Amino Acid Sequences of Peptides from a Tryptic Digest of a Urea-Soluble Protein Fraction (U.S.3) from Oxidized Wool

By M. C. CORFIELD, J. C. FLETCHER AND A. ROBSON*
Wool Industries Research Association, Torridon, Headingley Lane, Leeds 6

(Received 13 June 1966)

1. A tryptic digest of the protein fraction U.S.3 from oxidized wool has been separated into 32 peptide fractions by cation-exchange resin chromatography. 2. Most of these fractions have been resolved into their component peptides by a combination of the techniques of cation-exchange resin chromatography, paper chromatography and paper electrophoresis. 3. The amino acid compositions of 58 of the peptides in the digest present in the largest amounts have been determined. 4. The amino acid sequences of 38 of these have been completely elucidated and those of six others partially derived. 5. These findings indicate that the parent protein in wool from which the protein fraction U.S.3 is derived has a minimum molecular weight of 74000. 6. The structures of wool proteins are discussed in the light of the peptide sequences determined, and, in particular, of those sequences in fraction U.S.3 that could not be elucidated.

Although a mass of amino acid sequence data has been accumulated for globular proteins, corresponding information for fibrous proteins is fragmentary. It is not known how far the special structural and protective functions of fibrous proteins govern the amino acid sequences that occur in them, nor whether the mechanism of protein synthesis accepted for globular proteins is modified to produce the elaborate tertiary structures of fibrous proteins (Priestley & Speakman, 1966). Despite the impetus that early work on wool (Consdon, Gordon & Martin, 1947) gave to amino acid sequence studies on proteins, progress with wool has been hampered by the difficulty of obtaining demonstrably homogeneous protein fractions. Much effort has been devoted to this problem (reviewed by Crewther, Fraser, Lennox & Lindley, 1965), but the number and nature of the proteins present in the native wool fibre, and the sites of their synthesis, are still uncertain. The detailed chemical structure of wool is not only of technological importance. The follicle is a convenient small-scale structure for the study of the processes of growth, differentiation and keratinization, and as such is the subject of active biochemical and morphological investigations.

This paper describes work on the tryptic digest of a urea-soluble fraction (U.S.3) isolated in 36% yield from oxidized wool under comparatively mild conditions (Corfield, 1963). This work was undertaken not only to obtain data on amino acid

sequences, but also with the anticipation that a knowledge of the number of peptides present in the digest, and their amounts, would provide information about the size and complexity of the starting material. Preliminary reports of the work have been given (Cole, Corfield, Fletcher & Robson, 1966; Corfield, Fletcher, Myers & Robson, 1966).

MATERIALS AND METHODS

Materials

Enzymes. Crystalline trypsin was obtained from Armour Pharmaceutical Co. Ltd. (Eastbourne, Sussex), crystalline pepsin and chymotrypsin (3 × crystallized, salt-free) were from Koch-Light Laboratories Ltd. (Colnbrook, Bucks.), Pronase was from Calbiochem (Los Angeles, Calif., U.S.A.) and subtilisin and carboxypeptidase A (treated with diisopropyl phosphorofluoridate) were from Sigma Chemical Co. (St Louis, Mo., U.S.A.).

Chemicals. Phenyl isothiocyanate (Koch-Light Laboratories Ltd.) was redistilled and stored at 5°. DNS† chloride was supplied by British Drug Houses Ltd. (Poole, Dorset). [¹³¹I]Pipsyl chloride was prepared as described by Keston, Udenfriend & Cannan (1949) with an activity of about 20 mc/m-mole.

Preparation of fraction U.S.3

Samples of freshly prepared oxidized wool (50g.) were fractionated by the method of Corfield (1963). Fraction U.S.3 was obtained in 36% yield and was stored at room temperature.

† Abbreviations: DNS, 1-dimethylaminonaphthalene-5-sulphonyl; pipsyl, iodobenzene-*p*-sulphonyl; in amino acid sequences CySO₃H and MetO₂ refer to cysteic acid and methionine sulphone respectively.

* Present address: Department of Textile Industries, University of Leeds.

Fig. 1. Digestion of fraction U.S.3 with trypsin. Fraction U.S.3 (4.65 g.) in water (200 ml.) was digested with trypsin (30.4 mg.) at pH 8.5. The pH was kept constant by the addition of 0.118N-NaOH. The percentage of lysine+arginine peptide bonds broken is calculated from the amino acid composition of fraction U.S.3 and the assumption that the basic groups liberated have an average pK_a 7.5. The reaction was stopped after 90 min. by the addition of formic acid.

Digestion of fraction U.S.3 with trypsin

Fraction U.S.3 (4.6 g.) was dissolved in water (200 ml.) under N_2 at 37° and the solution brought to pH 8.5 by the addition of 0.2N-NaOH from an autotitrator. Trypsin (30 mg.) was added, and the reaction was followed by observing the uptake of alkali (Fig. 1). After 86 min., the reaction was stopped by the addition of formic acid (3.6 ml.), when a white precipitate immediately formed.

Separation of peptides

Initial separation on resin. The acidified reaction mixture from the tryptic digestion, including the precipitated material, was applied to the top of a column (130 cm. $\times$ 7 cm.) of Zeo-Karb 225 (X2) resin (200 mesh) and washed in. Volatile buffers (see Fig. 2) were pumped down the column at a rate of 200 ml./hr. through a mixing vessel of capacity 15 l., and the effluent was collected in 200 ml. fractions. Samples (1 ml.) from each fraction were hydrolysed with 2N-NaOH (1 ml.) at 100° for 1 hr. and the pH was adjusted by the addition of 5.2N-acetic acid (1 ml.). Modified ninhydrin reagent (Moore & Stein, 1954) (1 ml.) was added and the mixture heated at 100° for 15 min. After the addition of ethanol-water (1:1, v/v; 5 ml.) the extinction at 570 m μ was measured. Other samples (1 ml.) from each fraction were examined by chromatography on paper in butan-1-ol-acetic acid-water (4:1:5, by vol.) and by high-voltage electrophoresis on paper in acetic acid-formic acid buffer, pH 1.85, to assist in locating peptides. The fractions were combined as shown in Fig. 2(a) to give crude peptide fractions designated 1T/1-32. In the fractionation of a second, similarly prepared, tryptic digest of fraction U.S.3, the digestion mixture was applied to the column in a more acidic buffer (pH 2.4 instead of pH 8.5) to improve the resolution of the faster-moving peptides. In this case the fractions were combined as shown in Fig. 2(b) to give crude peptide fractions designated 2T/1-42.

Refractionation on resin. Fractions 1T/2-32 were sub-

jected to a second fractionation on columns (150 cm. $\times$ 1.8 cm.) of the same resin used for the initial fractionation. Each peptide fraction was suspended in the chromatographic buffer (5 ml.) and shaken for 2 hr. Insoluble material, designated by the letter R, was removed by centrifuging, washed with water, ethanol and ether, and dried. Soluble material was applied to the column, which was eluted at the rate of 8 ml./hr. with the pyridine-formic acid buffers indicated in Table 1. Peptides were located in the effluent fractions (8 ml.) by reaction with ninhydrin after alkaline hydrolysis as described above, and the fractions were combined as shown. Peptides were recovered in the dry state by freeze-drying. The fractions produced at this stage are designated by Greek letters, e.g. 1T/3 α , β and γ .

Separation by paper chromatography. The peptide fractions thus obtained were fractionated further by paper chromatography. A sample of each fraction (6 mg.) was dissolved in water or aqueous ammonia and applied as a line, 15 cm. long, to a sheet of Whatman 3MM filter paper. The paper was hung for 1 hr. in a tank containing the lower phase of butan-1-ol-acetic acid-water (4:1:5, by vol.), followed by development overnight with the upper phase of this mixture. Some peptides required longer periods of elution, up to 3 days, and in these cases the bottom edge of the paper was serrated to avoid non-uniform flow of the solvent. The paper was dried in a current of warm air and two guide strips, each containing the outer 1 cm. portion of the line on which the peptide mixture was applied, were cut off, treated with cadmium-ninhydrin reagent (Heilman, Barollier & Watzke, 1957) and left overnight in an ammonia-free atmosphere for the colour to develop. Bands corresponding to peptides thus revealed were marked on the unstained piece of paper, numbered serially from the origin, and cut out, and the peptides eluted into small glass vials by the method of Blackburn & Lee (1966a) with aq. 10% (v/v) pyridine. The peptides were stored dry at -40°. The R_F values are indicated in Fig. 3.

Separation by high-voltage paper electrophoresis. When peptides required further purification, a solution of the peptide was applied as a line 10 cm. long about 30 cm. from one end of a sheet of Whatman 3MM filter paper (100 cm. $\times$ 15 cm.) that had previously been wetted with its own weight of pH 1.85 buffer (21 ml. of 98% formic acid and 74.2 ml. of acetic acid/l.). Electrophoresis at 90 v/cm. and 11° for 40 min. was carried out in the apparatus described by Blackburn (1965) with the paper placed so that the point of application was sited at the anode end. After drying the paper, peptides were located on guide strips with ninhydrin and recovered as described above. Peptides thus obtained were designated by the letter H and by a number indicating the order of the observed band measured from the anode end.

Amino acid analysis of peptides

Peptides were hydrolysed in sealed tubes for 20 hr. with 5N-HCl at 105°. Analysis was performed by the automatic method of Corfield & Robson (1962) modified as described by Cole *et al.* (1966). In those cases where a less complete analysis sufficed, the method of Atfield & Morris (1961) as modified by Blackburn & Lee (1963), or the method of Corfield & Simpson (1965), was used.

Fig. 2. Chromatography of tryptic digests (*a*, 1T/1; *b*, 2T/1) of fraction U.S.3 (approx. 5g.) on a column (130cm. $\times$ 7cm.) of Zeo-Karb 225 (X2) in pyridine-formic acid and pyridine-acetic acid buffers. The effluent fractions were combined as shown by the numbered bars above each chromatogram. Samples (1ml.) of each effluent fraction (195ml.) were taken for alkaline hydrolysis and colour development with ninhydrin. Buffers: A, 16.25ml. of pyridine and 51ml. of formic acid/l.; B, 81.25ml. of pyridine and 43.5ml. of formic acid/l.; C, 162.5ml. of pyridine and 99ml. of acetic acid/l.; D, 162.5ml. of pyridine and 4.5ml. of acetic acid/l.

Table 1. *Rechromatography of peptides 1T/2-32*

Peptides were eluted from Zeo-Karb 225 resin (see the text) with the buffers shown and recovered from the fractions indicated.

Peptide fraction	Fraction nos. of subfractions				Pyridine-formic acid buffer	
	α	β	γ	δ	Molarity	pH
1T/2	12-19	20-29	42-53		0.20	2.50
1T/3	42-50	64-72	122-142		0.20	3.00
1T/4	45-54	150-168	169-190		0.20	3.00
1T/5	30-40	41-50	90-105		0.40	3.21
1T/6	36-40	45-57	129-138	139-168	0.40	3.21
1T/7	44-59	108-123	124-135		0.46	3.30
1T/8	48-70	105-119	120-143		0.46	3.30
1T/9	50-77	114-126	127-146	153-163	0.55	3.42
1T/10	100-113	114-132			0.55	3.42
1T/11	57-63	66-74	131-140		0.60	3.56
1T/12	54-64	113-130	131-153		0.60	3.56
1T/13	28-32	80-84			0.70	3.67
1T/14	42-48	81-87			0.70	3.67
1T/15	33-40	66-75	77-81		0.75	3.75
1T/16	53-66	120-130	131-139		0.80	3.78
1T/17	63-69	70-80	120-134	135-150	0.85	3.84
1T/18	62-72	122-127	129-133	134-140	0.90	3.95
1T/19*	56-62	64-69	70-76	77-86	0.95	4.09
1T/20	75-81	82-89	129-139	161-172	1.00	4.23
1T/21	98-108	145-155			1.05	4.38
1T/22	125-137	140-150	151-160	170-182	1.10	4.45
1T/23	81-89	120-133	145-156		1.10	4.45
1T/24†	62-69	70-80	81-90	107-117	1.20	4.60
1T/25	105-125	134-146	172-180		1.27	4.62
1T/26	153-167				1.30	4.70
1T/27	97-108	177-193	202-213		1.35	4.75
1T/28	140-151	152-164			1.40	4.80
1T/29	129-136	137-144			1.40	4.83
1T/30	71-78	141-154			1.40	4.83
1T/31	163-180				1.40	4.90
1T/32	121-130	200-214			1.42	4.95

* 1T/19 ϵ , 90-97; 1T/19 ζ , 120-128; 1T/19 η , 130-140; 1T/19 θ , 143-149; 1T/19 κ , 156-164.

† 1T/24 ϵ , 139-153; 1T/24 ζ , 159-169.

Amino acid composition of fractions 1T/2-32

The average composition of this group of fractions was obtained by two methods: (a) by calculation from the measured compositions of fraction U.S. 3 and fraction 1T/1, assuming that fraction 1T/1 represents one-third of the weight of fraction U.S. 3; (b) by direct determination on a sample produced as follows. Fraction U.S. 3 was treated with trypsin as described above and after acidification with formic acid was eluted from a column (37 cm. $\times$ 6.5 cm.) of Zeo-Karb 225 (X2) resin with 6.3 l. of 0.1 N-acetic acid followed by 3.6 l. of 2 M-pyridine-acetic acid buffer, pH 6.7. Material, 3T/1, recovered from the first 1.8 l. of the effluent was found by amino acid analysis to correspond to fraction 1T/1, indicating that the material, 3T/2, recovered from the remainder of the effluent was equivalent to fractions 1T/2-32 (Table 2).

Criteria of purity of peptides

Peptides were considered pure if they showed less than 10% by weight of contamination according to the following

criteria: (a) fractionation by paper electrophoresis and paper chromatography revealed only one component; (b) the peptide consisted of amino acids in simple integral proportions; (c) the peptide had a single *N*-terminal amino acid.

Determination of the amino acid sequences of peptides

Preparation and analysis of partial hydrolysates. The following treatments were utilized to produce partial hydrolysis of the tryptic peptides: (a) 11 N-HCl at 37° or 5.7 N-HCl at 100°; (b) 0.03 N-HCl at 105° (Schultz, Allison & Grice, 1962); (c) digestion with chymotrypsin, pepsin, subtilisin or Pronase as described by Ambler (1963). The mixtures of peptides thus obtained were separated by high-voltage electrophoresis on paper at pH 1.85 or pH 6.5 (100 ml. of pyridine and 3.6 ml. of acetic acid/l.) and located and recovered as described above. The amino acid compositions of the peptides were determined by the methods of Blackburn & Lee (1963) or Corfield & Simpson (1965).

Fig. 3. Separation of peptides by one-dimensional paper chromatography in butan-1-ol-acetic acid-water (4:1:5, by vol.). The mixture of peptides (6mg.), dissolved in water or 0.2N-ammonia (80-200 μ l.), was applied as a line, 15cm. long, to a sheet of Whatman 3MM filter paper. Development of the chromatogram and the location and elution of the peptide bands are described in the text. Solid blocks denote peptide bands that give an intense stain with cadmium-ninhydrin reagent; open blocks denote bands giving stains of faint to medium intensity.

Table 2. Amino acid composition of peptide fractions from tryptic digest of fraction U.S.3

Results, expressed as the no. of residues/1000 residues, are not corrected for decomposition of amino acids during hydrolysis.

Amino acid	Fraction 1T/1 (obs.)	Fraction 3T/1 (obs.)	Mean	Fractions 1T/2-32 (calc.)	Fraction 3T/2 (obs.)	Mean
Ala	52	65	59	68	67	67
Arg	45	53	49	91	65	78
Asp	55	50	52	98	91	95
CySO ₃ H	161	166	164	26	50	38
Glu	60	58	59	182	155	168
Gly	88	100	94	51	50	51
His	4	5	4	8	5	6
Ile	30	28	29	40	47	44
Leu	45	47	46	116	120	118
Lys	12	9	11	49	48	48
MetO ₂	4	4	4	8	8	8
Phe	16	17	16	30	27	29
Pro	118	109	114	12	30	21
Ser	149	130	139	78	76	77
Thr	88	88	88	49	49	49
Tyr	14	23	19	38	44	41
Val	59	50	54	60	70	65
Total	1000	1002	1001	1004	1002	1003
No of residues/tryptic peptide			16.7			8.0
Mean residue weight			110			115

Edman degradation of peptides. This was carried out as described by Gray & Hartley (1963a). It was used only in a subtractive way.

N-Terminal residues of peptides. Reaction with DNS chloride was carried out as described by Gray & Hartley (1963b). After hydrolysis of the DNS peptide for 16 hr. in 5N-HCl, DNS-amino acids were separated by paper electrophoresis at pH 4.4 (Gray & Hartley, 1963b), or by thin-layer chromatography on silica gel (Cole, Fletcher & Robson, 1965). *N*-Terminal residues of peptides were also identified by reaction with [¹³¹I]pipsyl chloride (Fletcher, 1967) or by reaction with acrylonitrile (Fletcher, 1966). 1-Fluoro-2,4-dinitrobenzene (Sanger, 1945) was occasionally used (DNP method), but the amount of material available was generally insufficient. Deamination with nitrous fumes (Consden *et al.* 1947) was used to limited extent at the beginning of this work but was later abandoned as unreliable.

C-Terminal residues of peptides. Carboxypeptidase A was used as described by Canfield (1963). The liberated amino acids were separated by high-voltage electrophoresis on paper.

Measurement of yields of peptides

From paper strips. The weights of peptides obtained were calculated from the amino acid compositions of their hydrolysates. Values expressed in μ moles refer to the yield from 4.65 g. of fraction U.S.3.

From columns. In those cases where eluted fractions consisted of single peptides, yields were determined directly by weighing. When an eluted fraction consisted of two or more peptides the proportions of the mixture were calculated from the amino acid composition of its hydrolysate, and the yield of each peptide was then calculated from the total weight of peptides. This procedure was adopted to avoid errors inherent in measurements of yields of peptides eluted from paper strips.

Digestion of fraction 3T/1 with pepsin

The fraction was treated with pepsin (substrate/enzyme ratio 50:1) in aq. 5% (v/v) formic acid (250 ml./g. of substrate) at 40° for 48 hr. The course of reaction was followed by measuring the colour produced by reaction with ninhydrin of samples removed at intervals.

RESULTS

Structures of peptides isolated from the first tryptic digest of fraction U.S.3

In the sequences given below, the *C*-terminal residue is usually inferred from the specificity of trypsin. When two basic amino acid residues are present in a peptide, and one has been located elsewhere in the sequence, the other is assumed to be *C*-terminal.

Peptide 1T/4 β 2H2. This was found to be:

Peptide 1T/4 β 6. No *N*-terminus was detected

(deamination, DNS, pipsyl and DNP methods). Hydrolysis with concentrated hydrochloric acid at 37° for 24 hr. gave: Thr-(Glu,Leu) (DNP method); Thr-Glu (DNP method); (Asp,Glu₂,Leu,Thr): (Glu,Leu); Asp-Lys. Hydrolysis with chymotrypsin for 16 hr. gave: (Glu₂,Leu,Thr); Asp-Lys. The probable sequence is therefore:

Peptide 1T/4 γ 2. The *N*-terminus was: Glu (pipsyl and DNS methods). Hydrolysis with 0.03N-hydrochloric acid at 105° for 24 hr. gave: CySO₃H-(Ala,Glu) (pipsyl method); Glu-(Ala,Gly,Ser) (pipsyl method); (CySO₃H,Glu); (Gly,Ser)-Arg. Hydrolysis with concentrated hydrochloric acid at 37° for 48 hr. gave: (Ala,CySO₃H,Glu); (Ala,Glu); Ser-Gly-Arg (deamination method); Gly-Arg. The probable sequence is therefore:

Peptide 1T/4 γ 3. The *N*-terminus was: Glu (pipsyl method). Hydrolysis with 0.03N-hydrochloric acid at 105° for 40 hr. gave: (CySO₃H,Glu,Val); (Gly,Ser)-Arg. Hydrolysis with concentrated hydrochloric acid at 37° for 48 hr. gave: (Ala,CySO₃H,Glu,Val); Glu-Ala-Asp (DNP and carboxypeptidase methods); Glu-Ala (DNP method); Ser-Gly-Arg (DNP method); Gly-Arg (DNP method). The probable sequence is therefore:

Peptide 1T/5 γ 1. The *N*-terminus was: Ala (pipsyl and deamination methods). Hydrolysis with concentrated hydrochloric acid at 37° for 24 hr. gave: Ala-Gly (deamination method); Ser-(CySO₃H,Gly) (deamination method); Ser-(Arg,CySO₃H,Gly) (deamination method); Gly-Arg. The probable sequence is therefore:

Peptide 1T/5 γ 2. The *N*-terminus was: Ala (deamination method). Hydrolysis with concentrated hydrochloric acid at 37° for 24 hr. gave: Ser; (Ala,Gly). The probable sequence is therefore:

Peptide 1T/6 γ 1. This was found to be:

Peptide 1T/6 γ 2. The *N*-terminus was: Glu (deamination method). Hydrolysis with 0.03N-hydrochloric acid at 105° for 48 hr. gave: (Glu,Ser). Hydrolysis with concentrated hydrochloric acid at 37° for 7 hr. gave: (Glu,Ser); (Ala,Glu,Ser); Asp-Arg. The probable sequence is therefore:

Peptide 1T/6γ4. The *N*-terminus was: Ile (pipsyl and deamination methods). Hydrolysis with concentrated hydrochloric acid at 37° for 48 hr. gave: $\text{CySO}_3\text{H-Ala-Lys}$ (deamination method); $\text{CySO}_3\text{H-Ala}$ (deamination method); $(\text{Ala}, \text{CySO}_3\text{H}, \text{Leu})$; (Ile, Leu) . The probable sequence is therefore:

Peptide 1T/6δ4. The *N*-terminus was: Ile (DNS method). Hydrolysis with concentrated hydrochloric acid at 37° for 48 hr. gave: $(\text{CySO}_3\text{H}, \text{Gly})$; $(\text{CySO}_3\text{H}, \text{Leu}_2)$; $(\text{CySO}_3\text{H}, \text{Gly}_2)\text{-Lys}$; (Ile, Leu) ; Gly-Lys . The probable sequence is therefore:

Peptide 1T/8β2. This was the free amino acid:

Peptide 1T/8γ1. The *N*-terminus was: Ala or Ser (DNS method). Hydrolysis with concentrated hydrochloric acid at 37° for 24 hr. gave: $(\text{CySO}_3\text{H}, \text{Ser})$; $(\text{CySO}_3\text{H}, \text{Gly}, \text{Ser})$; (Phe, Ser) ; (Ala, Ser) ; Ser-Arg ; $(\text{Gly}, \text{Ser})\text{-Arg}$. Hydrolysis with chymotrypsin for 16 hr. gave: $(\text{Ala}, \text{Phe}, \text{Ser}_2)$; $(\text{CySO}_3\text{H}, \text{Gly}, \text{Ser}_2)\text{-Arg}$. Hydrolysis with Pronase for 5 hr. gave: $(\text{CySO}_3\text{H}, \text{Gly}, \text{Ser}_2)$; $(\text{CySO}_3\text{H}, \text{Gly}, \text{Ser})\text{-Arg}$; Ser-Ala (DNS method); Ser-Arg . The probable sequence is therefore:

Peptide 1T/8R3. No *N*-terminus was detected. Hydrolysis with concentrated hydrochloric acid at 37° for 48 hr. gave: $\text{Asp-(Asp, Glu, Leu, Ser)}$ (DNS method); $\text{Leu-(Asp, Glu, Ser)}$ (DNS method); $\text{Thr-(Ala, Glu)-Arg}$ (DNS method); (Glu, Thr) ; (Leu, Thr) ; $(\text{Ala}, \text{Glu}, \text{Ser}, \text{Thr})\text{-Arg}$. Hydrolysis with 0.03N-hydrochloric acid at 105° for 16 hr. gave: $(\text{Ala}, \text{Glu})\text{-Arg}$; $\text{Thr-(Ala, Glu)-Arg}$ (DNS method); $\text{Ser-(Ala, Glu, Thr)-Arg}$ (DNS method); (Glu, Leu) . The probable sequence is therefore:

Peptide 1T/9γ5. The *N*-terminus was: Asp-Val ... (leucine aminopeptidase method); Asp (pipsyl method). Hydrolysis with concentrated hydrochloric acid at 37° for 48 hr. gave: $(\text{Asp}, \text{CySO}_3\text{H}, \text{Val})$; $(\text{Ala}, \text{Asp}, \text{CySO}_3\text{H}, \text{Val})$; Leu-Arg . Hydrolysis with 0.03N-hydrochloric acid at 105° for 20 hr. gave: (Ala, Leu) ; $(\text{Ala}, \text{Leu})\text{-Arg}$; Leu-Arg . The probable sequence is therefore:

Peptide 1T/10R. The *N*-terminus was: Leu-Asp ... (Edman-DNS method); Leu (pipsyl method). Hydrolysis with concentrated hydrochloric acid at 37° for 7 days gave: $(\text{Ala}, \text{Asp}, \text{Leu}, \text{Val})\text{-Arg}$; Val-

$(\text{Asp}, \text{Leu})\text{-Arg}$ (deamination method); Val-Asp-(Ala-Glu) (Edman-DNS method); Asp-(Ala, Glu) (DNS method); Leu-(Asp, Val) (deamination method); Asp-Arg ; Ala-Pro (DNS method); (Asp, Leu) . Hydrolysis with 0.03N-hydrochloric acid at 105° for 2 days gave: $(\text{Ala}, \text{Asp}, \text{Glu}, \text{Leu}, \text{Val})\text{-Arg}$; Thr-Val (deamination method). Hydrolysis with chymotrypsin for 4 days gave: $(\text{Glu}, \text{Val})\text{-Leu}$ (specificity). The probable sequence is therefore:

Peptide 1T/11α2. The *N*-terminus was: Ser-Lys ... (Edman-DNS method). Hydrolysis with concentrated hydrochloric acid at 37° for 72 hr. gave: $\text{CySO}_3\text{H-Glu}$ (DNS method); $\text{Ser-(CySO}_3\text{H}, \text{Glu}, \text{Lys})$ (DNS method); $(\text{CySO}_3\text{H}, \text{Lys}, \text{Ser})$; (Lys, Ser) ; Glu-(Ile, Lys) (DNS method). Hydrolysis with 0.03N-hydrochloric acid at 105° for 7 hr. gave: $\text{Ser-(CySO}_3\text{H}, \text{Glu}, \text{Lys})$ (DNS method); (Lys, Ser) . The probable sequence is therefore:

Peptide 1T/11α3. The *N*-terminus was: Ala (pipsyl and DNS methods). Hydrolysis with concentrated hydrochloric acid at 37° for 96 hr. gave: $\text{MetO}_2\text{-Ala}$ (DNS method); (Ala, Lys) ; $(\text{Ala}, \text{Glu}, \text{Lys})$; $\text{Ala-Leu-CySO}_3\text{H}$ (Edman-DNS method); $(\text{CySO}_3\text{H}, \text{Leu}, \text{Lys})$; (Asp, Glu) . Hydrolysis with 0.03N-hydrochloric acid at 105° for 7 hr. gave: $\text{MetO}_2\text{-(Ala, CySO}_3\text{H}, \text{Leu}_2, \text{Lys})$ (DNS method); $(\text{Ala}, \text{CySO}_3\text{H}, \text{Leu}, \text{MetO}_2)\text{-Leu}$ (carboxypeptidase method); Ala-(Glu, Lys) (DNS method). The probable sequence is therefore:

Peptide 1T/11α3A. The *N*-terminus was: Asp (DNS method). Hydrolysis with concentrated hydrochloric acid at 37° for 72 hr. gave: $(\text{Ala}, \text{Asp}, \text{Glu}_2, \text{Leu})$; Glu-Leu (DNS method); $(\text{Glu}_2, \text{Leu})$; $(\text{Ala}, \text{Glu}_2, \text{Leu})$; Ser-Arg . Hydrolysis with 0.03N-hydrochloric acid at 105° for 7 hr. gave: $\text{Ala-(Glu}_2, \text{Leu}, \text{Ser})\text{-Arg}$ (DNS method); (Leu, Ser) ; $(\text{Ala}, \text{Glu}_2, \text{Leu}, \text{Ser})$. The probable sequence is therefore:

Peptide 1T/12R. The *N*-terminus was: Leu-Val-Val ... (Edman-DNS method); Leu (pipsyl method). Hydrolysis with concentrated hydrochloric acid at 37° for 72 hr. gave: $(\text{Asp}, \text{Glu}, \text{Ile})$; (Asp, Ile) ; (Ala, Asp) ; $(\text{Asp}, \text{Glu}, \text{Ile}, \text{Val})$; Ala-Lys . Hydrolysis with 0.03N-hydrochloric acid at 105° for 48 hr. gave: $(\text{Leu}, \text{Val}_2)$; $(\text{Leu}, \text{Val})\text{-Val-Glu}$ (carboxypeptidase method); Ala-Lys ; Ile . The probable sequence is therefore:

Peptide 1T/12γ1. The *N*-terminus was: Thr-Glu ... (Edman-DNS method). Hydrolysis with concentrated hydrochloric acid at 37° for 72 hr. gave: Thr-(Glu,Leu) (DNS method); Glu-Glu-Lys; Glu-Lys; (Glu,Ile,Leu) (DNS-Ile after Edman degradation; *N*-terminus not detected). The probable sequence is therefore:

Thr-Glu-Leu-Ile-Glu-Glu-Lys

Peptide 1T/13α3. No *N*-terminus was detected. Hydrolysis with Pronase for 5 hr. gave: Glu-Ser (DNS method); (Glu,Ser)-Arg; Ser-Arg; (Glu,Ser)₂. The probable sequence is therefore:

Ser-Glu-Ser-Arg

Peptide 1T/13α4. This was found to be:

Thr-Lys

Peptide 1T/13β2. This was the free amino acid:

Arg

Peptide 1T/14β1H3. The *N*-terminus was: ?-Glu-Ser ... (Edman-DNS method). Hydrolysis with 5.7*N*-hydrochloric acid at 105° for 20 min. gave: Asp-(Glu,Ser) (DNS method); Glu-Ser (DNS method); (Ala,Glu,Ser). Hydrolysis with Pronase for 5 hr. gave: (Ala,Glu,Ser); (Asp,Glu,Ser); (Ala,Asp,Glu,Ser); (Glu,Ser); Ala-Arg. The probable sequence is therefore:

Asp-Glu-Ser-Ala-Arg

Peptide 1T/17γ2. The *N*-terminus was: Ser-Glu ... (Edman-DNS method). Hydrolysis with concentrated hydrochloric acid at 37° for 72 hr. gave: Ser-Glu (DNS method); Ser-(Glu,Leu) (DNS method); Leu-Gly (DNS method); (Asp,Gly)-Arg; Asp-Arg; (Asp,Glu,Gly,Leu,Ser). Hydrolysis with 0.03*N*-hydrochloric acid at 105° for 7 hr. gave: (Glu,Gly,Leu,Ser). The probable sequence is therefore:

Ser-Glu-Leu-Gly-Asp-Arg

Peptide 1T/17γ5. The *N*-terminus was: Leu-Leu-Glu ... (Edman-DNS method). Hydrolysis with concentrated hydrochloric acid at 37° for 72 hr. gave: Leu-(Glu,Leu) (DNS method); (Glu,Leu); Glu-Gly (DNS method); Glu-Glu-Arg; Glu-Arg; (Glu₂Gly,Leu₂). Hydrolysis with 0.03*N*-hydrochloric acid at 105° for 7 hr. gave: (Glu₂Gly,Leu₂). The probable sequence is therefore:

Leu-Leu-Glu-Gly-Glu-Glu-Arg

Peptide 1T/18β1. The *N*-terminus was: Ala-Glu ... (Edman-DNS method). Hydrolysis with concentrated hydrochloric acid for 37° for 72 hr. gave: Ala-Glu (DNS method); Leu-Ala (DNS method); Ser-Arg; Ala-(Asp,Glu,Tyr) (DNS method); Asp-

(Ala,Leu) (DNS method); (Ala,Asp,Glu); (Ala,Asp₂Glu,Ile/Leu,Tyr). Hydrolysis with 0.03*N*-hydrochloric acid at 105° for 24 hr. gave: (Ala,Glu); (Ala,Ile/Leu,Ser)-Arg. The probable sequence is therefore:

Ala-Glu-Asp-Tyr-Asp-Leu-Ala-Ile-Ser-Arg

Peptide 1T/18γ2. The *N*-terminus was: Val-Leu-Asp-Glu ... (Edman-DNS method). Hydrolysis with concentrated hydrochloric acid at 37° for 24 hr. gave: (Leu,Val); Val-Leu-(Asp-Glu) (Edman-DNS method); Thr-Arg; (Glu,Thr)-Arg. The probable sequence is therefore:

Val-Leu-Asp-Glu-Thr-Arg

Peptide 1T/18δ8. The *N*-terminus was: Ala (DNS method). Hydrolysis with concentrated hydrochloric acid at 100° for 20 min. gave: Thr-(Ile,Tyr,Val)-Arg (DNS method); Ile-Arg; (Ile,Val,Tyr)-Arg; (Tyr,Val). Hydrolysis with Pronase for 5 hr. gave: (Ala,Thr,Val); (Ile,Tyr,Val). The probable sequence is therefore:

Ala-Thr-Val-Tyr-Ile-Arg

Peptide 1T/19γ1. This was found to be:

Ser-Arg

Peptide 1T/19ζ1. This was found to be:

Glu-Arg

Peptide 1T/19ζ2. The *N*-terminus was: Glu-?-MetO₂ ... (Edman-DNS method). Hydrolysis with 0.03*N*-hydrochloric acid at 105° for 6 hr. gave: (Glu,Leu,MetO₂,Phe,Thr). Hydrolysis with pepsin for 72 hr. gave: (Glu,MetO₂,Thr); (Asp₂Leu,Phe)-Arg. Hydrolysis with Pronase for 5 hr. gave: Glu-Thr-MetO₂ (Edman-DNS method); Asp-Asp-Arg. The probable sequence is therefore:

Glu-Thr-MetO₂-(Leu,Phe)-Asp-Asp-Arg

Peptide 1T/19η3. The *N*-terminus was: Phe (DNS method). Hydrolysis with Pronase for 4 hr. gave: Leu-Glu-Glu (Edman-DNS method); Glu-Glu; Phe; Leu; Asp-Lys. The probable sequence is therefore:

Phe-Leu-Glu-Glu-Asp-Lys

Peptide 1T/19κ2. The *N*-terminus was: Gly-Ile ... (Edman-DNS method). Hydrolysis with concentrated hydrochloric acid at 37° for 48 hr. gave: Gly-Ile (DNS method); (CySO₃H,Gly,Ile,Tyr); Ser-Arg. Hydrolysis with Pronase for 6.5 hr. gave: (CySO₃H,Gly,Ile); Tyr; Ser-Arg. The probable sequence is therefore:

Gly-Ile-CySO₃H-Tyr-Ser-Arg

Peptide 1T/2081. This was found to be:

Ala-Arg

Peptide 1T/21 α 1. No *N*-terminus was detected. Hydrolysis with pepsin for 48hr. gave: Ala-(Glu, Val) (DNS method); (Asp, Glu, Phe); (Ala, Glu, Leu); (Glu, Val); Val-Lys; (Glu, Val)-Lys. Hydrolysis with Pronase for 5hr. gave: (Ala, Glu, Thr); (Ala, Glu); (Ala, Leu). Hydrolysis with subtilisin for 17hr. gave: (Ala₂, Thr). The probable sequence is therefore:

(Asp, Phe)-Glu-Leu-Ala-Thr-Ala-Glu-Val-Lys

Peptide 1T/21 β 4. The *N*-terminus was: Ile (DNS method). Hydrolysis with Pronase for 5hr. gave: Val-Glu-Arg (DNS method); Glu-Arg. The probable sequence is therefore:

Ile-Glu-Val-Glu-Arg

Peptide 1T/23 β 1. The *N*-terminus was: Ile (DNS method). Hydrolysis with concentrated hydrochloric acid at 37° for 48hr. gave: (Glu, Ile, Ser); (Thr, Tyr)-Arg; Tyr-Arg. The probable sequence is therefore:

Ile-(Glu, Ser)-(Asp, Glu, Leu)-Thr-Tyr-Arg

Peptide 1T/23 γ 3. The *N*-terminus was: Ile-Leu ... (Edman-DNS method). Hydrolysis with Pronase for 5hr. gave: Glu-Arg. The probable sequence is therefore:

Ile-Leu-Glu-Arg

Peptide 1T/24 δ 5. The *N*-terminus was: Leu (DNS method). Hydrolysis with 0.03*N*-hydrochloric acid at 105° for 6hr. gave: (Glu, Leu, Tyr). Hydrolysis with pepsin for 48hr. gave: (Glu, Leu); (Glu, Leu, Tyr); Leu-(Glu₂, Tyr) (DNS method); Glu-Ile-Arg (DNS method). Hydrolysis with Pronase for 5hr. gave: Ile-Arg (DNS method); Glu-Tyr (DNS method). The probable sequence is therefore:

Leu-Glu-Tyr-Glu-Ile-Arg

Peptide 1T/24 ϵ 4. The *N*-terminal was: Leu (acrylonitrile method). Hydrolysis with 5.7*N*-hydrochloric acid at 105° for 30min. gave: Leu-Ala (DNS method); (Glu, Leu); Glu-Lys; (Leu, Ser, Tyr). Hydrolysis with Pronase for 6hr. gave: Ala-Ser (Edman-DNS method); Glu-Lys; Tyr; Leu. Hydrolysis with subtilisin for 5hr. gave: (Leu, Ser, Tyr); (Glu, Leu)-Lys. The probable sequence is therefore:

Leu-Ala-Ser-Tyr-Leu-Glu-Lys

Peptide 1T/24 ζ 3. This was found to be:

Val-Arg

Table 3. Structure of peptides isolated in low yield from tryptic digest 1T of fraction U.S.3

Peptide	Yield (μ mole)	Sequence
1T/3 γ 3	0.3	Asp-(CySO ₃ H, Pro)-Gly-Glu-Ser-Val-Arg
1T/4 β 1	0.2	CySO ₃ H-Glu-Asp-Ser-Ser-Lys
1T/7 γ 5	0.6	CySO ₃ H-Leu-Asp-Ser-Asp-Arg
1T/9 δ 5	0.3	Leu-(Asp, Glu, Gly, Leu, Val)-Lys
1T/13 α 5	0.2	Ala-Lys
1T/17 δ 2	0.5	Gly-Ser-Arg
1T/29 β 1	0.5	Leu-Gly-Glu-Arg

Peptide 1T/27 α 4. No *N*-terminus was detected. Hydrolysis with 5.7*N*-hydrochloric acid at 100° for 40min. gave: (Leu, Thr). Hydrolysis with subtilisin for 4hr. gave: (Gly, Thr)-Arg; (Gly, Leu, Thr)-Arg. The probable sequence is therefore:

(Glu, His)-Leu-Thr-Gly-Arg

Peptide 1T/30 β 1. The *N*-terminus was: Thr (DNS method). Hydrolysis with 5.7*N*-hydrochloric acid at 100° for 30min. gave: Thr-(Glu, Lys, Ser, Tyr) (DNS method); (Lys, Tyr); (Glu, Ser, Tyr); Glu-Ser (acrylonitrile method); (Glu, Ser)-Arg; Glu-Arg. Hydrolysis with Pronase for 5hr. gave: (Lys, Thr); (Lys, Thr, Tyr); (Glu, Ser); Ser-Glu-Arg (acrylonitrile method); Glu-Arg. The probable sequence is therefore:

Thr-Lys-Tyr-Ser-Glu-Arg

Minor peptides. The structures of other peptides observed in the tryptic digest of fraction U.S.3 are given in Table 3.

Yields of peptides obtained from fraction 2T/32

More accurate determinations of the weights of some peptides in the tryptic digest were obtained by examination of this fraction by methods that avoided elution from paper strips before the measurement of yields. Fraction 2T/32 (130mg.) was shaken with 20ml. of water and the insoluble residue, 2T/32R (58mg.), separated by centrifuging. The soluble material and the residue were each fractionated on a column (150cm. $\times$ 1.8cm.) of De-Acidite G (2-3% cross-linked; 200 mesh) by using gradient elution with an initial buffer of 0.2*M*-pyridine-acetic acid, pH 6.5, changing to 2*N*-acetic acid. Examination of the six components thus obtained by electrophoresis on paper showed that three consisted of single peptides and the others all contained pairs of peptides. The yields of these peptides were determined as described above and are shown in Table 4.

Table 4. *Yields of peptides in fraction 2T/32*

The yields of peptides from fraction 2T/32 (see Fig. 2b), which was subfractionated by column procedures only, are summarized and compared where possible with the yields of the same peptides isolated from tryptic digest 1T by fractionation procedures involving separation by paper electrophoresis or paper chromatography. Yields are based on 5g. of fraction U.S.3.

Peptide	Yield (μ moles)	Equivalent peptide	Yield (μ moles)
Val-Arg	60.2	1T/24 ζ 3	3
(Arg,Leu,Ser)	29.1		
(Arg,Asp,Glu ₂ Ile,Leu,Ser,Thr,Tyr)	25.1	1T/23 β 1	1.2
Leu-Ala-Ser-Tyr-Leu-Glu-Lys	37.0	1T/24 ϵ 4	8
(Ala ₂ ,Arg,Asp,Glu,Leu ₂ ,Lys,Ser ₂ ,Tyr)	7.5		
Leu-Glu-Tyr-Glu-Ile-Arg	23.6	1T/24 δ 5	4
(Ala ₂ ,Arg,Asp ₃ ,Glu ₃ ,His,Leu ₃ ,Ser,Thr,Val)	7.7	1T/24R2	1.1

Investigation of fraction 3T/1

This fraction represents one-third of the weight of fraction U.S.3 and contains a large proportion of the cysteic acid peptides. These peptides were strongly retained by columns of De-Acidite G and could not be eluted from the resin. They were eluted as a single peak, with very little evidence of fractionation, by gel filtration with water or phenol solutions. They also moved as a single band during zone electrophoresis in both neutral and acid buffers.

The intractability of this fraction necessitated the use of further hydrolysis to provide material more amenable to fractionation. After 48hr. digestion with pepsin a peptide mixture was obtained that gave 17 discernible bands on paper electrophoresis in pH 1.85 buffer. However, it proved impossible to isolate pure peptides from the peptic digest by ion-exchange chromatography, paper chromatography or paper electrophoresis; this was due partly to the complexity of the digest and partly to the tailing undergone by the negatively charged peptides during separation by paper methods.

DISCUSSION

Earlier work suggested that fraction U.S.3 consists of a number of large polypeptides derived from a single structural unit of the wool fibre by oxidative scission of its disulphide bonds and limited hydrolytic fission of its labile peptide bonds (Corfield, 1963). Since the fractionation procedure used to isolate it restricts peptide-bond fission to a minimum, the polypeptides of fraction U.S.3 are probably closer to the composition of the parent structure than are wool fractions of somewhat similar amino acid composition, such as α -keratose (Corfield, Robson & Skinner, 1958) or *S*-carboxymethyl-keratine A (Gillespie, O'Donnell, Thompson & Woods, 1960). On the assumption,

therefore, that the polypeptides of fraction U.S.3 contain in more or less degree all the elements of primary structure that form the parent unit, some idea of the size of the latter may be derived from the number and nature of the peptides obtained in high yield from a tryptic digest of fraction U.S.3.

The structures of the peptides isolated to date from the digest show that normal splitting of peptide bonds has occurred, because only one of the peptides present in high yield has an amino acid residue other than arginine or lysine at the C-terminus. Moreover, since only 5% of these peptides contain more than one basic amino acid residue reaction with trypsin must have been substantially complete. Having established the specificity of reaction with trypsin, therefore, the major difficulty in calculating the molecular weight of the parent protein unit from the peptides in the tryptic digest was to distinguish peptides of the precursor from those originating in protein impurities. Most of the peptides were isolated in yields greater than 1 μ mole or less than 0.1 μ mole, and we therefore set the limiting yield arbitrarily at 0.75 μ mole.

Table 5 shows the amino acid composition of fractions 1T/2-32 compiled from the 58 peptides considered on this basis to form the major part of the primary structure of the precursor, together with the calculated amino acid composition of fractions 1T/2-32. The good agreement between the observed and calculated values suggests that the selection of the relevant peptides cannot be much in error. There can be no doubt, however, that some peptides of structural significance were present in the digest that could not be isolated in sufficient quantity and purity for sequence studies, e.g. peptides in fractions 1T/15 and 1T/16.

If we assume that the contribution of protein impurities to the weight of fraction U.S.3 is negligible, i.e. the material consists entirely of polypeptide chains of the parent protein, an

Table 5. *Composition of major peptides from fractions 1T/2-32*

Peptide	Yield (μ -moles)	Amino acid residues																		Total
		Ala	Arg	Asp	Cys-O ₃ H	Glu	Gly	His	Ile	Leu	Lys	Met-O ₂	Phe	Pro	Ser	Thr	Tyr	Val		
1T/2 β 1	1.8		1		2									1	2			1	7	
1T/4 β 2H2	3				1						1								2	
1T/4 β 6	5			1		2				1	1					1			6	
1T/4 γ 2	5	2	1	1	1	2	1								1				9	
1T/4 γ 3	1.5	2	1	1	1	2	1								1			1	10	
1T/5 γ 1	9	1	1		1		2								1				6	
1T/5 γ 2	3	1					1								1				3	
1T/6 γ 1	4		1		1														2	
1T/6 γ 2	2	1	1	1		1									1				5	
1T/6 γ 4	5	1			1				1	1	1								5	
1T/6 δ 2	3	1	1	1	1	1	2			1				1	2			1	12	
1T/6 δ 3	2				1		2		1	1	1				1			1	8	
1T/6 δ 4	2				1		2		1	2	1								7	
1T/7 β 1	1.5	2		2	1	2				1	1				1	1		1	12	
1T/8 β 2	5										1								1	
1T/8 γ 1	1.5	1	1		1		1						1		4				9	
1T/8R2	7	1	1	1	1	2	2			2					2			2	14	
1T/8R3	5	1	1	2		2				2					1	2			11	
1T/9 γ 5	3	1	1	1	1					1								1	6	
1T/10R	20	2	1	4		2				3				1		1		3	17	
1T/11 α 2	0.7				1	2			1		2				1				7	
1T/11 α 3A	1.7	1	1	1		2				1					1				7	
1T/11 α 3	1.3	2		1	1	1				2	2	1							10	
1T/12 γ 1	7					3			1	1	1					1			7	
1T/12R	12	1		2		1			1	1	1							2	9	
1T/13 α 3	1.4		1			1									2				4	
1T/13 α 4	0.9										1					1			2	
1T/13 β 2	3		1																1	
1T/14 β 2	0.8	3		1		2	1			2	1		1			1		1	13	
1T/14 β 1H3	2	1	1	1		1									1				5	
1T/17 γ 2	1.2		1	1		1	1			1					1				6	
1T/17 γ 5	1.2		1			3	1			2									7	
1T/18 β 1	2	2	1	2		1			1	1					1		1		10	
1T/18 γ 2	4		1	1		1				1						1		1	6	
1T/18 δ 8	5	2	1							1						1	1	1	7	
1T/19 γ 1	1.9		1												1				2	
1T/19 ζ 1	0.9		1			1													2	
1T/19 ζ 2	2		1	2		1				1		1	1			1			8	
1T/19 η 3	4			1		2				1	1		1						6	
1T/19 η 5	1.0	1	1	1		1	1			1			1		1				8	
1T/19 η 6	3		1	1		1	1		1	2								1	8	
1T/19 κ 2	0.75		1		1		1		1						1		1		6	
1T/20 δ 1	1.9	1	1																2	
1T/21 α 1	0.9	2		1		2				1	1		1			1		1	10	
1T/21 β 4	4		1			2			1									1	5	
1T/23 β 1	1.2		1	1		2			1	1					1	1	1		9	
1T/23 γ 3	1.1		1			1			1	1									4	
1T/24 β 2	2		1	1	1	3				1			1	1	1		2	1	13	
1T/24 δ 5	4		1			2			1	1							1		6	
1T/24 ϵ 4	8	1				1				2	1				1		1		7	
1T/24 ζ 3	3		1															1	2	
1T/24R	1.1	2	1	3		3		1		3					1	1		1	16	
1T/25 α 3	0.8			1		2	1			2	1				1		1	1	10	
1T/27 α 1	1.6	1	1	1	1	2			1	1						1	1		10	
1T/27 α 4	4		1			1	1	1		1						1			6	
1T/27 γ 6	1.8	2		2		1	1		1		1		2				1		11	
1T/29 α 1	0.8			1		3	2	1		3	1				1		1	2	15	
1T/30 β 1	1.5		1			1					1				1	1	1		6	
Totals		39	38	41	20	67	25	3	15	50	22	2	9	4	35	17	13	25	425	
Calc. totals		29	33	40	16	71	22	3	19	50	20	3	12	9	33	21	17	27	425	

approximate molecular weight for the latter can be derived as follows. Fractions 1T/2-32 contain 58 major peptides constituted from 425 amino acid residues with a mean residue weight of 116, whereas the amino acid residues of fraction 1T/1, which weighed 1.57 g., or one-third of the weight of fraction U.S.3, have a mean residue weight of 110 and must therefore number about 225. Hence the parent protein might be expected to consist of some 650 amino acid residues, to yield about 70 peptides on digestion with trypsin and have a molecular weight of approx. 74000. Its estimated 24 cystine residues should give 48 cysteic acid residues on oxidation, and of these 20 have been completely or partly characterized; the remainder must be in the 13 tryptic peptides yet to be isolated from fraction 1T/1.

By far the most unsatisfactory aspect of the work on which the foregoing argument is based is the very poor recovery of peptides from the first tryptic digest of fraction U.S.3. If the 4.65 g. of fraction U.S.3 digested by trypsin had consisted wholly of a parent protein unit with a molecular weight of 74000 it should have provided 63 μ moles of every one of the 58 peptides identified. Clearly, however, this estimate must be too high since it takes no account of extraneous protein, incomplete digestion of the structural unit, or our failure to identify all the relevant peptides. Even so, with the notable exceptions of peptides 1T/10R and 1T/12R, the yields of 56 peptides, based on the amounts recovered from weighed samples subjected to paper chromatography or electrophoresis, fell within the range 0.75-9 μ moles, only a small fraction of the expected value of, say, 50 μ moles. Early in the investigation we suspected that some of these yields were extremely low, since the peptides concerned were isolated in fractions of a milligram whereas their chromatograms showed that they comprised the bulk of the 6 mg. of sample usually applied. Incomplete elution of peptides from paper strips was thus implicated. The comparatively high yields of peptides 1T/10R and 1T/12R, which partially separated from solution in a pure state during the removal of buffer salts and water by freeze-drying, confirmed this view. The serious nature of the losses incurred in this way was later emphasized by the amounts of peptides determined in fraction 2T/32 compared with the amounts of the same peptides isolated from the fraction 1T digest by a succession of separatory procedures culminating in paper methods (Table 4). When account is taken of the additional losses to be expected from the overlapping of peptide peaks in the initial fractionation on resin and from the incomplete dissolution of the material recovered from this column before re-chromatography, the yields of peptides believed to

be part of the primary structure of the parent protein are probably not far off the 'expected value' of about 50 μ moles.

Two points must be mentioned in connexion with the sequence work. First, the oxidation of proteins modifies tryptophan residues in a way not yet understood, and the location of these residues in wool must be regarded as a separate problem. Secondly, a proportion of the residues reported as glutamyl and aspartyl must in fact be glutaminyl and asparaginy residues. The allocation of amide groups will have to be based on a further study of the electrophoretic behaviour of the relevant peptides and the products of their hydrolysis by procedures that preserve amide groups intact.

One of the disappointments of this investigation has been our inability to resolve the fraction 1T/1, one-third of fraction U.S.3, into its constituent peptides. This leads to the speculation that this material may not in fact be a mixture of peptides but a resistant core of polypeptide whose basic residues are not susceptible to hydrolysis by trypsin because of the proximity of negatively charged cysteic acid residues. Any high-sulphur moiety of the wool structure after oxidation might be expected to be equally resistant to trypsin and to possess all the properties of fraction 1T/1. This doubt will be clarified when more information is available from an analysis of the chymotryptic digest of fraction U.S.3.

Our sequence data may be compared with those of Fell, La France & Ziegler (1960) and M. Fell (unpublished work; quoted by Crewther *et al.* 1965) for α -keratose, of Blackburn & Lee (1966b) for oxidized wool, and of Consden, Gordon & Martin (1949) and of Consden & Gordon (1950) for the virgin fibre.

Of the dipeptide sequences found in wool by Consden *et al.* (1949) all those containing aspartic acid or glutamic acid occur in fraction U.S.3 with the exception of Glu-CyS. Only about half the dipeptide sequences containing cysteic acid that have been found in wool have also been identified in fraction U.S.3, but this is to be expected since more than half of the cysteic acid residues in fraction U.S.3 are in the fraction 1T/1. About two-thirds of the sequences shown by Fell *et al.* (1960) to occur in α -keratose have also been found in fraction U.S.3 (Table 6). Asterisks against the 'corresponding peptides from fraction U.S.3' indicate that in these cases there is some doubt that the pairs are identical, but the equivalences in amino acid composition are sufficiently close to focus attention on them since differences in amino acid composition and sequence may reflect errors in analysis rather than differences in structure. Indeed, when the difficulties involved in the isolation and purification of the relevant peptides

Table 6. *Tentative identification of certain tryptic peptides from α -keratose with tryptic peptides from fraction U.S.3*

Peptides from α -keratose are those obtained by Fell *et al.* (1960) and M. Fell (unpublished work; quoted by Crewther *et al.* 1965).

Peptides from α -keratose	Corresponding peptide from U.S.3
(Arg, Glu(NH ₂) ₂ , Val, Ile)-Arg	
Ala-(Thr, Val, Ile)-Arg	1T/1888*
Ser-Ser-Arg	
Phe-Arg	
Gly-Ser-Arg	1T/1782
Ser-Arg	1T/1902
Ser-Lys	1T/11 α 2
Gly-Arg	in 1T/4 γ 3
Ala-Lys	1T/13 α 5
Ala-Arg	1T/2081
Lys-Lys	1T/8 β 2
Ser-(His, Asp ₂ , Tyr, Ile)-Arg	
Ala-(Gly ₅ , Glu, Ile)-Arg	
Ser-(Asp, Glu, Leu)-Arg	1T/17 γ 2a*
Gly-(CySO ₃ H, Gly, Ser, Ala)-Arg	1T/5 γ 1*
(Asp, Val)-Arg	
Leu-(Ser, Glu, Ala, Tyr, Leu)-Lys	1T/24 ϵ 4
Ala-(Asp, Glu, Ile, Thr, Phe)-Lys	1T/21 α 1*
(Asp, Glu, Leu, Ile)-Lys	
Phe-(Asp, Glu ₂ , Leu)-Lys	1T/19 η 3
Leu-(Thr, Glu, Leu)-Lys	1T/12 γ 1*
Glu-Ile-Lys	in 1T/11 α 2
(Glu, CySO ₃ H ₂)	
Asp-(CySO ₃ H, Asp, Glu, Ala, Val, Leu)-Arg	in 1T/10R*
Glu-(CySO ₃ H, Asp, Ser ₁₋₂ , Gly, Glu ₄₋₅ , Ala ₄)-Arg	1T/4 γ 2*

* Correspondence is not exact (see the Discussion section).

Table 7. *Identification of certain tryptic peptides from oxidized wool with tryptic peptides from fraction U.S.3*

Peptides from oxidized wool are those obtained by Blackburn & Lee (1966b).

Peptides from oxidized wool	Corresponding peptide from fraction U.S.3
Ser-Arg	1T/19 γ 1
Thr-Lys	1T/30 β 1
Ala-Lys	1T/13 α 5
CySO ₃ H-Lys	1T/482H2
Ala-Phe	
Ser-Gly-Arg	1T/1782
Ser-Glu-Leu-Gly-Asp-Arg	1T/17 γ 2
Asp-Val-CySO ₃ H-Ala-Leu-Arg	1T/9 γ 5
Leu-Leu-Glu-Gly-Glu-Glu-Arg	1T/17 γ 5
(Lys, Ser, Glu)	
(Lys, Ser, Asp)	
(Lys, Leu, Thr, Glu ₂ , Asp)	1T/486
(Arg, Leu, Thr)	
(Arg, Ala, Ser, Leu)	
(Arg, Ala, Ser, Leu, Glu ₂ , Asp)	1T/11 α 3A
(Arg, Val, Leu, Thr, Glu, Asp)	1T/18 γ 2

Table 8. *Frequency of occurrence of certain dipeptide sequences in fraction U.S.3*

The observed frequencies of occurrence of the given dipeptide sequences in the 248 dipeptide sequences characterized in the 58 major tryptic peptides are compared with the frequencies expected on the basis of a random arrangement of the amino acids involved.

Sequence	Observed no.	Calculated no.
Glu-Glu	5	6.3
Asp-Asp	1	2.1
Glu-Asp	5	3.6
Asp-Glu	3	3.6
Acid-acid	14	15.5
Glu-Arg	5	4.5
Asp-Arg	4	2.6
Glu-Lys	2	2.4
Asp-Lys	2	1.4
Acid-base	13	10.7
Base-acid	9-11	10.7
Base-base	≥ 2	7.4

are considered, agreement between our findings and those of Fell *et al.* (1960) for α -keratose is fairly good. Correspondence between some of the structures of the tryptic peptides of fraction U.S.3 and peptides identified by Blackburn & Lee (1966b) during a preliminary investigation of the tryptic digest of oxidized wool is considerably better (Table 7). There is complete agreement on the determined sequences, and the five peptides out of the 16 that were identified in oxidized wool, but apparently are not major peptides in the tryptic digest fraction of U.S.3, were all small.

In the so-called high-sulphur fractions from wool, such as γ -keratose and SCMKB, and in fraction 1T/1, proline residues occur with a frequency of 1 in 10 and half-cystine residues with a frequency of 1 in 5. The frequency of occurrence of proline residues among the peptides in Table 3, however, is less than 1 in 100, and there is not a single proline residue among the peptides listed in Tables 6 and 7. The frequencies of occurrence of

half-cystine (cysteic acid) residues among the peptides in Tables 5, 6 and 7 are 1 in 21, 1 in 24 and 1 in 32 respectively. It seems likely that the primary structures of the high-sulphur regions of wool are represented by only a few of the peptides isolated. These data therefore do not resolve the problem whether wool consists of two proteins with high and low sulphur contents, or almost entirely of a single protein with high-sulphur

and low-sulphur moieties as structural features (Corfield, 1963; Blackburn & Lee, 1966b; Robson, 1966).

Our results, however, help to resolve one minor structural problem. Suggestions have been made from time to time that polar residues are concentrated in groups along the peptide chains of wool. Table 8 shows the frequencies of occurrence of certain dipeptide sequences of this type in fraction U.S.3 compared with those expected on a random arrangement of the amino acid residues concerned. It is apparent from these results that there is no especial tendency for acidic amino acid residues to occur in pairs, nor is there any evidence of excessive association of acidic and basic residues. Moreover, the low yields of free arginine and lysine observed in the tryptic digest indicate that basic residues rarely occur in adjacent positions, although the relatively large number of dipeptides observed suggests that the sequence 'base-X-base' might have some special significance for the structure of wool.

The authors are indebted to Mr C. Myers for maintaining and running the amino acid analyser, and to Miss P. C. Craven, Miss T. M. Kelham, Miss P. B. Teale and Mr M. Cole for skilful technical assistance. This work is sponsored by the Agricultural Research Service of the U.S. Department of Agriculture under the Authority of Public Law 480.

REFERENCES

- Ambler, R. P. (1963). *Biochem. J.* **89**, 349.
 Atfield, G. N. & Morris, C. J. O. R. (1961). *Biochem. J.* **81**, 606.
 Blackburn, S. (1965). *Meth. biochem. Anal.* **13**, 1.
 Blackburn, S. & Lee, G. R. (1963). *Biochem. J.* **87**, 1P.
 Blackburn, S. & Lee, G. R. (1966a). 3^e Congr. int. Recherche Textile Lainière, vol. 1, p. 313.
 Blackburn, S. & Lee, G. R. (1966b). 3^e Congr. int. Recherche Textile Lainière, vol. 1, p. 321.
 Canfield, R. E. (1963). *J. biol. Chem.* **238**, 2698.
 Cole, M., Corfield, M. C., Fletcher, J. C. & Robson, A. (1966). 3^e Congr. int. Recherche Textile Lainière, vol. 1, p. 335.
 Cole, M., Fletcher, J. C. & Robson, A. (1965). *J. Chromat.* **20**, 616.
 Consden, R. & Gordon, A. H. (1950). *Biochem. J.* **46**, 8.
 Consden, R., Gordon, A. H. & Martin, A. J. P. (1947). *Biochem. J.* **41**, 590.
 Consden, R., Gordon, A. H. & Martin, A. J. P. (1949). *Biochem. J.* **44**, 548.
 Corfield, M. C. (1963). *Biochem. J.* **86**, 125.
 Corfield, M. C., Fletcher, J. C., Myers, C. & Robson, A. (1966). 3^e Congr. int. Recherche Textile Lainière, vol. 1, p. 345.
 Corfield, M. C. & Robson, A. (1962). *Biochem. J.* **84**, 146.
 Corfield, M. C., Robson, A. & Skinner, B. (1958). *Biochem. J.* **68**, 348.
 Corfield, M. C. & Simpson, E. C. (1965). *J. Chromat.* **17**, 420.
 Crewther, W. G., Fraser, R. D. B., Lennox, F. G. & Lindley, H. (1965). *Advanc. Protein Chem.* **20**, 191.
 Fell, M., La France, N. H. & Ziegler, K. (1960). *J. Text. Inst.* **51**, T797.
 Fletcher, J. C. (1966). *Biochem. J.* **98**, 34c.
 Fletcher, J. C. (1967). *Biochem. J.* **102**, 815.
 Gillespie, J. M., O'Donnell, I. J., Thompson, E. O. P. & Woods, E. F. (1960). *J. Text. Inst.* **51**, T703.
 Gray, W. R. & Hartley, B. S. (1963a). *Biochem. J.* **89**, 379.
 Gray, W. R. & Hartley, B. S. (1963b). *Biochem. J.* **89**, 59P.
 Heilman, J., Barollier, J. & Watzke, E. (1957). *Hoppe-Seyl. Z.* **309**, 219.
 Keston, A. S., Udenfriend, S. & Cannan, R. K. (1949). *J. Amer. chem. Soc.* **71**, 249.
 Moore, S. & Stein, W. H. (1954). *J. biol. Chem.* **211**, 907.
 Priestley, G. C. & Speakman, P. T. (1966). *Nature, Lond.*, **209**, 1336.
 Robson, A. (1966). *Text. Inst. Industr.* **4**, 37.
 Sanger, F. (1945). *Biochem. J.* **39**, 507.
 Schultz, J., Allison, M. & Grice, M. (1962). *Biochemistry*, **1**, 694.