Reprinted from Nature, Vol. 282, No. 5738, pp. 527-529, November 29 1979

N-terminal amino acid sequencing of prolamins from wheat and related species

Jean-Claude Autran*, Ellen J.-L. Lew, Charles C. Nimmo & Donald D. Kasarda

Food Proteins Research Unit, Western Regional Research Center, Agricultural Research, SEA, US Department of Agriculture, Berkeley, California 94710

The gliadins of common bread wheat (*Triticum aestivum* L.) make up the major storage protein fraction of grain endosperm and comprise a complex mixture of proteins with similar amino acid compositions and properties[1]. Two-dimensional methods of electrophoresis[2,3], in which separations are based mainly on net charge, separate gliadins of a single wheat variety into 30–46 components and there is considerable variation among the electrophoretic patterns of different wheat varieties[1]. The gliadins seem to be encoded by clusters of duplicated genes[2] and may constitute a multigene family[4]. The fact that several purified gliadin components have closely similar N-terminal amino acid sequences[5,6] led us to consider the possibility that, despite the complexity of the mixture, gliadins have sufficient homology in their N-terminal sequences to allow estimation of the homology by sequencing the unfractionated mixture. Here, we provide evidence for two major groups of gliadins in wheat based on N-terminal sequencing, show that similar groups of prolamins (a general term for equivalent proteins in other species) are found in related species of *Triticum* and *Aegilops* that may have contributed genomes to polyploid wheats with *ABD* genome composition, and show that only one of these groups is found in rye (*Secale cereale* L.).

Wheat and other grain samples used are listed in Table 1. The prolamins of common wheat, *Triticum monococcum*, and rye were prepared from endosperm only (flour) whereas those of other samples, of which we had only small amounts, were prepared from de-germed grain that was ground in a Wiley mill. Prolamins were prepared according to the procedure of Charbonnier[7] except that the final product was dissolved in 0.01 M acetic acid or 60% ethanol-water and dialysed thoroughly against distilled water. The small amount of precipitate that formed (5% or less of the amount dissolved) was removed by centrifugation and the clear supernatant was freeze-dried. Rye prolamins were less soluble in water—about two-thirds of the preparation precipitated—and both the precipitate and the supernatant were freeze-dried separately for use in sequencing. Yields of prolamins ranged from 3 to 8% of starting material. Electrophoretic patterns of typical preparations, which illustrate the complexity of the mixtures, are shown in Fig. 1.

Automatic sequencing[8] was carried out on a Beckman sequencer model 890B with DMAA peptide program 111374. Eight cycles were considered sufficient for our purposes. (One sample, *T. monococcum*, was sequenced only through six cycles.) Residue identification was by gas chromatography (GC) according to the procedure of Pisano *et al.*[9]. In all cases, only a few phenylthiohydantoin (PTH) amino acids (sometimes only one) were identified at each cycle of the Edman degradation and recoveries were sufficient to account for 70–80% of the protein sample when moisture contents (10%), probable sequencer yields (65% assumed[10]) and prolamin molecular weights (36,000 assumed[1]) were taken into account. Typical results are given in Table 2.

The major PTH amino acids at each cycle usually corresponded to the sequence Val-(Arg?)-Val-Pro-Val-Pro-Gln-Leu- for all species of *Triticum* and *Aegilops* except for the wheat variety 'Justin' and an accession of *Triticum urartu* (G-1876), where it was the minor sequence. This sequence corresponds to that of α-gliadins[5,6]; we shall refer to it as the α-type sequence, although Bietz *et al.*[6] found this sequence also in a β-gliadin and a γ-gliadin (γ_1-gliadin). Arg and His were not determined by our GC method. We assume Arg to be an important amino acid at position 2 by analogy with the α-gliadin sequence[5,6] and identified it at cycle 2 in two samples by liquid chromatography.

Table 1 Ratio of α-type to γ-type sequence from cycle 4

Species and genome designations	α/γ Ratio
T. aestivum L. em. Thell., 'Scout 66' (*ABD*)	1.6
T. aestivum L. em. Thell., 'Justin' (*ABD*)	0.9
T. dicoccoides Körn., 64–1148 (*AB*)	1.1
T. dicoccum Schubl., Khapli emmer, 63–2356 (*AB*)	1.3
T. boeoticum Boiss. em. Schiem., G–2512 (*A*?)	2.2
T. monococcum L., Univ. Manitoba (*A*?)	1.3
T. urartu Tum., G–3151 (*A*?, *B*?)	1.1
T. urartu Tum., G–1876 (*A*?, *B*?)	0.7
Ae. speltoides Tausch, U. C., Davis (*B*?)	1.4
Ae. squarrosa L., 0–623 (*D*)	1.9
Ae squarrosa L., 64–1154 (*D*)	1.9
Secale cereale L., 'Frontier'	—

A second sequence corresponding to Asn-(Met/Ile)-Gln-Val-(Val/Asp)-Pro-Gln-Gly- was present in all the samples of *Triticum* and *Aegilops*. This sequence is similar to the γ_2- and γ_3-gliadin sequences described by Bietz *et al.*[6] (Table 2); we shall refer to it as the γ-type sequence. It was the minor sequence in most samples except for 'Justin' and *T. urartu* G-1876, where it was present in amounts about equal to (or greater than) the amount of α-type sequence. *T. urartu* G-1876, however, had only a trace amount of Asn at cycle 1; the major amino acid was Val (40 nm) followed by Ala (17 nm). We usually found both Met and Ile at cycle 2, but our sample of *T. monococcum* showed only Ile at cycle 2 whereas our sample of *Aegilops speltoides* showed only Met at cycle 2.

On the basis of N-terminal sequences, our results indicate two major homologous groups of prolamins in the species of *Triticum* and *Aegilops* we examined. Because these species were fairly representative of the two genera, this may be true in general. Based on sequencer yields, we estimate that these two groups account for about 70–80% of our prolamin preparations. A fairly clear difference between the α-type and γ-type sequences occurs at residue 4, where Pro is found in the former and Val or Gln in the latter (Table 2). Pro and Val were the only two amino acids we identified at cycle 4 for all species of *Triticum* and *Aegilops* with the exception of one accession of *T. urartu*

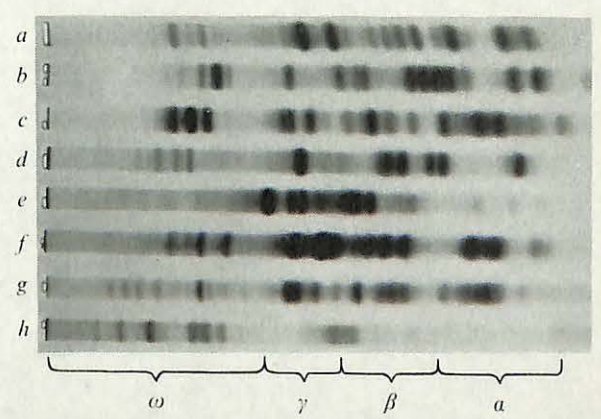

Fig. 1 Polyacrylamide gel electrophoresis[16] of prolamins from different species, aluminium lactate buffer, pH 3.2, migration from left (+) to right (−): *a*, *T. monococcum*; *b*, *T. boeoticum* (G-2512); *c*, *T. urartu* (G-3151); *d*, *Ae. squarrosa* (0-623); *e*, *Ae. speltoides*; *f*, *T. dicoccoides* (64-1148); *g*, *T. aestivum* ('Scout 66'); *h*, *S. cereale* ('Frontier'). The α-, β-, γ- and ω-regions correspond to the usual assignments of electrophoretic mobilities for gliadin patterns of common wheats[1].

* Permanent address: Institut National de la Recherche Agronomique, Laboratoire de Technologie des Blés Durs et du Riz, 34060 Montpellier, France.

Table 2 Recovery of PTH amino acids for different species (nmol)

Species	Cycle of Edman degradation 1	2	3	4	5	6	7	8
T. aestivum 'Scout 66'	Val (97) Asn (37)	* Met (40) Ile (22)	Val (65) Gln (54) Leu (12)	Pro (75) Val (47)	Val (69) Asp (28)	Pro (91)	Gln (71) Ser (50)	Leu (64) Gly (37)
T. aestivum 'Justin'	Val (55) Asn (51)	* Met (42) Ile (40)	Val (50) Gln (39) Leu (32)	Pro (53) Val (63)	Val (41) Asp (24)	Pro (104)	Ser (56) Gln (20)	Leu (46) Gly (42)
T. dicoccoides 64–1148	Val (64) Asn (33)	* Met (33) Ile (29)	Val (63) Gln (29) Leu (21) Phe (5)	Pro (76) Val (72)	Val (68) Asp (16)	Pro (109)	Ser (70) Gln (25)	Leu (59) Gly (41)
Ae. squarrosa 64–1154	Val (119) Asn (26) Ala (12)	* Ile (39) Met (17)	Val (91) Gln (25)	Pro (94) Val (47)	Val (53) Asp (9)	Pro (99)	Gln (43)	Leu (39) Gly (10)
S. cereale 'Frontier' (soluble fraction)	Asn (59) Ala (24) Leu (20)	Met (97) Gln (87)	Gln (95) Leu (78)	Val (35) Asn (15)	Pro (48) Gly (41) Asn (15)	Pro (59) Ser (57)	Ser (73) Gln (14)	Gly (53) Gln (43)
α-gliadin[5,6]	Val	Arg	Val	Pro	Val	Pro	Gln	Leu
γ_2-gliadin[6]	Asn Pro	Ile	Gly Gln	Val	Asp Val Gln	Pro Gln	Trp	Gly Leu
γ_3-gliadin[6]	Asn Pro	Met	Gly Gln	Val Gln	Asp Val	Pro Gln	Trp Gln	Gly

Recoveries are reported on a basis of 10-mg samples.

* Arg is likely to be an important residue at position 2, but Arg and His were not determined by our GC method: in addition, the quantitation of Asn and Gln was less accurate than that of the other amino acids.

(G-1876), which also had a small amount of Leu at this position. The ratio of Pro to Val at cycle 4 may provide a measure of the relative amounts of the two groups of proteins in a sample. These ratios are listed in Table 1, and range from 0.8 to 2.2. They sometimes differed for different varieties or samples of the same species. For example, the two hexaploid bread wheats, 'Scout 66' and 'Justin', had ratios of 1.6 and 0.8, respectively. The large amount of α-type sequence in 'Scout 66' is consistent with the large amounts of α-gliadins in this variety[11]. In contrast, 'Justin' is low in α-gliadins[11]. The presence of both α-type and γ-type prolamins in species of *Triticum* and *Aegilops* indicates the close relationship of these genera, and supports the proposal of Morris and Sears[12] that *Aegilops* be incorporated into the genus *Triticum*.

Both the water-soluble and the water-insoluble prolamins from 'Frontier' rye differed from those of the other species examined in lacking the α-type sequence, although we might not have detected a sequence corresponding to about 10% or less of the mixture. The sequences of these rye prolamin preparations were not identical to one another but both were fairly close to the γ-type found in species of *Triticum*[6] and *Aegilops*. Although minor amino acids were identified at most cycles of the rye prolamin degradation, these did not correspond to known gliadin sequences. The electrophoretic patterns of rye prolamins at *p*H 3 vary somewhat according to variety[13], but they do not show a significant amount of staining in the region of mobility corresponding to α-gliadins[13]. As we have noted above, however, the α-type of sequence has been found for proteins with other electrophoretic mobilities[6], so the agreement between our sequencing results and the electrophoretic patterns may not be essential.

The presence of only the γ-type sequence in our sample of rye prolamin may indicate that rye is closer evolutionarily than *Triticum* and *Aegilops* to the common ancestor that gave rise to the sub-tribe Triticinae[14], which includes both wheat and rye, within the family Gramineae. Differentiation may have included the development and amplification of genes coding for α-type prolamins in *Triticum* and *Aegilops*. This speculation is based on results for only one variety of rye; further work on additional varieties and types is needed to allow its generalisation to the species as a whole.

The great homology among the mixtures of prolamins we have studied enables considerable sequencing information to be obtained from them without isolating single protein components. This has also been demonstrated[15] for the zein proteins, the prolamins of maize (*Zea mays*). This homology is notable in consideration of the complexity of the mixtures. The prolamins of species of *Triticum* and *Aegilops* consist of many components that may differ in molecular weight as well as in charge; prolamin preparations from various species of these genera were resolved into at least four (usually more) bands by SDS electrophoresis (E. C. Cole, J. G. Fullington and D.D.K., unpublished results). Small differences in amino acid composition have been found for gliadin components that differ in electrophoretic mobility (aluminium lactate buffer, *p*H 3.2)[1], and gliadin proteins corresponding to particular bands of the *p*H 3.2 patterns exhibited simple inheritance in crosses between wheat varieties that differed in their electrophoretic patterns[2]. Regardless of whether the complexity of gliadins and other prolamin mixtures results from multiple genes, post-translational modifications or some other mechanism, the high degree of homology we found for their N-terminal sequences suggests that this region has some important function and thus tends to be conserved.

We thank B. L. Johnson, J. G. Waines, P. Mattern, V. A. Johnson, C. O. Qualset, J. Dvorak, B. L. Jones, and R. Bieter for providing us with grain samples. J.-C.A. is a Visiting Scientist. Reference to a company and/or product name by the Department is only for information and does not imply approval or recommendation to the exclusion of others that may also be suitable.

Received 2 July; accepted 20 September 1979.

1. Kasarda, D. D., Bernardin, J. E. & Nimmo, C. C. in *Advances in Cereal Science and Technology* Vol. 1 (ed. Pomeranz, Y.) 158–236 (American Association of Cereal Chemists, St Paul, Minnesota, 1976).
2. Mecham, D. K., Kasarda, D. D. & Qualset, C. O. *Biochem. Genet.* **16,** 831–853 (1978).
3. Wrigley, C. W. & Shepherd, K. W. *Ann. N. Y. Acad. Sci.* **209,** 154–162 (1973).
4. Hood, L. *Fedn Proc.* **35,** 2158–2167 (1976).
5. Kasarda, D. D., Da Roza, D. A. & Ohms, J. I. *Biochim. biophys. Acta* **351,** 290–294 (1974).
6. Bietz, J. A., Huebner, F. R., Sanderson, J. E. & Wall, J. S. *Cereal Chem.* **54,** 1070–1083 (1977).
7. Charbonnier, L. *C.r. hebd. Séanc. Acad. Sci., Paris* **271** Ser. D, 2042–2045 (1970).
8. Edman, P. & Begg, G. *Eur J. Biochem.* **1,** 80–91 (1967).
9. Pisano, J. J., Bronzert, T. J. & Brewer, H. B. Jr *Analyt. Biochem.* **45,** 43–59 (1972).
10. Niall, H. in *The Proteins,* Vol. III (ed. Neurath, H.) 179–238 (Academic, New York, 1977).
11. Platt, S. G., Kasarda, D. D. & Qualset, C. O. *J. Sci. Fd Agric.* **25,** 1555–1561 (1974).
12. Morris, R. & Sears, E. R. in *Wheat and Wheat Improvement* (eds Quisenberry, K. S. & Reitz, L. P.) 19–87 (American Society of Agronomy, Madison, Wisconsin, 1967).
13. Konarev, V. G., Gavrilyuk, I. P., Gubareva, N. K. & Peneva, T. I. *Cereal Chem.* **56,** 272–278 (1979).
14. Smith, D. B. & Flavell, R. B. *Biochem. Genet.* **12,** 243–256 (1974).
15. Bietz, J. A., Paulis, J. W. & Wall, J. S. *Cereal Chem.* **56,** 327–332 (1979).
16. Kasarda, D. D., Bernardin, J. E. & Qualset, C. O. *Proc. natn. Acad. Sci. U.S.A.* **73,** 3646–3650 (1976).

Printed in Great Britain by Macmillan Production Ltd., Basingstoke, Hampshire