Structural Studies on Wheat (*Triticum aestivum*) Proteins Lacking Phenylalanine and Histidine Residues

By DAVID G. REDMAN

Flour Milling and Baking Research Association, Chorleywood, Rickmansworth, Herts. WD3 5SH, U.K.

(Received 4 March 1975)

1. Three very similar proteins, each of approx. 120 amino acid residues but lacking phenylalanine and histidine, were isolated from wheat (*Triticum aestivum*) flour in sufficient quantities for further structural studies. 2. Each protein, after reduction and carboxymethylation, was cleaved at the three methionine residues with CNBr to give four major peptides, which were isolated. These peptides are suitable for future sequencing studies, as the sums of their amino acid compositions are in good agreement with those of the whole proteins. 3. The *N*- and *C*-terminal peptides were identified. 4. Evidence from amino acid analyses, *N*-terminal amino acids and electrophoretic mobilities of the peptides suggests a high degree of homology between the proteins. Definite differences in *C*-terminal amino acids and the number of glycine, alanine and arginine residues were found in the *C*-terminal peptides.

A major albumin of hexaploid wheat that lacks phenylalanine and histidine residues has been described by several workers (Ewart, 1969; Feillet & Nimmo, 1970; Shainkin & Birk, 1970). Later studies (Cantagalli *et al.*, 1971; Silano *et al.*, 1973), indicated that a closely related series of such proteins existed, although pure wheat varieties were not used. The proteins were strong inhibitors of mid-gut α -amylase of the larva of the yellow mealworm (*Tenebrio molitor*), but inactive against other α -amylases (Shainkin & Birk, 1970; Silano *et al.*, 1973). In the present work, large-scale preparations were undertaken as a preliminary to sequence studies. It became evident that multiple forms of the albumin existed, even in a pure variety, and three proteins were isolated. Sequence studies then became of even greater importance, as they would pinpoint the differences within this family of proteins. This paper describes the isolation and partial characterization of the peptides obtained after CNBr treatment of the three reduced and carboxymethylated proteins.

Protein nomenclature

Various names have been given to wheat albumins lacking phenylalanine and histidine residues. These are ALB 13B (Feillet & Nimmo, 1970) based on the system of Feillet & Bourdet (1968); AmI₁ (Shainkin & Birk, 1970); and the 0.28, 0.32, 0.35, 0.39 and 0.48 family of proteins described by Silano *et al.* (1973). Although the proteins described in the present paper are undoubtedly related to the proteins isolated by all these workers, for simplicity the terms S, M and F will be used denoting the slowest, mid-mobility and fastest proteins respectively during electrophoresis at pH 3.2.

Materials and Methods

Materials

Flours were prepared from Cappelle-Desprez wheats on a Bühler mill to approx. 70% extraction.

Suppliers of reagents and other materials were as follows: Sephadex G-25 (medium grade) and G-50 (fine grade), Pharmacia Fine Chemicals, Uppsala, Sweden; DEAE-cellulose (DE52) and CM-cellulose (CM52), Whatman Biochemicals Ltd., Maidstone, Kent, U.K.; sulphoethyl-cellulose, Serva, Heidelberg, West Germany; CNBr, Ralph N. Emanuel Ltd., Wembley, U.K.; dansyl chloride and toluene-4-sulphonic acid, BDH Chemicals Ltd., Poole, Dorset, U.K.; 2-mercaptoethanol, 5,5'-dithiobis(2-nitrobenzoic acid) (Ellman's reagent) and L-homoserine, Koch-Light Laboratories, Colnbrook, Bucks., U.K.; tryptamine [3-(2-aminoethyl)indole] hydrochloride, iodoacetic acid and L-homoserine, Sigma Chemical Co., St. Louis, Mo. U.S.A.; α -amino- γ -butyrolactone (homoserine lactone) hydrobromide, Aldrich Chemical Co., Milwaukee, Wis., U.S.A.; anhydrous hydrazine, Pierce Chemical Co., Rockford, Ill., U.S.A. Polyamide layer sheets (Cheng Chin Trading Co. Ltd., Taipei, Taiwan) were obtained through BDH Chemicals Ltd.

Carboxypeptidase A (bovine pancreas) was from Schwarz Bioresearch Inc., Orangeburg, N.Y., U.S.A., carboxypeptidase C (1000 units/mg of enzyme) from Röhm GmbH, Darmstadt, West Germany, and insulin from Boots Pure Drug Co. Ltd., Nottingham, U.K.

Preparation of mixture of proteins S+M+F

Extraction of flour and (NH₄)₂SO₄ precipitation were as described by Cantagalli *et al.* (1971) except

that flour was magnetically stirred for 2.5 h at approx. 20°C. The washed final precipitate from 1500 g of flour was suspended in 0.1 M-NaCl-0.02% NaN₃ (25 ml), dialysed against this solution for 5 h and, after centrifugation at 1800 g for 20 min, the supernatant was fractionated on a column (170 cm × 5.5 cm) of Sephadex G-50 swollen and eluted with the NaCl-Na₃N solution. The last of the three major peaks (detected on this and other columns by absorption at 254 nm with an L.K.B. Uvicord spectrophotometer) was dialysed against water before freeze-drying.

Separation of protein S from proteins M+F

The material from two such preparations was dissolved in 25 ml of 0.02 M-Tris-HCl buffer, pH 8.6, dialysed against 1 litre of this buffer and loaded on a column (55 cm × 3.5 cm) of DEAE-cellulose equilibrated in the same solution. Elution was by a convex NaCl gradient. The mixing flask contained 1800 ml of the buffer, and the reservoir 0.15 M-NaCl in buffer. The two major peaks with centres at approx. 1000 ml (M+F) and 1400 ml (S), omitting shoulders on leading edges, were dialysed against water and freeze-dried.

Final purification of proteins S, M and F

Equivalent peaks from two DEAE-cellulose runs were combined for CM-cellulose chromatography on columns (55 cm × 3.5 cm) equilibrated in 16 mM-sodium acetate-4 mM-acetic acid, pH 5.2. Proteins were dissolved in 25 ml of this buffer and dialysed against the buffer before loading.

Protein S was eluted from CM-cellulose with a convex NaCl gradient. The mixing flask contained 1800 ml of acetate buffer and the reservoir 0.15 M-NaCl in this buffer. Protein S (peak centre approx. 300 ml), separated from a shoulder on its leading edge, was dialysed against water and freeze-dried.

Proteins M and F were separated on CM-cellulose by elution with starting buffer to give two incompletely resolved major peaks with maxima at approx. 500 and 700 ml. The leading edge (M) and trailing edge (F) and overlap region were dialysed against water and freeze-dried. The proteins in the overlap region were subjected to chromatography under the same conditions to yield more proteins M and F. Re-chromatography of proteins M and F under the same conditions gave single peaks.

Carboxymethylation of proteins

Reduction and carboxymethylation were carried out by the methods of Crestfield *et al.* (1963) as described by Leach (1966a,b) with the following modifications. The pH of the reaction mixture during carboxymethylation was kept at approx. 7.5 by the addition of 2 M-NaOH. At the end of the reaction,

2-mercaptoethanol (sufficient to react with one-third of the iodoacetic acid originally used) was added to avoid reaction of iodoacetic acid with amino acid residues other than cysteine in subsequent stages. After a further 10 min the pH was adjusted to 4 with acetic acid, and the solution dialysed against 0.01 M-acetic acid for 24 h. The suspension of carboxymethylated protein was freeze-dried.

Cleavage at methionine residues

Proteins (usually carboxymethylated) were treated with CNBr in 70% (v/v) formic acid (Gross & Witkop, 1961; Steers *et al.*, 1965) for 24 h in the dark at approx. 20°C. Protein concentration was 10 mg/ml, with a CNBr/protein ratio of 1:1 (w/w). At the completion of the reaction, the solution was diluted fivefold with water and freeze-dried.

Fractionation of CNBr peptides

(1) *Gel filtration.* Sephadex G-50 was swollen in 9.1% (v/v) acetic acid for 65 h (prolonged swelling was essential) before being packed into a column (135 cm × 2.4 cm). A 5 cm layer of Sephadex G-25 on the top of the column improved resolution by eliminating channelling during loading. Up to 150 mg of CNBr peptides was loaded in 5 ml of 9.1% acetic acid, the same solvent being used for elution. Fractions corresponding to peaks (see Fig. 1 and the Results section) were freeze-dried.

(2) *Ion-exchange chromatography of gel-filtration fractions.* CM-cellulose columns (40 cm × 1.6 cm) were equilibrated at pH 3.5 (5 mM-sodium acetate adjusted with acetic acid). Samples from one Sephadex gel-filtration run were loaded in 5 ml of this buffer. A convex NaCl gradient was applied with a mixing vessel containing 170 ml of acetate buffer and the reservoir 0.17 M-NaCl in this buffer. Fractions, concentrated when necessary by freeze-drying, were desalted on columns (135 cm × 2.4 cm) of Sephadex G-25 or G-50 in 9.1% acetic acid before freeze-drying.

Sulphoethyl-cellulose columns (40 cm × 1.6 cm) were equilibrated at pH 2.0 (13 mM-HCl-50 mM-KCl). Samples were loaded in 5 ml of this buffer. A convex pH gradient was applied with a mixing vessel containing 170 ml of the pH 2.0 buffer and the reservoir 16 mM sodium acetate-4 mM-acetic acid, pH 5.2. Fractions were desalted on Sephadex G-50 as described above.

Electrophoretic characterization

Starch-gel electrophoresis was performed at pH 3.2 in aluminium lactate buffer (Ewart, 1966). CNBr peptides were run for 3.5 h at 10 V/cm. Polyacrylamide-gels were run at pH 8.9 under the conditions described by Davis (1964).

Amino acid analysis

Samples were hydrolysed in constant-boiling HCl under N₂ at 105°C for 24h, before analysis on a Technicon 120cm single-column analyser. The protein S was hydrolysed also for 48 and 72h to obtain correction factors for destruction or incomplete liberation of amino acids. The same factors were used to correct the 24h values for proteins M and F. All other analyses were uncorrected. Performate oxidation before 24h hydrolysis was by the method of Hirs (1956) except that after reaction the solution was directly freeze-dried, before suspending in water and freeze-drying again.

Tryptophan was determined on the analyser by using the elution system of Redman (1975), after hydrolysis in toluene-*p*-sulphonic acid in the presence of 3-(2-aminoethyl)indole *in vacuo* for 23h (Liu & Chang, 1971). Peptides were hydrolysed in this same mixture.

N-Terminal amino acid detection

Dansylation was performed in 0.4M-NaHCO₃-0.5% dansyl chloride in acetone (1:1, v/v) for 2.5h at 37°C. After drying *in vacuo* the derivatives were hydrolysed in constant-boiling HCl for 6h at 110°C. HCl was removed *in vacuo* and the tube contents were applied in pyridine-water (1:1, v/v), to both sides of a polyamide plate (7.5cm×7.5cm). A dilute standard mixture of dansyl amino acids was co-spotted on one side. Development was as described by Gray (1972), except that toluene replaced benzene in the second solvent.

C-Terminal amino acid detection

(1) The hydrazinolysis procedure of Akabori *et al.* (1952) as described by Press *et al.* (1966) was followed by amino acid analysis.

(2) Carboxypeptidase A was brought into solution at pH 8.3 by the method of Mitz as quoted by Hirs *et al.* (1960). Reaction mixtures (0.6ml) contained 40 EC units of enzyme, and 0.18 μmol of protein or 0.1 μmol of peptide. Incubation was at 37°C for 5h.

(3) Carboxypeptidase C digestion was in 2.8mm-citric acid-7.2mm-trisodium citrate, pH 5.3, for 5.5h at 37°C. Reaction mixtures (0.6ml) contained 0.25 mg of enzyme, and 0.18 μmol of protein or 0.1 μmol of peptide. Proteins were performate-oxidized before digestion.

Thiol determination

The Ellman (1959) reagent was used in sodium dodecyl sulphate solution as described by Habeeb (1972).

Results

Each protein was isolated in a yield of approx. 0.005% by weight of the flour (i.e. approx. 0.05% with respect to protein).

Electrophoresis of proteins

The three proteins each ran as single bands towards the cathode on starch-gel electrophoresis. The mobility of protein S was 0.96 with respect to proteins M and F. Protein F was shown to be very slightly faster migrating than protein M after prolonged electrophoresis.

On polyacrylamide-gel electrophoresis each protein ran predominantly as a pure band towards the anode, with minor faster moving components. Mobilities of the proteins with respect to Bromophenol Blue were S, 0.42; M, 0.31; F, 0.31.

CNBr treatment of non-reduced protein S gave a single band on starch-gel electrophoresis (mobility 1.07 with respect to protein S).

Amino acid analyses of proteins (Table 1)

The values are calculated to a mol.wt. of approx. 13000 (Feillet & Nimmo, 1970; Ewart, 1972; Silano *et al.*, 1973). The proteins contained no thiol groups.

N- and C-Terminal amino acids of proteins

These are shown in Table 2.

Gel filtration of CNBr peptides

The elution patterns of the CNBr peptides on Sephadex G-50 are shown in Fig. 1 (similar patterns from all three proteins). Major peptides were found in peaks A and B, and the corresponding starch-gel electrophoresis patterns are in Fig. 2. Peptides from protein F were electrophoretically identical with those from protein M. Peaks X and Y were not always as well defined and gave rise to variable electrophoretic patterns. However, in the particular case depicted in Fig. 1, representing CNBr peptides of a protein S sample, peak X gave bands as shown in Fig. 2. Peak Y gave a band of mobility equal to peptides 1.

Insulin (mol.wt. 5733) emerged at 320ml (peak centre) from the same column under the same conditions.

Electrophoresis and nomenclature of CNBr peptides

Peptides are referred to as S-CNBr-1 etc., the number referring to the position in the electrophoretic patterns (Fig. 2). Peaks A consisted of three major peptides and peaks B one major peptide. The latter (peptides 4) were used without further purification as

Table 1. *Amino acid compositions of proteins S, M and F*

Values are residues/protein molecule. See the text for details.

Protein	Residues			Nearest integers		
	S	M	F	S	M	F
Asp	9.07	9.03	9.11	9	9	9
Thr	3.15	3.11	3.14	3	3	3
Ser	8.69	9.27	9.54	9	9	10
Glu	12.59	12.34	12.75	13	12	13
Pro	9.91	9.29	9.90	10	9	10
Gly	11.30	10.25	10.40	11	10	10
Ala	10.06	10.06	10.11	10	10	10
Val	13.04	12.87	11.92	13	13	12
Cys	8.20*	7.96	8.09	10*	10*	10*
Met	3.08†	2.89	3.03	3	3	3
Ile	2.08	2.03	2.06	2	2	2
Leu	8.98	8.75	9.07	9	9	9
Tyr	3.69	4.07	4.01	4	4	4
Phe	Trace	Trace	Trace	0	0	0
Lys	6.26	5.94	5.95	6	6	6
His	Trace	Trace	Trace	0	0	0
Arg	5.75	6.74	6.99	6	7	7
Trp	3.77	3.85	3.76	4	4	4
Total	—	—	—	122	120	122

* Values of 8.47 cysteic acid residues and 9.34 *S*-carboxymethylcysteine residues were obtained on samples of performate-oxidized protein S and of reduced and carboxymethylated protein S respectively. The Cys value may therefore be 10 residues for each protein (zero cysteine residues).

† 2.90 residues as methionine sulphone on performate-oxidized protein S.

Fig. 1. Gel-filtration elution profile of CNBr peptides (30 mg) from Sephadex G-50

Similar patterns were obtained for the fragments from all three proteins. For details see the text.

they contained only minor slower-moving impurities. Protein S differed from proteins M and F in that peptide 2 was of lower mobility. This peptide from the latter two proteins behaved identically with peptide 4 on starch-gel electrophoresis.

Separation of peptides CNBr-1, -2 and -3

Peaks A (peptides 1+2+3) were fractionated on CM-cellulose.

Peak A from protein S gave one peak at approx.

Fig. 2. Starch-gel electrophoretic patterns (pH 3.2) of CNBr peptides before and after gel-filtration

Peaks A and B are as depicted in Fig. 1. For details see the text. ▨, Major bands; —, minor bands; ----, faint bands.

120ml (peptides 1+2) and another (low absorbance) at approx. 180ml (peptide 3 plus the minor band of slightly lower mobility).

Peaks A from both proteins M and F behaved similarly to that from protein S except that peptides

1 and 2 were resolved. The return to baseline after peptide 1 was almost complete, but slight cross-contamination of peptides 1 and 2 was observed on subsequent starch-gel electrophoresis.

Peptides 1+2 from protein S were separated on sulphoethyl-cellulose. Peptide 2 (pure on starch-gel electrophoresis) was eluted between pH 2.02 and 2.28, and peptide 1 between pH 3.19 and 4.26. In this case peptide 1 ran as a closely spaced doublet on starch-gel electrophoresis unless it was first dissolved in trifluoroacetic acid and left for 2 h at approx. 20°C before drying *in vacuo*. This treatment converted the slower band to the mobility of the faster, i.e. equal to the normal mobility of peptides 1. It is probable that some C-terminal homoserine lactone became converted into homoserine during the preparation. Some of the peptide would have been above pH 4 overnight at room temperature after separation on the column. Apparently the peptides with C-terminal homoserine and homoserine lactone were separable at pH 3.2 on starch-gel electrophoresis. Lactonization in trifluoroacetic acid (Ambler, 1965) before electrophoresis, by 'neutralization' of the carboxyl group, increased mobility towards the cathode.

Amino acid analyses of CNBr peptides

These are in Tables 3 and 4. The peptides contained no half-cystine, methionine, phenylalanine and histidine residues. Difficulty was encountered in analysis of homoserine plus homoserine lactone arising from the action of CNBr on methionine residues. Homoserine (two commercial samples)

and homoserine lactone each gave two peaks, corresponding to the free acid and the lactone (homoserine emerging before glutamic acid and the lactone before lysine). Treatment of homoserine with trifluoroacetic acid or 2M-NH₃ before analysis,

Table 2. *N*- and C-terminal amino acids of proteins and CNBr peptides

See the text for details. Abbreviations: H, hydrazinolysis; CPA, carboxypeptidase A; CPC, carboxypeptidase C; Hse, homoserine; -, positions not determined.

Protein or peptide	<i>N</i> -Terminal*	C-Terminal†
S	Ser	Val (H,CPC)
M	Ser	Val (H,CPA,CPC)
F	Ser	Ala (CPC); none detected (H)
S-CNBr-1	Val	-
M-CNBr-1	Val	Hse (CPA)
F-CNBr-1	Val	-
S-CNBr-2	Lys	-
M-CNBr-2	Lys	Val (CPA)
F-CNBr-2	Lys	Ala (CPC)
S-CNBr-3	Leu	-
M-CNBr-3	Leu	Hse (CPA)
F-CNBr-3	Leu	-
S-CNBr-4	Ser	-
M-CNBr-4	Ser	Hse (CPA)
F-CNBr-4	Ser	-

* No other *N*-terminal amino acids were detected in significant amounts.

† Major amino acid liberated.

Table 3. Amino acid compositions of CNBr peptides

Values are residues/peptide molecule. See the text for details. Abbreviation: CmCys, *S*-carboxymethylcysteine

Protein ...	CNBr-1 peptides			CNBr-2 peptides			CNBr-3 peptides			CNBr-4 peptides		
	S	M	F	S	M	F	S	M	F	S	M	F
CmCys	5.21	4.89	4.86	1.98	1.73	2.00	1.59	1.30	1.21	1.94	1.97	2.04
Asp	5.58	5.12	5.22	3.10	2.84	2.97	0.81	0.47	0.35	1.04	1.12	1.24
Thr	0.22	0.24	0.22	0.95	0.93	0.93	0.20	0.16	0.15	1.98	1.91	1.89
Ser	3.16	3.11	3.02	1.89	1.85	1.79	1.64	1.45	1.33	2.67	2.76	2.76
Hse	1.00	1.25	1.16	Absent	Absent	Absent	1.29	1.28	1.13	1.29	1.00	1.20
Glu	6.15	6.07	6.00	1.24	1.09	1.16	6.40	6.03	5.68	0.16	0.27	0.40
Pro	1.74	1.91	2.06	5.56	5.17	5.28	1.89	1.55	1.39	2.09	2.23	2.07
Gly	2.45	2.34	2.31	3.10	1.84	2.11	4.30	3.69	3.73	3.12	2.96	2.88
Ala	2.70	2.65	2.86	4.88	3.94	4.88	0.90	0.47	0.37	3.06	2.93	2.84
Val	3.95	4.01	3.83	4.35	4.18	3.95	4.14	3.77	3.58	1.00	1.08	1.10
Ile	1.18	1.08	1.10	1.03	0.97	0.98	0.32	0.13	0.10	0.04	0.05	0.07
Leu	3.87	3.75	3.70	1.09	1.06	1.06	3.91	3.06	3.08	1.07	1.10	1.21
Tyr	0.27	0.23	0.35	1.91	1.44	1.81	1.26	1.09	1.01	1.02	0.96	0.93
Lys	1.05	1.07	1.03	1.91	1.97	2.04	2.38	2.16	2.23	0.99	0.92	0.97
Trp	0.91	0.83	0.86	0.72	0.56	0.58	0.20	0.06	Trace	1.31	1.53	1.50
Arg	1.98	1.99	2.03	0.07	1.04	1.02	2.80	2.91	2.77	1.01	1.08	1.07

to convert the material entirely into the lactone or the free acid respectively (Ambler, 1965), again resulted in both peaks on the analyser. Thus extensive interconversion took place on the column. Much of the homoserine and homoserine lactone appeared between the two peaks as indicated by a rise in the baseline after homoserine. Voelkert & Grant (1970) found limited opening of the lactone ring during amino acid analysis at 50°C, i.e. 10°C lower than in the present studies. In all the cases described in the present work, homoserine was the major peak. On the basis of these investigations, it was concluded that the most satisfactory way to assess the amount of homoserine plus homoserine lactone was to measure the area of only the former peak. Homoserine gave a peak with a ninhydrin colour yield of 27% of that of the internal standard norleucine (570nm). Homoserine lactonized in trifluoroacetic acid gave rise to a homoserine peak with a colour yield of 28%. Homoserine converted entirely into the free acid by 2M-NH₃ treatment gave a peak with a colour yield of 35%. Therefore in order to assess the amount of homoserine plus homoserine lactone in peptide hydrolysates the area of the former peak was multiplied by a factor of 3.7 (i.e. 100/27), before relating it to the area of the norleucine peak. An error of approx. 30% could have arisen if the material was present entirely as the free acid when loaded on the column. However, this is unlikely after acid hydrolysis. It is stressed that colour constants should be assessed for each analytical system. When this factor was applied to an analysis of a hydrolysate of CNBr-treated, carboxymethylated protein S, a value of 2.60 homoserine residues was found compared with a theoretical 3 if all the methionine residues had reacted. A methionine peak corresponding to 0.17 residue was detected in the same hydrolysate. This agreement was considered to be satisfactory.

The peptide compositions have been calculated so that the sum of their content of threonine, isoleucine, tyrosine, lysine and arginine fitted best with the composition of the appropriate protein. Reasonable values for homoserine contents were obtained after this procedure. Errors in estimates of the size of any of the peptides would become evident in future sequencing studies.

N- and C-terminal amino acids of CNBr peptides

These are shown in Table 2.

Discussion

Amino acid analyses (Table 1) of the three proteins suggest their close sequential relationship. Consideration of both their compositions and electrophoretic mobilities leads to the conclusion that

Table 4. Integer values of amino acid residues in CNBr peptides

Protein ...	S										M										F									
	CNBr-1					CNBr-2					CNBr-3					CNBr-4					CNBr-1					CNBr-2				
	Total	1-2	3	4	5	Total	1-2	3	4	5	Total	1-2	3	4	5	Total	1-2	3	4	5	Total	1-2	3	4	5	Total	1-2	3	4	5
CmCys	5	2	3	1	1	5	2	3	1	1	5	2	3	1	1	5	2	3	1	1	5	2	3	1	1	5	2	3	1	1
Asp	5-6	3	1	0	0	3	1	0	0	0	3	1	0	0	0	3	1	0	0	0	3	1	0	0	0	3	1	0	0	0
Thr	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0
Ser	3	2	1	0	0	2	1	0	0	0	2	1	0	0	0	2	1	0	0	0	2	1	0	0	0	2	1	0	0	0
His	1	0	1	1	0	1	0	1	1	0	1	0	1	1	0	1	0	1	1	0	1	0	1	1	0	1	0	1	1	0
Glu	6	1	6-7	0	0	1	6	1	0	0	1	6	1	0	0	1	6	1	0	0	1	6	1	0	0	1	6	1	0	0
Pro	2	5-6	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Gly	2-3	3	4	3	3	3	4	3	3	3	3	4	3	3	3	3	4	3	3	3	3	4	3	3	3	3	4	3	3	3
Ala	3	5	1	1	1	5	1	1	1	1	5	1	1	1	1	5	1	1	1	1	5	1	1	1	1	5	1	1	1	1
Val	4	1	4	4	4	1	4	4	4	4	1	4	4	4	4	1	4	4	4	4	1	4	4	4	4	1	4	4	4	4
Ile	1	1	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0
Leu	4	1	4	1	1	1	4	1	1	1	1	4	1	1	1	1	4	1	1	1	1	4	1	1	1	1	4	1	1	1
Tyr	0	2	1	1	1	0	2	1	1	1	0	2	1	1	1	0	2	1	1	1	0	2	1	1	1	0	2	1	1	1
Lys	1	2	2	2	2	1	2	2	2	2	1	2	2	2	2	1	2	2	2	2	1	2	2	2	2	1	2	2	2	2
Trp	1	1	0	1	1	1	0	1	1	1	1	0	1	1	1	1	0	1	1	1	1	0	1	1	1	1	0	1	1	1
Arg	2	0	3	1	1	0	3	1	1	1	0	3	1	1	1	0	3	1	1	1	0	3	1	1	1	0	3	1	1	1
Total	40-42	33-34	32-34	23-24	23-24	40	31-32	27-31	23-24	121-122	40	31-32	27-31	23-24	121-127	40	31-32	27-31	23-24	121-127	40	31-32	27-31	23-24	121-127	40	31-32	27-31	23-24	121-127

* Cys residues.

† Met residues.

protein S resembles most closely the 0.39 protein, and proteins M and F the 0.28 protein of Silano *et al.* (1973). Minor bands evident on electrophoresis at alkaline pH value may have been partially deamidated forms, or further minor members of the same family (cf. Silano *et al.*, 1973). In view of the presence of phenylalanine and histidine in only trace amounts, contamination with non-related proteins was low. Proteins M and F were very similar in electrophoretic mobilities but differed in C-terminal amino acids. Protein S, lacking one arginine residue, was readily distinguished from the other proteins.

CNBr cleavage gave four major peptides from each protein, as expected from the knowledge that these proteins each consist of a single chain (Ewart, 1969, 1972; Silano *et al.* 1973) and contain three methionine residues. The similarity of equivalent peptides (Fig. 2; Table 3) emphasizes the probability of there being sequential likenesses in the proteins. The sums of the amino acid compositions (Table 4) are in good agreement with those of the intact proteins, an essential prerequisite to sequential studies. Amide contents of proteins and peptides were not determined, but in view of the lack of satisfactory methods it was considered that identification of individual residues during sequencing would be preferable.

CNBr-1 peptides from all three proteins are probably identical (Table 3). Differences were found between CNBr-2 peptides. The absence of an arginine residue in peptide S-CNBr-2 is in accord with the absence of an arginine residue in the intact protein when compared with proteins M and F. Peptide S-CNBr-2 apparently contains one more glycine residue, and peptide M-CNBr-2 one less alanine residue, than the equivalent peptides from the other two proteins. Other dissimilarities may not be significant. CNBr-3 peptides were the least pure as indicated by the deviation of the ratios of amino acid residues from integral values. Precise location of these peptides in column eluates was hindered by low optical absorbance, a fact which is now explained by the lack of tryptophan residues. Possible differences exist in several amino acid residues. The three CNBr-4 peptides appear identical.

Peptides M-CNBr-2 and F-CNBr-2 were shown to arise from the C-terminal regions of the proteins by the absence of homoserine (Table 4), and similarity to the corresponding proteins with respect to carboxypeptidase treatment (Table 2). Peptide S-CNBr-2 may also be placed at the C-terminus by the absence of homoserine and by assuming homology of the three proteins. All other CNBr peptides contained homoserine (Table 4) and this was shown to be liberated from the M-CNBr peptides by carboxypeptidase (Table 2).

N-Terminal amino acid determinations showed that peptides CNBr-4 are from the N-terminal regions of the proteins.

The ordering of peptides CNBr-1 and CNBr-3 in the chains has not been definitely established. However, preliminary studies involving partial characterization of materials in peaks X and Y (Fig. 1) (D. G. Redman, unpublished work), suggest that the order of the fragments in the proteins may be, reading from the N-terminus, CNBr-4-CNBr-1-CNBr-3-CNBr-2.

No peptides were detached from the non-reduced protein S by CNBr treatment. Therefore the sections corresponding to the CNBr peptides from this protein were linked in the intact protein by disulphide bonds. The detection of carboxymethylcysteine residues in all the CNBr peptides is in accord with this result. Also, the carboxymethylcysteine residues in those CNBr peptides possessing only two such residues (Table 4) could not have been linked originally to each other. The same considerations very probably apply to proteins M and F, again by reason of homology.

A notable feature of the compositions of the CNBr peptides is that 12 of the 13 glutamine residues are in peptides CNBr-1 and CNBr-3, i.e. the central region of the protein of approx. 70 residues. The isolated peptides, because of their size and also the distribution of residues susceptible to common proteolytic enzymes, seem well suited to sequencing studies.

I thank Dr. J. A. D. Ewart for discussion during the course of the work, and Mr. D. Smith for experimental assistance.

References

- Akabori, S., Ohno, K. & Narita, K. (1952) *Bull. Chem. Soc. Jpn.* **25**, 214-218
- Ambler, R. P. (1965) *Biochem. J.* **96**, 32p
- Cantagalli, P., Di Giorgio, G., Morisi, G., Pocchiari, F. & Silano, V. (1971) *J. Sci. Food Agr.* **22**, 256-259
- Crestfield, A. M., Moore, S. & Stein, W. H. (1963) *J. Biol. Chem.* **238**, 622-627
- Davis, B. J. (1964) *Ann. N. Y. Acad. Sci.* **121**, 404-427
- Ellman, G. L. (1959) *Arch. Biochem. Biophys.* **82**, 70-77
- Ewart, J. A. D. (1966) *J. Sci. Food Agr.* **17**, 526-532
- Ewart, J. A. D. (1969) *J. Sci. Food Agr.* **20**, 730-733
- Ewart, J. A. D. (1972) *J. Sci. Food Agr.* **23**, 701-705
- Feillet, P. & Bourdet, A. (1968) *Ann. Technol. Agr.* **17**, 217-225
- Feillet, P. & Nimmo, C. C. (1970) *Cereal Chem.* **47**, 447-464
- Gray, W. R. (1972) *Methods Enzymol.* **25**, 121-138
- Gross, E. & Witkop, B. (1961) *J. Am. Chem. Soc.* **83**, 1510-1511

- Habeeb, A. F. S. A. (1972) *Methods Enzymol.* **25**, 457–464
- Hirs, C. H. W. (1956) *J. Biol. Chem.* **219**, 611–621
- Hirs, C. H. W., Moore, S. & Stein, W. H. (1960) *J. Biol. Chem.* **235**, 633–647
- Leach, S. J. (1966a) in *A Laboratory Manual of Analytical Methods of Protein Chemistry* (Alexander, P. & Lundgren, H. P., eds.), vol. 4, p. 63, Pergamon Press, Oxford
- Leach, S. J. (1966b) in *A Laboratory Manual of Analytical Methods of Protein Chemistry* (Alexander, P. & Lundgren, H. P., eds.), vol. 4, pp. 69–70, Pergamon Press, Oxford
- Liu, T. Y. & Chang, Y. H. (1971) *J. Biol. Chem.* **246**, 2842–2848
- Press, E. M., Piggot, P. J. & Porter, R. R. (1966) *Biochem. J.* **99**, 356–366
- Redman, D. G. (1975) *J. Chromatogr.* **104**, 178–179
- Shainkin, R. & Birk, Y. (1970) *Biochim. Biophys. Acta* **221**, 502–513
- Silano, V., Pocchiari, F. & Kasarda, D. D. (1973) *Biochim. Biophys. Acta* **317**, 139–148
- Steers, E., Jr., Craven, G. R. & Anfinsen, C. B. (1965) *J. Biol. Chem.* **240**, 2478–2484
- Voelkert, E. & Grant, D. R. (1970) *Anal. Biochem.* **34**, 131–137