C*-terminal non-helical segments were included, as sequence similarity over the remainder of these segments was not significant.

REFERENCES

- Aggarwal, B. B., Henzel, W. J., Moffat, B., Kohr, W. J. & Harkins, R. N. (1985) *J. Biol. Chem.* **260**, 2334–2344
- Berchtold, M. W., Heizmann, C. W. & Wilson, K. J. (1982) *Eur. J. Biochem.* **127**, 381–389
- Conway, J. F. & Parry, D. A. D. (1988) *Int. J. Biol. Macromol.* **10**, 79–98
- Crewther, W. G. (1976) *Proc. Int. Wool Text. Res. Conf.* 5th, 1975, vol. 1, pp. 1–101
- Crewther, W. G., Dowling, L. M., Gough, K. H., Inglis, A. S., McKern, N. M., Sparrow, L. G. & Woods, E. F. (1976) *Proc. Int. Wool Text. Res. Conf.* 5th, 1975, vol. 2, pp. 233–242
- Crewther, W. G., Dowling, L. M., Gough, K. H., Marshall, R. C. & Sparrow, L. G. (1980) in *Fibrous Proteins: Scientific, Industrial and Medical Aspects* (Parry, D. A. D. & Creamer, L. K., eds.), vol. 2, pp. 151–159, Academic Press, London and New York
- Dopheide, T. A. A. (1973) *Eur. J. Biochem.* **34**, 120–124
- Dowling, L. M., Gough, K. H., Inglis, A. S. & Sparrow, L. G. (1979) *Aust. J. Biol. Sci.* **32**, 437–442
- Dowling, L. M., Parry, D. A. D. & Sparrow, L. G. (1983) *Biosci. Rep.* **3**, 73–78
- Dowling, L. M., Crewther, W. G. & Inglis, A. S. (1986) *Biochem. J.* **236**, 695–703
- Elleman, T. C. (1977) *Aust. J. Biol. Sci.* **30**, 33–45
- Fraser, R. D. B., MacRae, T. P. & Rogers, G. E. (1972) in *Keratins: Their Composition, Structure and Biosynthesis* (Kugelmass, I. N., ed.), Charles C. Thomas, Springfield
- Holcomb, G. N., James, S. A. & Ward, D. N. (1968) *Biochemistry* **7**, 1291–1296
- Inglis, A. S., Rubira, M. R. & Strike, P. McK. (1981) *J. Biochem. Biophys. Methods* **4**, 279–285
- Jones, L. N. (1976) *Biochim. Biophys. Acta* **446**, 515–524
- Krieg, T. M., Schafer, M. P., Cheng, C. K., Filpula, D., Flaherty, P., Steinert, P. M. & Roop, D. R. (1985) *J. Biol. Chem.* **260**, 5867–5870
- Marshall, R. C. & Blagrove, R. J. (1979) *J. Chromatogr.* **172**, 351–356
- Marshall, R. C., Gillespie, J. M., Inglis, A. S. & Frenkel, M. J. (1980) *Proc. Int. Wool Text. Res. Conf.* 6th 1980, vol. 2, pp. 147–158
- McKern, N. M., O'Donnell, I. J., Stewart, D. J. & Clark, B. L. (1985) *J. Gen. Microbiol.* **131**, 1–6
- Quax, W., Khan, P. M., Quax-Jeuken, Y. & Bloemendal, H. (1985) *Gene* **38**, 189–196
- Rogers, G. E. (1985) *Ann. N.Y. Acad. Sci.* **455**, 403–425
- Rosenberg, M., Ray Chaudhury, A., Shows, T. B., LeBeau, M. M. & Fuchs, E. (1988) *Mol. Cell. Biol.* **8**, 722–736
- Sparrow, L. G. & Inglis, A. S. (1980) *Proc. Int. Wool Text. Res. Conf.* 6th 1980, vol. 2, pp. 237–246
- Steinert, P. M. & Parry, D. A. D. (1985) *Annu. Rev. Cell Biol.* **1**, 41–65
- Swart, L. S., Joubert, F. J. & Parris, D. (1976) *Proc. Int. Wool Text. Res. Conf.* 5th, 1975, vol. 2, pp. 254–264
- Trevor, K. & Oshima, R. G. (1985) *J. Biol. Chem.* **260**, 15885–15891
- Tyner, A. L., Eichman, M. J. & Fuchs, E. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 4683–4687
- Woods, E. F. (1979) *Aust. J. Biol. Sci.* **32**, 423–435

Received 19 January 1989/7 March 1989; accepted 16 March 1989