NOTE

N-Terminal Amino Acid Sequences Show that D Hordein of Barley and High Molecular Weight (HMW) Secalins of Rye Are Homologous with HMW Glutenin Subunits of Wheat

P. R. SHEWRY,[1] A. S. TATHAM,[1] D. J. PAPPIN,[2] and J. KEEN[2]

The alcohol-soluble storage proteins of barley (hordeins), wheat (gliadins and glutenin subunits), and rye (secalins) can be divided into three groups on the basis of their amino acid compositions and the chromosomal locations of their structural genes (Shewry et al 1984c). The validity of two of these groups has been confirmed by comparison of partial and complete amino acid sequences. The first comprises C hordein, ω-gliadins, and ω-secalins, and the second B and γ-type hordeins, γ-secalins (M_r 40,000 and M_r 75,000), α-type and γ-type gliadins, and at least the major low molecular weight subunits of glutenin (Shewry et al 1986). The third group comprises the high molecular weight (HMW) subunits of wheat glutenin, HMW secalins of rye, and D hordein of barley. Although the HMW subunits of wheat glutenin have been studied in detail (Shewry et al 1984b, Field et al 1987, Halford et al 1987) because of their proposed role in determining breadmaking quality (Payne et al 1984), only preliminary characterization of D hordein and HMW secalins has been reported (Field et al 1982, Kreis et al 1984).

In order to determine partial amino acid sequences of these two groups of proteins, we purified fractions from milled grain of the mutant barley line Risø 1508 and rye cultivar Gazelle. D hordein was prepared as described by Kreis et al (1984). Secalins were extracted with 70% (v/v) aqueous (aq) ethanol followed by 50% (v/v) aq propan-1-ol containing 2% (v/v) 2-mercaptoethanol and 1% (v/v) acetic acid. The supernatant from the second extraction was mixed with 2 vols of 1.5*M* aq NaCl and stood overnight at 4° C to precipitate prolamins, which were then reduced and pyridylethylated (Friedman et al 1970). HMW secalins were purified by ion-exchange chromatography on CM cellulose at pH 4.6 (Shewry et al 1984b) followed by gel filtration on Sephadex G75 in 0.1*M* acetic acid.

Whereas the D hordein fraction gave one predominant band when separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), several bands were present in the HMW secalin fraction (Fig. 1). This is because rye is an outbreeding species, and cultivars are mixtures of genotypes rather than pure lines (Shewry et al 1983, 1984a).

The N-terminal amino acid sequences of the two fractions were determined using automated solid phase Edman degradation (Laursen 1971), with identification of PTH amino acids by reversed-phase high-performance liquid chromatography (see Shewry et al 1987). Use of a two-stage procedure allowed the determination of the first 40 residues of each fraction. In this procedure the protein was initially sequenced for 20 (HMW secalin) or 30 (D hordein) cycles, and the analyses were then repeated for 40 cycles with the coupling times increased for the amides and acidic residues and the cleavage times increased for the proline residues identified in the first run. The successful analysis of the D hordein fraction contrasts with previous attempts, when it appeared to be N-terminally blocked (Field et al 1982). We do not know the reason for this difference.

The sequences are aligned with those of wheat HMW glutenin subunits in Figure 2. Hexaploid bread wheat (*Triticum aestivum*) has six HMW subunit genes, two each on chromosomes 1A, 1B, and 1D (Harberd et al 1986). They are of two types, called x and y, which appear to have arisen from a gene duplication that predated the divergence of the progenitors of the three genomes (Payne et al 1981). Not all of these genes are expressed, genotypes of bread wheat containing 3, 4, or 5 subunits, which are numbered in order of increasing mobility on SDS-PAGE (Payne et al 1981). Figure 1 includes data for x-type and y-type subunits encoded by all three genomes of bread wheat and for x-type subunits from diploid species thought to be related to the progenitors of the A and D genomes (*T. monococcum* and *T. tauschii* [*Aegilops squarrosa*], respectively). All of the bread wheat sequences with the exception of 1Ax1 are deduced from genomic DNA, the 1Ay sequence being deduced from a silent (i.e., not expressed) gene.

Although the HMW secalin fraction was a mixture of proteins, sequence heterogeneity was only observed at one position (glycine

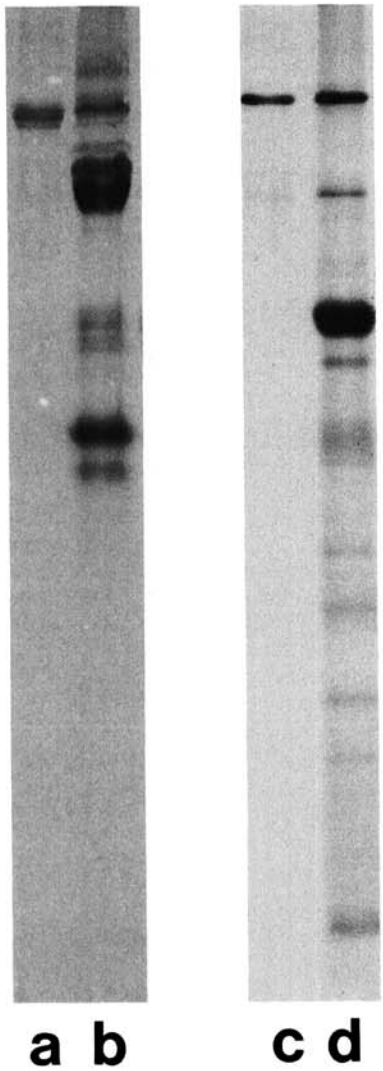

Fig. 1. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis of **a,** the purified high molecular weight secalin fraction; **b,** total secalins from rye cultivar Gazelle; **c,** the purified D hordein fraction; and **d,** total hordeins from the mutant barley line Risø 1508.

[1]AFRC Institute of Arable Crops Research, Rothamsted Experimental Station, Harpenden, Herts. AL5 2JQ, U.K.
[2]Department of Biochemistry, University of Leeds, Leeds LS2 9JT, U.K.

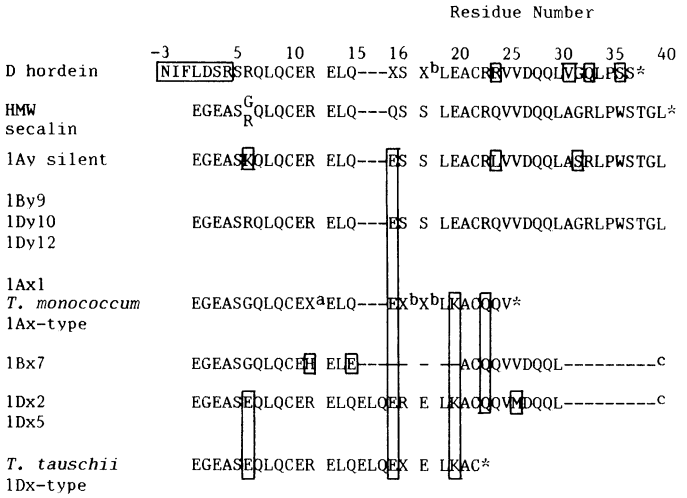

Fig. 2. Alignment of the N-terminal amino acid sequences determined for D hordein and high molecular weight (HMW) secalins with those reported for HMW subunits of wheat glutenin. The wheat proteins are referred to by the chromosomal locations of their structural genes (1A, 1B, or 1D), their type (x or y), and their positions on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (1, 2, 5, 7, 9, 10, 12, in order of increasing mobility) following the nomenclature of Payne et al (1981). The wheat sequences are from Shewry et al 1984b (1Ax1, 1Dx from *T. tauschii*), Forde et al 1985 (silent 1Ay), Sugiyama et al 1985 (1Dx2), Thompson et al 1985 (1Dy12), Halford et al 1987 (1By9), and F. C. Greene et al (*unpublished*) (1Dx5, 1Dy10). The 1By7 sequence is based on the partial protein sequence of Shewry et al (1984b) and the unpublished genomic sequence of Anderson and Greene (*personal communication*). Standard single letter abbreviations are used: alanine (**A**), cysteine (**C**), aspartic acid (**D**), glutamic acid (**E**), phenylalanine (**F**), glycine (**G**), histidine (**H**), lysine (**K**), leucine (**L**), asparagine (**N**), proline (**P**), glutamine (**Q**), arginine (**R**), serine (**S**), threonine (**T**), valine (**V**), tryptophan (**W**), unidentified (**X**). Letters followed by superscripts a or b may correspond to arginine or serine, respectively, which are often difficult to identify positively. Superscript c indicates deletion corresponding to residues 31–48 of the y-type subunits. An asterisk indicates the last residue determined by automated Edman degradation. The boxes enclose sequences in the wheat and barley proteins that differ from those in the HMW secalin.

and arginine at position 6). This indicates that the component proteins are highly homologous. Apart from this heterogeneity, the HMW secalin sequence differs at only one position from those of the 1By and 1Dy subunits of bread wheat and at two additional positions from the protein encoded by the silent 1Ay gene. In contrast, the 1Bx and 1Dx subunits differ in the insertion and/or deletion of blocks of residues. The additional tripeptides present after residues 15 of subunits 1Dx2 and 1Dx5 are identical at the protein (Glu.Leu.Gln) and codon (GAG.CTC.CAG) levels to residues 13 to 15 of the same proteins and may have resulted from a duplication of this region. The directly determined sequences of the 1Ax subunits are too short to make meaningful comparisons.

The D hordein sequence can also be aligned with those of the HMW glutenin subunits, although it differs at the N-terminus where the conserved Glu.Gly.Glu.Ala sequence is replaced by an unrelated block of seven residues. The sequence is again most closely related to those of the y-type subunits of bread wheat, without the insertions or deletions that are present in the x-type proteins. Apart from the seven residues at the N-terminus, it differs from the HMW secalin at five out of the 31 identified positions (including position 6 of the HMW secalin), and from the 1By and 1Dy subunits at six positions.

The sequences reported here confirm the homology of the HMW secalins, HMW subunits of glutenin, and D hordein. They also show that the N-terminal regions of the barley and rye components are most closely related to the y-type HMW glutenin subunits encoded by the B and D genomes. Although this suggests that the y-type subunits may be more closely related to the ancestral HMW subunit type than are x-type subunits, more extensive sequence analyses are required to establish this conclusively.

LITERATURE CITED

FIELD, J. M., SHEWRY, P. R., MIFLIN, B. J., and MARCH, J. F. 1982. The purification and characterization of homologous high molecular weight storage proteins from grain of wheat, rye and barley. Theor. Appl. Genet. 62:329.

FIELD, J. M., TATHAM, A. S., and SHEWRY, P. R. 1987. The structure of a high molecular weight subunit of wheat gluten. Biochem. J. 247:215.

FORDE, J., MALPICA, J.-M., HALFORD, N. G., SHEWRY, P. R., ANDERSON, O. D., GREENE, F. C., and MIFLIN, B. J. 1985. The nucleotide sequence of a HMW glutenin subunit gene located on chromosome 1A of wheat (*Triticum aestivum* L.). Nucl. Acids Res. 13:6817.

FRIEDMAN, M., KRULL, L. H., and CAVINS, J. F. 1970. The chromatographic determination of cystine and cysteine residues as S-β-(4-pyridylethyl) cysteine. J. Biol. Chem. 245:3868.

HALFORD, N. G., FORDE, J., ANDERSON, O. D., GREENE, F. C., and SHEWRY, P. R. 1987. The nucleotide and deduced amino acid sequences of an HMW glutenin subunit gene from chromosome 1B of bread wheat (*Triticum aestivum* L.) and comparison with those of genes from chromosome 1A and 1D. Theor. Appl. Genet. 75:117.

HARBERD, N. P., BARTELS, D., and THOMPSON, R. D. 1986. DNA restriction-fragment variation in the gene family encoding high molecular (HMW) glutenin subunits of wheat. Biochem. Genet. 24:579.

KREIS, M., SHEWRY, P. R., FORDE, B. G., RAHMAN, S., BAHRAMIAN, M. B., and MIFLIN, B. J. 1984. Molecular analysis of the effects of the mutant *lys 3a* gene on the expression of *Hor* loci in developing endosperms of barley (*Hordeum vulgare*). Biochem. Genet. 22:231.

LAURSEN, R. A. 1971. Solid phase Edman degradation. An automatic peptide sequencer. Eur. J. Biochem. 20:89.

PAYNE, P. I., HOLT, L. M., and LAW, C. N. 1981. Structural and genetical studies of the high molecular weight subunits of wheat glutenin: I. Allelic variation in subunits amongst varieties of wheat (*T. aestivum* L.). Theor. Appl. Genet. 60:229.

PAYNE, P. I., HOLT, L. M., JACKSON, E. A., and LAW, C. N. 1984. Wheat storage proteins: Their genetics and their potential for manipulation by plant breeding. Philos. Trans. R. Soc. London Ser. B 304:359.

SHEWRY, P. R., PARMAR, S., and MIFLIN, B. J. 1983. The extraction, separation and polymorphism of the prolamin storage proteins (secalins) of rye. Cereal Chem. 60:1.

SHEWRY, P. R., BRADBERRY, D., FRANKLIN, J., and WHITE, R. D. 1984a. The chromosomal locations and linkage relationships of the structural genes for the prolamin storage proteins (secalins) of rye. Theor. Appl. Genet. 69:63.

SHEWRY, P. R., FIELD, J. M., FAULKS, A. J., PARMAR, S., MIFLIN, B. J., DIETLER, M. D., LEW, E. J-L., and KASARDA, D. D. 1984b. Purification and N-terminal amino acid sequence analysis of high molecular weight (HMW) gluten polypeptides of wheat. Biochim. Biophys. Acta 788:23.

SHEWRY, P. R., MIFLIN, B. J., and KASARDA, D. D. 1984c. The structural and evolutionary relationships of the prolamin storage proteins of barley, rye and wheat. Philos. Trans. R. Soc. London Ser. B. 304:297.

SHEWRY, P. R., TATHAM, A. S., FORDE, J., KREIS, M., and MIFLIN, B. J. 1986. The classification and nomenclature of wheat gluten proteins: A reassessment. J. Cereal Sci. 4:97.

SHEWRY, P. R., PARMAR, S., and PAPPIN, D. J. C. 1987. Characterization and genetic control of the prolamins of *Haynaldia villosa*: Relationship to cultivated species of the Triticeae (rye, wheat and barley). Biochem. Genet. 25:309.

SUGIYAMA, T., RAFALSKI, A., PETERSON, D., and SOLL, D. 1985. A wheat HMW glutenin subunit gene reveals a highly repeated structure. Nucl. Acids Res. 13:8729.

THOMPSON, R. D., BARTELS, D., and HARBERD, N. P. 1985. Nucleotide sequence of a gene from chromosome 1D of wheat encoding a HMW-glutenin subunit. Nucl. Acids Res. 13:6833.

[Received May 17, 1988. Accepted August 8, 1988.]