N-Terminal Amino Acid Sequence of Wheat Proteins that Lack Phenylalanine and Histidine Residues

By DAVID G. REDMAN

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

(Received 6 January 1976)

The 24 residues of the *N*-terminal CNBr peptide from a wheat albumin, that lacks phenylalanine and histidine, have been sequenced. Three of the assignments were made partly by analogy with two other proteins, as evidence is presented that all three proteins are probably identical in this region. Extra evidence for the sequences of the α -chymotryptic peptides derived from the *N*-terminal CNBr peptides of the three proteins, and also for their assembly, has been deposited as Supplementary Publication SUP 50063 (11 pages) at the British Library Lending Division, Boston Spa, Wetherby, West Yorkshire LS23 7BQ, U.K., from whom copies may be obtained on the terms given in *Biochem. J.* (1976) 153, 5. The nature of such evidence is stated in the text of this present communication.

The CNBr peptides from three wheat albumins that lack phenylalanine and histidine residues were partially characterized by Redman (1975). The peptides from the *N*-terminal regions of these three proteins were identical in integral amino acid composition, in electrophoretic mobility and in possessing *N*-terminal serine. They contained 23–24 amino acid residues (1–2 tryptophan residues). Kasarda *et al.* (1974) and Patey *et al.* (1975) reported identity of the *N*-terminal sequences between α_1 - and α_2 - and between γ_1 - and γ_2 -gliadins from wheat respectively. The sequences of α - and γ -gliadins matched at seven of the first 12 residues. This present communication enables a contrast to be made in this region with major wheat albumins, as sequence studies on this class of protein have not been reported previously.

The nomenclature in the previous paper (Redman, 1975) is retained, the *N*-terminal CNBr peptides being termed S-CNBr-4, M-CNBr-4 and F-CNBr-4, where S, M and F are the proteins.

Experimental

Materials

Peptides S-CNBr-4, M-CNBr-4 and F-CNBr-4 were prepared as described previously (Redman, 1975).

α -Chymotrypsin ($3 \times$ crystallized, type 1) was from Sigma Chemical Co., St. Louis, MO, U.S.A. Tri-fluoroacetic acid and phenyl isothiocyanate (Sequenal grades) were from Pierce Chemical Co., Rockford, IL, U.S.A. Fluram (fluorescamine) was from Roche Products, London W1M 6AP, U.K. Sources of other

chemicals were as described previously (Redman, 1975).

Methods

α -Chymotryptic digestion. CNBr-4 peptides (4 mg) were digested in 0.5% (w/v) NH_4HCO_3 (0.5 ml) containing α -chymotrypsin (0.1 mg). Incubation was at 37°C for 5 h in the presence of a drop of toluene, before freeze-drying.

Preparation of peptide 'maps'. Electrophoresis at pH 3.5 and subsequent chromatography were as described by Ewart (1966). A maximum of 4 mg of peptides was loaded on a paper. Drying before chromatography and again before staining was for 90 min at 70°C. Peptide spots were located by dipping papers in a solution of fluorescamine (1 mg) in a mixture of acetone (300 ml) and pyridine (3 ml). (In early experiments up to 8 mg of fluorescamine was used. This did not apparently interfere with subsequent characterization, but it was found possible to decrease the quantity to 1 mg.) After the paper had been dried in a current of air for 30 min, spots were detected under u.v. light (360 nm).

Elution of peptides. Peptide spots, stained with fluorescamine, were cut out and eluted with 9.1% (v/v) acetic acid (approx. 1 ml), before freeze-drying.

Sequencing of peptides. The 'rapid Edman degradation' procedure of Gray and Smith as described by Gray (1972) was used. Subsequent dansylation and detection of derivatives on polyamide layers were as described by Redman (1975).

Partial acidic hydrolysis. Hydrolysis with 0.03 M-HCl was under N_2 at 105°C for 17 h, before freeze-drying.

Carboxypeptidase-C digestion. This was performed as described by Redman (1975).

Results and Discussion

Separation and analyses of α -chymotryptic peptides

Limited reaction during detection of peptides with fluorescamine gave spots that could be eluted from paper for direct amino acid analysis and sequencing.

The peptide 'maps' of the α -chymotryptic digests of peptides CNBr-4 showed five major spots (C1–C5), one weaker spot (C6) and several very weak ones. The positions of peptides C1–C6 are shown in Fig. 1. Peptide 'maps' of α -chymotryptic digests of CNBr-4 peptides from proteins S, M and F gave no indication that the peptides differed in structure. In confirmation of this, digests of approximately equal weight mixtures of peptides S-CNBr-4 and F-CNBr-4, and also of peptides S-CNBr-4 and M-CNBr-4, gave no spots outside the normal pattern.

Fig. 1. 'Map' of the α -chymotryptic peptides from peptides CNBr-4

Lys, Ser and Asp were included as markers on this occasion only. Electrophoresis at pH 3.5 (water/acetic acid/pyridine, 100:10:1, by vol.) was followed by chromatography in butan-1-ol/acetic acid/water (3:1:1, by vol.). Peptides were located with fluorescamine. For further details see the text.

Amino acid analyses of eluted materials are in Supplementary Publication SUP 50063.

Sequencing of α -chymotryptic peptides from peptide F-CNBr-4

The proposed sequences are shown in Fig. 2, and are based on the following evidence.

Peptide C1 (residues 1–4). The sequence of residues 1–3 was determined by dansyl-Edman degradation. Residue 4 was identified after carboxypeptidase-C treatment of peptide C1 and isolation of the products from peptide 'maps'. These were found to be Ser-Gly-Pro (dansyl-Edman degradation) and Trp (amino acid analysis).

The assignments were supported by amino acid analyses of peptides C1 from both peptides S-CNBr-4 and M-CNBr-4, and also carboxypeptidase-C treatment of peptide C1 from peptide S-CNBr-4.

Peptide C2 (residues 5–13). The sequence of residues 5–8 (6 = unidentified, 8 = Asx) was determined by dansyl-Edman degradation. Subtractive Edman degradation of peptide C2 from peptide S-CNBr-4 proved residues 5–7 (6 = Trp) (see Supplementary Publication 50063). Amino acid analyses of peptides C2 from both peptides S-CNBr-4 and M-CNBr-4 were consistent with the proposed composition.

The major part of the evidence for the sequence of peptide C2 from peptide F-CNBr-4 resulted from characterizing products of its hydrolysis in 0.03 M-HCl. Three major spots were obtained on a peptide 'map', one analysing as only aspartic acid (residue 8). The other two, peptides C2 α and C2 β , appeared on peptide 'maps' at similar positions to peptides C2 and C1 respectively, under the same conditions.

The sequences of peptide C2 α (residues 5–7; 6 = unidentified) and of residues 9–12 of peptide C2 β (residues 9–13) were both determined by dansyl-Edman degradation.

These same sequences were found in peptides C2 α and C2 β respectively from peptide S-CNBr-4 (dansyl-Edman degradation).

Amino acid analysis of peptide C2 β from peptide

Fig. 2. Amino acid sequence of peptide F-CNBr-4

Residues 6, 13 and 24 were determined partly by analogy with peptides obtained from peptides S-CNBr-4 and M-CNBr-4; for details see the text. The assembly of the α -chymotryptic peptides C1, C2, C3 and C4 is described in the text and in Supplementary Publication 50063. Peptide F-CNBr-4 constitutes the N-terminal 24 residues of the protein F, where CmCys (S-carboxymethylcysteine) and Hse (homoserine) residues are derived from Cys and Met residues respectively.

M-CNBr-4 was in agreement with the composition given in Fig. 2 (i.e. including Tyr) (see Supplementary Publication 50063).

Therefore, as expected from observations on other peptide chains (Schultz, 1967), hydrolysis in 0.03 M-HCl cleaved the C2 peptide chain on both sides of Asp (residue 8), yielding peptides C2 α and C2 β , and the evidence is sufficient to prove the sequence of peptide C2 to be that shown in Fig. 2 (8 = Asx).

Peptide C2 (prepared from a mixture of peptides S-CNBr-4 and M-CNBr-4) had a single electrophoretic mobility at pH 6.5 of $m = +0.48$ relative to aspartic acid, corrected for endo-osmosis. Calculation by the method of Offord (1966) predicts a mobility of $+0.47$ (for Asx = Asp) or $+0.26$ (for Asx = Asn), the former assignment evidently being the correct one.

Peptide C3 (residues 14–18). The sequence of residues 14–18 was determined by dansyl-Edman degradation.

Amino acid analysis of peptide C3 from peptide M-CNBr-4 supported the assignments.

Peptide C4 (residues 19–24). The sequence of residues 19–23 was determined by dansyl-Edman degradation. Amino acid analysis of peptide C4 from peptide M-CNBr-4 supported the assignments, and also located the homoserine residue known to be C-terminal in the CNBr-4 peptides (Redman, 1975).

Amino acid analysis showed that peptide C5 from peptide M-CNBr-4 was probably identical with peptide C4 except that the homoserine residue at the C-terminus was lactonized.

The detection of tryptophan, tyrosine and leucine as C-termini of peptides C1, C2, and C3 respectively is in accord with the known specificity of α -chymotrypsin (Neil *et al.*, 1966). The Trp-CmCys bond between residues 6 and 7 was relatively resistant, but the minor peptide C6 is thought to be residues 7–13 (see Supplementary Publication 50063).

Therefore peptides C1, C2, C3 and C4 from peptide F-CNBr-4 have been sequenced, the only positions obtained by analogy with peptides from other proteins being residues 6, 13, and 24. The location of Trp at position 6 is suggested by an unidentified 'blank' in the dansyl-Edman method, and the homoserine at position 24 is strongly supported by evidence on the whole peptide. Also, with the inclusion of these three residues, the sums of the amino acid residues of peptides C1, C2, C3 and C4 account exactly for the integral amino acid composition of peptide F-CNBr-4 (Redman, 1975), the presence of two tryptophan residues having been confirmed. Therefore assembly of these chymotryptic peptides completes the sequence.

Assembly of the chymotryptic peptides

Dansyl-Edman degradation of peptide S-CNBr-4 showed the three N-terminal amino acids to be Ser-Gly-Pro. This evidence places peptide C1 at the N-terminus of peptide CNBr-4.

Peptide C4, which contains the homoserine residue, is evidently the C-terminal peptide.

The sequence determination was completed after hydrolysis of peptides CNBr-4 preferentially at certain peptide bonds with 70% (v/v) formic acid and also with 0.03 M-HCl. Both treatments split the Asp-Pro bond between residues 8 and 9 (Fig. 2), but only in the latter case was the CmCys-Asp bond between residues 7 and 8 broken. Amino acid analysis of products obtained from peptide 'maps' showed the final assembly of the chymotryptic peptides to be that shown in Fig. 2. (Details are given in Supplementary Publication 50063, which also contains extra confirmatory evidence for the sequence.)

Conclusion

In view of the identical integral amino acid compositions of peptides S-CNBr-4, M-CNBr-4 and F-CNBr-4 (Redman, 1975) and apparent identity of chymotryptic peptide 'maps' (no peptide with more than nine residues), it is highly probable that these peptides have identical sequences. Comparative experiments and studies on mixtures of equivalent peptides derived from different albumins, described above, have supported this hypothesis. Therefore families of wheat proteins, each containing identical N-terminal amino acid sequences, occur in albumins, in addition to the previously reported observations on gliadins (Kasarda *et al.*, 1974; Patey *et al.*, 1975), although the actual sequences are entirely different in the two classes of proteins.

I thank Dr. J. A. D. Ewart, who initiated this project and advised on sequencing methods, and Mr. D. Smith for experimental assistance.

References

- Ewart, J. A. D. (1966) *J. Sci. Food Agric.* **17**, 30–33
- Gray, W. R. (1972) *Methods Enzymol.* **25**, 333–344
- Kasarda, D. D., Da Roza, D. A. & Ohms, J. I. (1974) *Biochim. Biophys. Acta* **351**, 290–294
- Neil, G. L., Neimann, C. & Hein, G. E. (1966) *Nature (London)* **210**, 903–907
- Offord, R. E. (1966) *Nature (London)* **211**, 591–593
- Patey, A. L., Evans, D. J., Tiplady, R., Byfield, P. G. H. & Matthews, E. W. (1975) *Lancet* **ii**, 718
- Redman, D. G. (1975) *Biochem. J.* **149**, 725–732
- Schultz, J. (1967) *Methods Enzymol.* **11**, 255–263