Amino Acid Sequences of α -Helical Segments from S-Carboxymethylkeratine-A

COMPLETE SEQUENCE OF A TYPE-II SEGMENT

W. GORDON CREWETHER, ADAM S. INGLIS and NEIL M. MCKERN
Division of Protein Chemistry, CSIRO, Parkville (Melbourne), Vic. 3052, Australia

(Received 2 November 1977)

1. The helical fragments obtained by partial chymotryptic digestion of S-carboxymethylkeratine-A, the low-sulphur fraction from wool, were fractionated into type-I and type-II helical segments in aqueous urea under conditions limiting carbamylation. 2. The amino acid sequence of a 109-residue type-II segment was completed by using the sequenator. 3. When the data were incorporated into a helical model of 3.6 residues per turn the hydrophobic residues generated a band aligned at a slight angle to the helical axis. This result is in accord with the postulated coiled-coil structure of the crystalline regions of α -keratin.

In the preceding paper, Hogg *et al.* (1978) proposed a partial tentative sequence for a helical segment from the low-sulphur fraction S-carboxymethylkeratine-A from wool. Similar segments (type II) constitute a major part of the α -helical fragments precipitated at pH4.0 from partial chymotryptic digests of S-carboxymethylkeratine-A (Crewther & Dowling, 1971). During storage of the material used for earlier studies (Hogg *et al.*, 1978) the homocitrulline content increased, with a concomitant decrease in lysine content. It was therefore presumed that the blocked N-terminus observed for this fraction was due to carbamylation of the N-terminal amino group by cyanate present in the urea buffers used in disintegrating the fragments and separating their constituent chain segments (Stark *et al.*, 1960).

In the present study an identical type-II helical segment was isolated under conditions that retained a free N-terminal amino group. The automatic sequenator was used to determine the N-terminal sequence of the intact chain, unknown sections of the sequences of tryptic peptides, the N-terminal sequences of large peptides obtained by acid cleavage at aspartic acid residues and the location of amide side chains. Together with the analyses given by Hogg *et al.* (1978), these data establish the complete amino acid sequence.

Experimental

Preparation of helix-enriched segments from S-carboxymethylkeratine-A

The preparation of the S-carboxymethylated low-sulphur protein fraction, S-carboxymethylkeratine-A, from Merino wool was as described by

Dowling & Crewther (1974). Chymotryptic digestion of this fraction and the recovery of helix-enriched segments from the digest were carried out as described by Crewther & Dowling (1971), except that 0.05 M-ethanolamine was included in urea buffers to act as a competitor for cyanate.

In addition, a higher pH (9.5) was used during gel filtration, and a lower temperature (2°C) was used for ion-exchange chromatography. Both of these procedures minimize the likelihood of carbamylation of amino groups (Dirnhuber & Schütz, 1948).

Isolation of tryptic peptides

To assign amide side chains and to complete the sequence data on tryptic peptides, 200 mg of the type-II helical segment was digested with trypsin and the peptides were fractionated by chromatography on DEAE-cellulose with an $(\text{NH}_4)_2\text{CO}_3$ gradient at pH8.5 (Hogg *et al.*, 1978). The pooled fractions were freeze-dried and peptides were isolated by high-voltage electrophoresis on Whatman 3MM paper at pH6.5 and pH9.0 (Hogg *et al.*, 1978).

Preparation of peptides by cleavage with dilute acid

Since aspartic acid residues were interspersed between the basic residues in the tentative sequence, dilute-HCl cleavage (Schultz, 1967) was used to generate large peptides overlapping the tryptic peptides. The type-II segment (80 mg) was heated at 110°C for 1 h in 12.5 mM-HCl, and the products were freeze-dried and fractionated on Sephadex G-50 (superfine grade) in 0.05 M- $\text{Na}_2\text{B}_4\text{O}_7$ containing 8 M-urea. Three columns (1.6 cm $\times$ 90 cm) in series were used with a flow rate of 4 ml/h.

Isolation of chymotryptic peptides

To confirm one portion of the sequence a chymotryptic digest of segment ChC'6 prepared as described by Hogg *et al.* (1978) was subjected to high-voltage paper electrophoresis for a short period (3000V; 20min) and the most-cationic peptides comprising four adjacent bands were eluted, analysed and their sequences determined manually.

Sequenator procedures

The sequenator used in these studies was of the type described by Edman & Begg (1967) but with a larger and faster spinning cup and a modified 6-way valve. The latter feature was designed to give a thinner protein film and hence more efficient reaction and extraction of the film with reagents entering the cup. The procedures used for the automatic analysis of the *N*-terminal sequences of the intact type-II chain and the acid-cleaved peptides were as described by Edman & Begg (1967) and Inglis & Burley (1977).

Sequences of tryptic peptides were determined automatically by the partition procedure described by Crewther & Inglis (1975) with 300–500nmol of peptide. All amide residues were identified unequivocally by t.l.c. of the amino acid phenylthiohydantoin (Inglis & Nicholls, 1973).

Polyacrylamide-gel electrophoresis

Fractions derived from the helix-enriched fragments were monitored by electrophoresis in 10% (w/v) polyacrylamide gels containing 8M-urea, as described by Sparrow & Crewther (1972).

Amino acid analysis

Amino acid analyses of the complete chains were obtained by hydrolysing 2–3mg with 1 ml of 6M-HCl

in vacuo at 110°C for 24h and analysing the hydrolysate on a Beckman automatic amino acid analyser, model 120C. Peptides (approx. 50nmol) were hydrolysed and analysed in a similar manner.

Results

Gel filtration in buffered 8M-urea of 10g of the helix-rich fragments from *S*-carboxymethylkeratine-A on a Sephadex G-75 column gave a similar elution pattern (Fig. 1) to that obtained by Crewther & Dowling (1971). Fraction ChC' (Fig. 1), mol.wt. approx. 12500, was chromatographed on a DEAE-cellulose column to yield fractions that were pooled as shown (Fig. 2). Amino acid analyses of samples of these fractions (Table 1) and high-voltage paper electrophoresis of tryptic digests showed that the first four fractions contained predominantly type-I

Fig. 1. Elution profile (280nm) of helical fragments from partial chymotryptic proteolysis of *S*-carboxymethylkeratine-A on a Sephadex G-75 column (15 cm × 120 cm). Flow rate was 300 ml/h, and eluent was 8M-urea/0.01 M-Tris/0.05 M-ethanolamine/1 mM-EDTA, pH adjusted to 9.5 with 6M-HCl.

Fig. 2. Elution profile (280nm) of fraction ChC' at 2°C, on a column of DEAE-cellulose (10 cm × 120 cm) with an elution flow rate of 400 ml/h

The buffer system 7M-urea/0.02 M-ethanolamine/0.01 M-Tris/1 mM-EDTA adjusted to pH 7.4 with 6M-HCl was used in conjunction with a linear KCl gradient (0.03–0.18 M; ----) in a total volume of 150 litres. For details of pooled fractions 1–8 see the text.

Table 1. Comparison of amino acid compositions of fractions 1–8 obtained by ion-exchange chromatography of fraction ChC' on DEAE-cellulose (Fig. 2) with values for type-I and type-II helical segments previously isolated

Amino acid	Content of amino acid residue in fraction (residues/100 residues)									
	Type I*	1	2	3	4	5	6†	7	8	Type II*
Ala	6.4	10.3	8.0	6.9	6.8	7.5	8.7	8.5	8.3	8.3
Arg	6.9	6.5	6.8	6.6	6.6	7.3	8.6	8.2	8.1	8.6
Asp	11.9	10.8	12.0	12.1	11.6	10.6	9.6	9.6	9.6	9.7
CmCys	3.6	3.0	3.2	3.2	3.6	3.7	3.6	3.3	4.0	3.8
Glu	20.3	21.0	20.3	19.9	19.8	19.9	21.2	22.4	22.1	21.9
Gly	1.9	2.4	1.7	1.9	2.5	3.1	3.2	3.2	3.3	2.9
His	0.9	0.4	0.8	0.9	0.9	0.6	0.3	0.7	0.7	0.8
Ile	5.4	4.6	5.1	5.3	5.1	4.5	3.5	3.6	3.6	3.6
Leu	15.8	14.8	14.4	15.7	15.3	14.5	13.6	12.7	12.5	12.9
Lys	6.9	7.9	7.8	7.2	6.5	5.8	5.1	5.2	5.1	5.2
Met	0.1	0.7	0.2	0.1	0.1	0.2	0.0	0.0	0.0	—
Phe	2.0	1.4	1.8	1.8	1.7	1.9	2.7	3.0	2.6	2.9
Pro	0.9	0.7	1.1	0.9	1.1	1.1	1.1	1.1	1.3	0.9
Ser†	6.1	4.5	6.0	6.1	6.0	6.3	5.8	5.5	5.6	5.5
Thr†	3.8	3.8	3.6	3.8	4.1	3.7	2.9	2.5	2.4	2.4
Tyr	2.5	2.1	2.3	2.4	2.6	3.0	4.3	4.1	4.0	4.5
Val	4.7	5.1	4.8	5.4	5.7	6.1	5.5	6.6	6.8	6.3

* Mean values for amino acid compositions of type-I and type-II helical segments isolated by Crewther & Dowling (1971).

† Uncorrected values.

‡ 5 mM-Phenol present during hydrolysis.

segments, whereas fractions 6–8 consisted of type-II segments (cf. Crewther & Dowling, 1971).

Polyacrylamide-gel electrophoresis in 8M-urea of the fractions indicated in Fig. 2 showed that all fractions were heterogeneous, although a densitometer tracing of the pattern obtained from fraction ChC'6 indicated that the major band constituted greater than 80% of the stained bands.

N-Terminal sequence of the type-II segment

The *N*-terminal sequence of segment ChC'6 (Fig. 3) was determined in the automatic sequenator with a loading of 10mg. The yield of the thiazolinone derivative of the *N*-terminal glutamine residue, as determined by A_{269} in ethanol, was 55% of the theoretical value. The low yield may be due to partial cyclization of the glutamine to a pyrrolidonecarboxylic acid residue (Doolittle, 1972). The identification of the first 53 residues confirmed the tentative partial sequence given by Hogg *et al.* (1978) and completed this portion of the sequence, including the location of amide side chains. Residue 41 (serine) was not determined unequivocally from the automatic sequence data and was confirmed by sequencing the tryptic peptide (residues 35–52) manually (K. Gough, personal communication).

Sequences of isolated peptides

Variations in the absorbance of the eluate from the ion-exchange chromatography of a tryptic digest of segment ChC'6 are shown in Fig. 4 and the pooled

fractions are indicated. By high-voltage electrophoresis of these fractions on paper, all the major tryptic peptides obtained in the previous study were isolated (Hogg *et al.*, 1978). Table 2 lists the amino acid compositions of selected peptides together with their amino acid sequences as determined with the sequenator and the designation of the corresponding peptide from segment ChC'9 (Hogg *et al.*, 1978). As with the *N*-terminal section of the molecule the sequences of the major peptides from fraction ChC'6 were identical with those from fraction ChC'9 (Hogg *et al.*, 1978). Further, it was possible to complete the sequence of the peptide corresponding to peptides T16b and T18b in the previous study, so making possible a statement of the complete sequence of this type-II helical segment (Fig. 3).

Isolation and sequence analyses of acid-cleaved peptides

The products obtained by cleavage of segment ChC'6 with dilute HCl were fractionated by gel filtration (Fig. 5), giving rise to fractions 1 (*N*-terminal serine), 2 (*N*-terminal carboxymethylcysteine) and 3 (*N*-terminal leucine). Automatic sequence analyses of these fractions showed that they were identical with the sections of the proposed total sequence indicated in Fig. 3. The sequences around five tryptic cleavage points were confirmed in this way.

Sequences of cationic chymotryptic peptides

The chymotryptic peptides of greatest basicity had the sequences Lys-Lys-Arg-Tyr, Lys-Arg-Arg-

Fig. 4. Elution pattern (230 nm) from DEAE-cellulose at pH 8.7 of tryptic peptides derived from fraction ChC'6
 ----, $(\text{NH}_4)_2\text{CO}_3$ gradient.

Table 2. Sequence analyses of major tryptic peptides from fraction ChC'6

Peptide*	Amino acid analysis	Amino acid sequence
3D (T6a)	Ala ₃ , Asx, Glx ₂ , Leu, Lys ₂ , Phe, Thr, Val	Ala-Thr-Ala-Glu-Asn-Glu-Phe-Val-Ala-Leu-Lys-Lys
5E2 (T15c)	Ala, Arg ₂ , Glx ₃ , Ile _{0.5} , Leu, Tyr, Val _{0.5}	Arg-Tyr-Glu-Glu-Glu-Val/Ile-Ala-Leu-Arg
8C2 (T16b)	Ala ₂ , Arg, Asx ₃ , Glx, Ile, Leu ₃ , Lys, Phe, Ser, Thr, Val	Lys-Ser-Asp-Leu-Glu-Ala-Asn-Val-Glu-Ala-Leu-Ile-Gln-Glu-Thr-Asp-Phe-Leu-Arg
8D2 (T19e)	Ala, Arg, Asx ₂ , CmCys, Leu, Lys, Tyr, Val	Asp-Val-Asp-CmCys-Ala-Tyr-Leu-Arg-Lys
8E1 (part T18c)	Arg, Glx ₃ , Ile, Leu, Tyr	Leu-Tyr-Glu-Glu-Glu-Ile-Arg
10C (T21C)	Ala ₃ , Arg, Asx, CmCys, Glx ₃ , Gly, Ser	Glu-Ala-Glu-CmCys-Ala-Glu-Ala-Asp-Ser-Gly-Arg

* The corresponding peptide in fraction ChC 9 is shown in parentheses. Numerals in peptide designation indicate the fraction (Fig. 4) from which the peptide was isolated.

Tyr and Leu-Arg-Arg-Leu, thus confirming sequences 50-53 and 98-101.

Discussion

The relative simplicity of the elution pattern of fraction ChC' from DEAE-cellulose in this study as compared with fraction ChC in the previous study (Crewther & Dowling, 1971) indicates that the use of low temperatures and the inclusion of ethanolamine in aqueous buffers containing urea resulted in decreased chemical modification of the chains during fractionation. This was confirmed by a comparison of the amino acid analyses of the two fractions, the fraction, ChC', used in this study containing no homocitrulline. The difference is therefore attributable largely to prevention of carbamoylation of amino groups.

The amino acid analyses (Table 1) and the tryptic peptide 'maps' of the fractions obtained by ion-exchange chromatography of fraction ChC' show that subfractions ChC' 2-4 contain typical type-I helical segments and subfractions ChC' 6-8 contain type-II helical segments. Polyacrylamide-gel-electrophoresis patterns of these fractions indicate that at least three major bands correspond to type-I segments and three to type-II segments. The major polypeptide present in fraction ChC'6 appears to be identical with that in fraction ChC 9, the fraction used in the previous study (Hogg *et al.*, 1978).

The amino acid composition of the complete sequence outlined in Fig. 3 corresponds satisfactorily to the amino acid analysis of fraction ChC'6 (Table 1), apart from significant discrepancies in the contents of valine and basic amino acids. As these residues are known (Hogg *et al.*, 1978) to be involved in sub-

Fig. 5. Elution profile (280nm) of dilute-HCl-cleavage products of fraction ChC'6 passed over three serially connected columns (1.6cm×90cm) of Sephadex G-50 (superfine grade)

Flow rate was 4ml/h, elution with 8M-urea/0.05M-borate.

stitutions at positions 28, 52, 57, 71, 72, 107 and 108 the discrepancies are not unexpected. Thus in terms of residues per 109 residues the sum of lysine and arginine contents (15.0) from the analysis exactly equals the number of basic residues in the sequence, and the sum (19.5) of valine, alanine and isoleucine residues, which substitute one for the other (Hogg *et al.*, 1978), compares satisfactorily with the 19 positions occupied by these residues in the sequence.

Fig. 6 shows that, when consecutive heptapeptides in the amino acid sequence of the helical segment are superimposed, residues with hydrocarbon side chains occur chiefly at two sites in the heptapeptide sequence, three or four residues apart. Accordingly, when an α -helical model of the sequence is formed with 3.6 residues per turn these hydrocarbon side chains generate a band at a slight angle to the axis of the helix. Crick (1953) predicted that such a sequence of amino acids would readily form a coiled-coil conformation, which is in accord with X-ray-diffraction measurements on α -keratin (Fraser *et al.*, 1972) and on the helical fragments themselves (Suzuki *et al.*, 1973).

An α -helical model of the sequence also shows that many ionic interactions are possible between oppositely charged side chains on adjacent turns of the helix, i.e.: Arg-22 and -23 with Glu-19 and -26;

a	b	c	d	e	f	g
				Gln	Asn	Arg
Gln	Cys	Cys	Glu	Ser	Asn	Leu
Glu	Pro	Leu	Phe	Ser	Gly	Tyr
Ile	Glu	Thr	Leu	Arg	Arg	Glu
Ala	Glu	Cys	Ala	Glu	Ala	Asp
Ser	Gly	Arg	Leu	Ser	Ser	Glu
Leu	Asn	Ser	Leu	Gln	Glu	Val
Leu	Glu	Gly	Tyr	Lys	Lys/Arg	Lys/Arg
Tyr	Glu	Glu	Glu	Ile/Val	Ala	Leu
Arg	Ala	Thr	Ala	Glu	Asn	Glu
Phe	Val	Ala	Leu	Lys	Lys	Asp
Val	Asp	Cys	Ala	Tyr	Leu	Arg
Lys	Ser	Asp	Leu	Glu	Ala	Asn
Val	Glu	Ala	Leu	Ile	Gln	Glu
Thr	Asp	Phe	Leu	Arg	Arg	Leu
Tyr	Glu	Glu	Glu	Ile	Arg	Val
Leu						

Fig. 6. Amino acid sequence of the type-II helical segment arranged as repeating heptads by the convention of Stewart & McLachlan (1975)

Residues with hydrocarbon side chains occur chiefly within boxed positions.

Arg-34 with Asp-31, Glu-38; Lys-51, Arg-52 with Glu-47, -54 and -55; Arg-60 with Glu-56, -64; Lys-71, -72 with Asp-75; Arg-80 with Asp-83; Arg-99, -100 with Asp-96, Glu-103, -104.

These and interhelical ionic interactions probably account to a considerable degree for the observation by Speakman & Hirst (1933) that 'salt linkages' play a major part in stabilizing α -keratin.

We acknowledge with gratitude the manual-sequence data provided by Mr. K. H. Gough for several peptides.

References

- Crewther, W. G. & Dowling, L. M. (1971) *Appl. Polym. Symp. no. 18*, 1-20
- Crewther, W. G. & Inglis, A. S. (1975) *Anal. Biochem.* **68**, 572-585
- Crick, F. H. C. (1953) *Acta Crystallogr.* **6**, 685-688
- Dirnhuber, P. & Schütz, F. (1948) *Biochem. J.* **42**, 628-632
- Doolittle, R. F. (1972) *Methods Enzymol.* **25**, 231-244
- Dowling, L. M. & Crewther, W. G. (1974) *Prep. Biochem.* **4**, 203-226
- Edman, P. & Begg, G. (1967) *Eur. J. Biochem.* **1**, 80-91
- Fraser, R. D. B., MacRae, T. P. & Rogers, G. E. (1972) *Keratins*, C. C. Thomas, Springfield
- Hogg, D. McC., Crewther, W. G. & Dowling, L. M. (1978) *Biochem. J.* **173**, 353-363
- Inglis, A. S. & Burley, R. W. (1977) *FEBS Lett.* **73**, 33-37
- Inglis, A. S. & Nicholls, P. N. (1973) *Biochem. J.* **42**, 628-632
- Schultz, J. (1967) *Methods Enzymol.* **11**, 255-263
- Sparrow, L. G. & Crewther, W. G. (1972) *J. Text. Inst.* **63**, 572-585
- Speakman, J. B. & Hirst, M. C. (1933) *Trans. Faraday Soc.* **29**, 148-172
- Stark, G. R., Stein, W. H. & Moore, S. (1960) *J. Biol. Chem.* **235**, 3177-3181
- Stewart, M. & McLachlan, A. D. (1975) *Nature (London)* **257**, 331-333
- Suzuki, E., Crewther, W. G., Fraser, R. D. B., MacRae, T. P. & McKern, N. M. (1973) *J. Mol. Biol.* **73**, 275-278