WHEAT GLIADIN HOMOLOGY REVEALED THROUGH N-TERMINAL AMINO ACID SEQUENCE ANALYSIS[1]

J. A. BIETZ, F. R. HUEBNER, J. E. SANDERSON, and J. S. WALL, Northern Regional Research Center, Agricultural Research Service, U.S. Department of Agriculture[2], Peoria, IL 61604.

ABSTRACT Cereal Chem. 54(5) 1070-1083

Amino-terminal sequences of gliadin proteins were examined to reveal evolutionary relations, aid in classification, and relate their structures and properties. γ_{1-3}-, β_5-, and α_{8-12}-gliadins were isolated from Ponca hard red winter wheat. Disulfide bonds were reduced and alkylated, and the proteins were subjected to automated Edman degradation. Cleaved amino acids were identified by thin-layer and gas chromatography, spot tests, and amino acid analysis. α_{8-12}- and γ_1-gliadins have the sequence NH$_2$-Val-Arg-Val-Pro-Val-Pro-Gln-Leu-Gln-Pro-Gln-Asn-Pro-Ser-Gln-Gln-Gln-Pro-Gln-Glu-Gln-Val-Pro-Leu-Val-Glx-Glx-; β_5-gliadin also has this same sequence for at least 22 residues, except that residue 14 could not be identified. Thus, γ_1- and β_5-gliadins are homologous to α-gliadin, which is known to be

toxic to individuals with celiac disease. γ_2- and γ_3-gliadins each contained two or more polypeptides; the determined partial sequence of γ_2 was NH$_2$-(Asn/Pro)-Ile-(Gly/Gln)-Val-(Asp/Val/Gln) - (Pro/Gln) - Trp-(Gly/Leu)-Gln-Val-Gln-Trp-(Leu/Val)-(Pro/Gln)-Gln-; that of γ_3 was NH$_2$-(Asn/Pro)-Met-(Gly/Gln)-(Val/Gln)-(Asp/Val)- (Pro/Gln)-(Trp/Gln) - Gly-Gln-Val-Gln - (Trp/Leu)-(Val/Pro)-Pro-Gln-Gln-. Thus, γ_{2-3}-gliadins differ significantly from α_{8-12}-, β_5-, and γ_1-gliadins, but are still homologous to them; to a greater extent they resemble each other. All α-, β-, and γ-gliadins may originate from one or two ancestral genes. Gliadin classification based solely on electrophoresis fails to reveal these protein relations.

Gliadin is that fraction of wheat storage protein soluble in neutral 70% ethanol. It contains at least three types of proteins, usually referred to as high-molecular-weight (mol wt)-, omega-, and low-mol wt gliadins. Low-mol wt gliadin (hereafter referred to as gliadin), the most abundant type, itself contains at least 30 to 40 distinct electrophoretic species (1). These are usually classified as α-, β-, and γ-gliadins, in the order of decreasing electrophoretic mobility in acidic buffers. Gliadin proteins are single-chained, have similar amino acid compositions, and have mol wt near 36,500 (2). Gliadin is an important contributor to the gluten network formed in wheat-flour doughs due to its viscous nature upon hydration.

Since gliadin has no known function in wheat other than storage, the large number of these proteins suggests that some may be closely interrelated. Evidence for a common origin of gliadin proteins has been obtained by comparing peptide maps of individual gliadin proteins (3).

Gliadin is also of physiological interest, since it is toxic to individuals with celiac disease (4). α-Gliadin is particularly toxic (5), but β-gliadin and other gluten fractions also elicit immunological responses associated with celiac disease (6).

Thus, the sequences of gliadin proteins are of special interest, since they relate to gliadin's origin, toxicity, and functionality in dough. There is also interest in whether gliadin classification schemes based on size or charge are valid in terms

of protein sequence homology. To gain some answers to these questions, we chose at this time not to extensively analyze a single gliadin, as is being done by Kasarda and coworkers for α_2-gliadin (7,8), but to compare partial sequences of a large number of gliadins. This communication reports amino-terminal amino acid sequences of nine individual gliadins isolated from one wheat variety.

METHODS

Protein Isolation

Gliadin proteins were isolated from Ponca hard red winter wheat flour by sulfoethyl cellulose (SEC)chromatography followed by Sephadex G-50 gel filtration (9,10). A typical chromatographic separation (Fig. 1) reveals at least 12 peaks. Starch gel electrophoresis (SGE) demonstrated that most peaks are heterogeneous; the major component of each of the nine peaks indicated in Fig. 1 was then purified by Sephadex G-50 chromatography to homogeneity as assessed by SGE and sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)[3].

The chromatogram in Fig. 1 is similar to that obtained previously (9), except that resolution in the α-gliadin region is improved. The nomenclature used combines electrophoretic mobility (α, β, or γ) with numerical subscripts; the subscripts denote only the peak of origin, and are not related to any other system of nomenclature. The γ- and β-gliadins have been studied previously (3,9), but these αs have not been examined in detail. The α-gliadins exhibit slight

---

[3]Following completion of this study, these isolated gliadins were examined by high-resolution one-dimensional PAGE (31) by D. K. Mecham (Western Regional Research Center, USDA-ARS, Berkeley, CA), and some heterogeneity was observed. α_{8-11}-gliadins each contained, in addition to a major protein, 1–3 minor species differing slightly in electrophoretic mobility. α_{12}- and β_5-gliadins were homogeneous. γ_1-gliadin contained two major and one minor species, and γ_{2-3}-gliadins each seem to contain two proteins. The electrophoretic results for α_{12}-, β_5-, and γ_{2-3}-gliadins are consistent with the sequence data presented below; and, since single N-terminal amino acid sequences were found for α_{8-11}- and γ_1-gliadins, this additional heterogeneity strengthens our conclusions about homology among gliadin proteins.
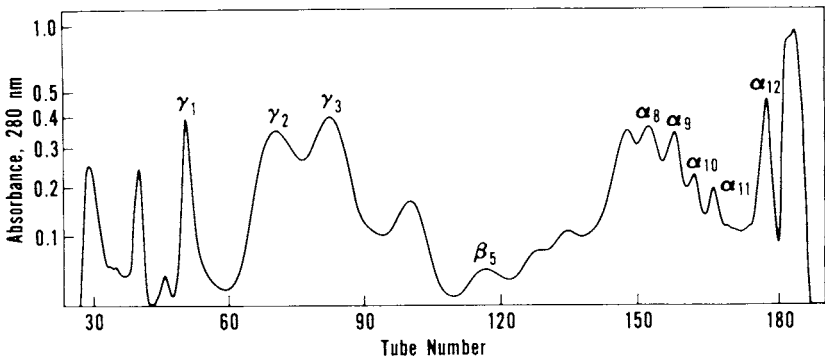

Fig. 1. Sulfoethyl cellulose chromatography of Ponca gliadin. Peaks are designated from which γ_{1-3}-, β_5-, and α_{8-12}-gliadins were isolated.

differences in electrophoretic mobility, but these differences do not necessarily correspond to their order of elution from SEC.

Protein Derivatization

Protein disulfide bonds were reduced with β-mercaptoethanol and converted to their S-carboxymethylated (CM) derivatives as described by Crestfield et al. (11), except that following alkylation, proteins were freed of excess reagents by dialysis rather than by gel filtration owing to the increased associative tendency of reduced and alkylated gliadin. S-Pyridylethyl (PE) derivatives of proteins were formed in a similar manner (12), but reduction time was decreased to 4 hr to limit the possibility of carbamylation owing to cyanate formation from urea. The 4-vinylpyridine used was freshly distilled under reduced (15 mm) pressure.

Electrophoretic Methods

SGE was performed as described by Huebner and Rothfus (10). SDS-PAGE was done at pH 8.9 (2,13) on 7.5% gels using a double-staining procedure (14). For gliadin proteins, this polyacrylamide concentration enhances resolution over that previously observed using 5% gels (2); increasing the gel concentration to 10 or 15% produced no further improvement.

Amino Acid Analysis

Proteins were hydrolyzed for 24 hr at 110°C by the method of Liu and Chang (15). Phenylthiohydantoin (PTH) amino acids were hydrolyzed at 130°C for 24 hr in ampoules sealed under vacuum, using 0.5–1.0 ml constant boiling HCl containing 0.05% β-mercaptoethanol (16). HCl was removed under vacuum over NaOH pellets, and samples were redissolved in pH 2.2 sample dilution buffer (Pierce: Rockford, IL).

Hydrolyzates were analyzed on a Beckman Model 121 amino acid analyzer (17) with L-α-amino-β-guanidino-propionic acid (Calbiochem) and norleucine as internal standards. Automatic computation of results was provided by an Infotronics integrator and an IBM 1130 computer (18).

Amino-Terminal Sequence Analyses

The automated Edman degradation procedure (19) was performed on a Beckman Model 890C sequencer using protein Quadrol programs 072172C and 122974. Sequencer-grade chemicals (Beckman and Pierce) were used throughout; dithioerythritol (Cyclo Chemical) was added (15 mg/l.) to butyl chloride to stabilize PTH-serine.

Samples (10–15 mg) dissolved in 0.4 ml 10% acetic acid were quantitatively transferred to the reaction cup and dried using sample application subroutine program 02772. Usually a preliminary cycle was run with no phenylisothiocyanate to wash and condition the sample and provide a uniform film. The degradation was then initiated, using double coupling (20) on the first cycle with program 122974. PTH-Norleucine (ca. 100 nmol) added to each sample tube served as internal standard. Usually 30 to 40 degradation cycles were run. The dried anilinothiazolinone derivatives of cleaved amino acids were converted to their PTH isomers, and PTH amino acids were extracted with ethyl acetate (16).

Repetitive yields of 93–96% were determined for apomyoglobin (Beckman);

similar repetitive yields were also obtained for isolated α- and γ_1-gliadins, based on valine recoveries at positions 1, 3, and 5. Initial yields for α-gliadin averaged 68%, based on a mol wt of 36,500 (2). These yields are similar to those usually obtained for pure proteins on a properly operating sequencer (21).

The isolated gliadins were examined as follows: γ_1 was subjected to three sequencer runs (one as the native protein and two as the CM derivative), γ_2 was analyzed four times (twice each as CM and PE derivatives), and γ_3 was analyzed five times (twice as the CM and three times as the PE derivative). The other gliadins were sequenced only as PE derivatives: β_5, α_8, α_{11}, and α_{12} were examined twice, α_9 three times, and α_{10} only once (due to lack of sample).

Examination of Sequencer Fractions

Sequencer fractions were first characterized by two-dimensional thin-layer chromatography (tlc) on 5 × 5 cm Cheng-Chin polyamide plates (Pierce) (22). BBOT fluorescent indicator was used (Packard: Downers Grove, IL). From 0.5 to 2.0% of each fraction was used, and identifications were made by comparison to a complete mixture of PTH amino acids (Pierce) on the reverse of each chromatogram; spot intensities were estimated on a numerical basis. Identifications by subsequent procedures are also greatly facilitated by this extremely rapid method.

Gas chromatography (gc) was done on a Packard 7400 series instrument using 4 ft × 2 mm coiled glass columns packed with Chromosorb W coated with 10% SP-400 (Beckman) (23). Helium flow rates were 30—60 cc/min; flame ionization detection was used. Oven temperature was increased, after a 3-min hold at 180° C, at 10° C/min to 285° C, and was held at this temperature for 7 to 12 min. Two-per cent aliquots of each ethyl acetate extract were examined as PTH and as trimethylsilyl derivatives, formed by on-column injection of equal volumes of sample and *N,O*-bis(trimethylsilyl) acetamide (Supelco, Inc.: Bellefonte, PA). Quantitation was by comparison to peak heights obtained for PTH amino acid standards, whose concentrations had been determined spectrophotometrically (19); corrections for recovery were made on the basis of the internal standard.

Aqueous phases from the ethyl acetate extraction of each fraction were examined for PTH-arginine, -histidine, and -PE cysteine. PTH-arginine was determined by the phenanthrenequinone method (21), and histidine by Pauly's diazo test (16). A new method was devised for detection of PTH-PE cysteine. Upon tlc (22), PTH-PE cysteine, synthesized from PE cysteine (Pierce) (19), had an R_f of 0.68 in solvent I (toluene:*n*-pentane:glacial acetic acid, 60:30:35 v/v/v) and 1.0 in solvent II (35% aqueous acetic acid), differentiating it from all other PTH amino acids. In practice, 2% aliquots of all aqueous phases were compared by one-dimensional tlc in solvent I to standard PTH-PE cysteine; positive and questionable results were further examined by the two-dimensional method.

Sequencer fractions were also characterized by amino acid analysis following HCl hydrolysis of PTH amino acids to free amino acids. Generally, at least one sequencer run of each protein studied was so analyzed. Amino acid analysis is especially useful for differentiating leucine and isoleucine, and for obtaining accurate quantitative data for glutamine plus glutamic acid and for asparagine plus aspartic acid.

Identification of Cleaved Amino Acids

All sequencer fractions were examined by six or seven independent tests, often in duplicate. Unambiguous sequence assignments require at least two positive identifications and essentially complete agreement among all methods. Also, quantitative data obtained by gc or amino acid analysis must show at least twice the amount of an amino acid present as was contained in the previous degradative cycle, except when a continued high yield of the previously identified amino acid, combined with no other positive identification, indicates that an amino acid occupies successive positions in the protein.

Some identifications require special consideration for gliadin proteins. Glutamine and glutamic acid may be difficult to differentiate or identify because they occur so frequently, often in sequence. Conversion of cleaved products to PTH derivatives causes partial acid hydrolysis of amide groups; thus, identifications of both glutamic acid and glutamine by tlc and gc signify glutamine, while the lack of any appreciable glutamine indicates glutamic acid. Glutamic acid may also be indicated by an increased gc recovery as compared to the previous fraction, accompanied by a decreased amount of glutamine. Asparagine and aspartic acid are differentiated and identified in an analogous manner. Arginine's identification depends on only one positive test, coupled with negative data from all other determinations. Tryptophan is easily identified by tlc but gives low gc recoveries. Serine seems especially difficult to identify: faint tlc spots or small gc peaks, a lack of any other positive data, and possible homology with other known protein sequences frequently have to suffice for an identification.

Gliadin proteins are subject to at least two general problems which interfere with their sequence analysis. Some nonspecific cleavage of peptide bonds, probably by heptafluorobutyric acid (HFBA) during the cleavage step, is inevitable during repetitive degradations. Also, freshly exposed amino-terminal glutaminyl residues may be subject to cyclization to pyroglutamyl residues under acidic conditions, such as the presence of HFBA. Pyroglutamic acid has no free amino groups, so proteins with N-terminal pyroglutamic acid are not amenable to further Edman degradation. The occurrence of three successive glutamine residues in gliadin seems to be a formidable barrier to extended sequence analyses, probably due to cyclization of N-terminal glutaminyl residues: the yield of cleaved residues decreases markedly following this sequence. For example, valine recoveries at positions 22 and 25 of α_{8-12}-, β_5-, and γ_1-gliadins were generally only about half of what would be predicted from initial repetitive yields. When, during a sequencer run, near-uniform tlc and gc results are obtained for three or more successive fractions and the yields of the major residues decrease to 10–20 nmol, the run is terminated.

RESULTS

Amino Acid Compositions

Table I presents amino acid compositions of all gliadins studied. Compositions of γ_{1-3}- and β_5-gliadins are very similar to previous determinations (3,9); compositions of isolated Ponca α_{8-12}-gliadins have not been previously reported.

The γ_1-, β_5-, and α_{8-12}-gliadins are very similar in amino acid composition (Table I). γ_{2-3}-Gliadins have compositions similar to each other, but differ significantly from all other gliadins with respect to contents of aspartic acid, proline, methionine, tyrosine, phenylalanine, and tryptophan. Like γ_1-gliadin, γ_{2-3}-gliadins are low in histidine and arginine. These data suggest a close relation among γ_1-, β_5-, and α_{8-12}gliadins. γ_{2-3}-Gliadins are similar, but differ from the other gliadins.

SDS-PAGE

Each gliadin, which was homogeneous by SGE, was also examined by SDS-PAGE (Fig. 2). Proteins are shown here in their reduced and alkylated forms, but unmodified proteins gave similar results. SGE resolves gliadins much better than does SDS-PAGE; however, SDS-PAGE may reveal unsuspected contaminants, as well as apparent molecular sizes of proteins.

All the gliadins behaved upon SDS-PAGE essentially as single components (Fig. 2), except that more streaking material in γ_3 indicated possible microheterogeneity. α_{8-12}-Gliadins had mobilities equal to the major band of whole gliadin. A slightly lower mobility for β_5-gliadin indicated increased molecular size. γ_1-Gliadin migrates at a rate between those of β_5- and α_{8-12}-gliadins; γ_2-gliadin has a mobility essentially equal to α_{8-12}-gliadins. γ_3 is the slowest migrating gliadin, indicating a higher mol wt, as was found previously (9). CM-γ_3-Gliadin has greater apparent molecular size than does PE-γ_3-gliadin (or native γ_3-gliadin) (Fig. 2, patterns 13−14); this may indicate the sensitivity of γ_3- gliadin's conformation to relatively minor structural changes, or may indicate how apparently anomalous mol wt results may be obtained through

TABLE I
Amino Acid Compositions of Ponca Gliadin Proteins

Amino Acid	Molar Percentage Recovered								
	γ_1	γ_2	γ_3	β_5	α_8	α_9	α_{10}	α_{11}	α_{12}
Aspartic acid	2.9	1.8	1.7	2.6	3.0	3.1	3.1	2.8	2.9
Threonine	1.7	2.0	2.2	1.7	1.6	1.6	1.6	1.7	1.8
Serine	5.3	4.9	4.2	5.3	5.2	5.0	4.9	5.1	4.8
Glutamic acid	41.7	39.1	39.6	39.6	37.2	37.1	38.9	36.9	36.4
Proline	15.1	18.7	18.9	15.8	15.5	16.7	15.4	15.7	16.2
Glycine	2.4	2.7	2.7	1.9	2.5	2.4	2.6	2.4	2.7
Alanine	3.4	3.0	3.2	3.5	2.9	2.7	2.9	2.8	2.9
Cysteine	1.8	1.9	2.0	1.8	1.9	1.1	1.8	2.3	1.8
Valine	3.8	3.4	3.7	4.0	4.0	3.5	3.4	4.1	4.4
Methionine	0.9	1.7	1.4	1.1	1.2	1.3	1.8	1.1	0.9
Isoleucine	3.9	3.7	3.5	3.8	4.1	4.1	4.0	3.9	4.2
Leucine	6.9	7.2	6.5	7.9	8.1	8.2	8.4	8.3	7.8
Tyrosine	3.1	0.5	0.4	2.5	3.1	3.1	3.0	3.0	3.1
Phenylalanine	3.7	5.2	5.6	3.7	3.9	4.2	3.0	3.8	3.9
Tryptophan	0.2	0.6	0.5	0.1	0.3	0.3	0.1	0.3	0.3
Lysine	0.1	0.7	0.7	0.2	0.5	0.5	0.7	0.7	0.4
Histidine	1.5	1.4	1.6	2.3	2.5	2.6	2.2	2.3	2.6
Arginine	1.6	1.5	1.6	2.3	2.4	2.5	2.3	2.6	2.9

SDS-PAGE (24). These results will be discussed further in relation to amino-terminal protein sequences.

Amino-Terminal Sequences of Gliadin Proteins

The sequence results for all gliadins examined are summarized in Table II[4].

α_8-*Gliadin*. Excellent single identifications were obtained through residue 13, except that a small amount of tryptophan at position 7 indicates microheterogeneity, or the presence of a minor gliadin with a tryptophan/glutamine substitution. No positive identification could be made at residue 14, but it is probably serine, as indicated by the lack of positive data for any other amino acid, by comparison to sequence data for the other proteins, and by the known difficulty in identifying serine. Identifications at positions 19–21 were not definite; assignment to position 20 of glutamic acid rather than glutamine was by comparison to sequence data of the other α-gliadins.

---

[4]Detailed analytical data used for determining these sequences are available from the authors upon request.

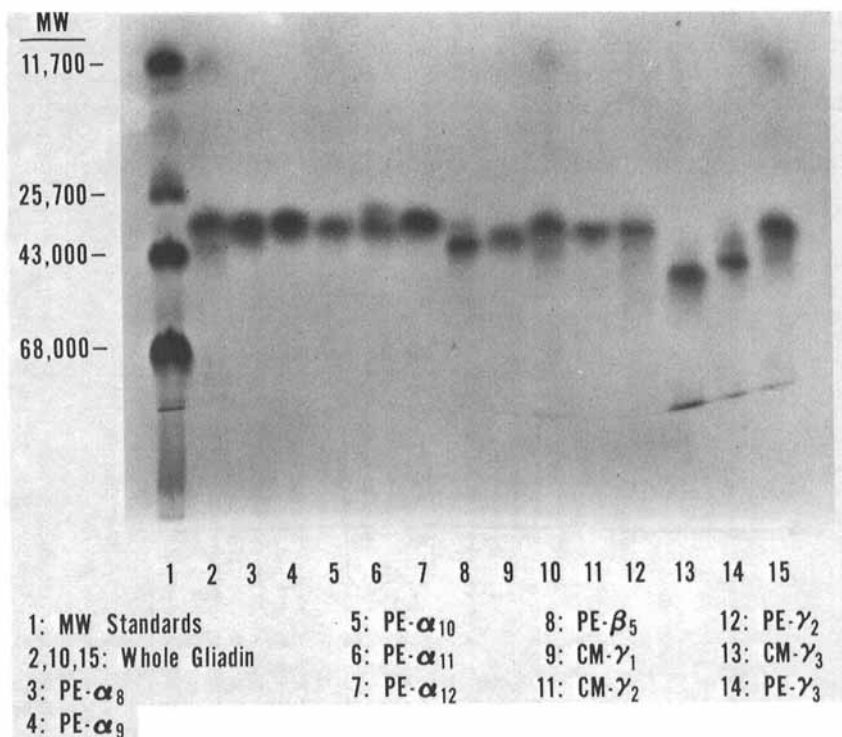|	
1: MW Standards	5: PE·α_{10} 8: PE·β_5 12: PE·γ_2
2,10,15: Whole Gliadin	6: PE·α_{11} 9: CM·γ_1 13: CM·γ_3
3: PE·α_8	7: PE·α_{12} 11: CM·γ_2 14: PE·γ_3
4: PE·α_9	

Fig. 2. SDS-PAGE of Ponca gliadin proteins. 1, mol wt standards (cytochrome c, chymotrypsinogen A, ovalbumin, and bovine serum albumin); 2, 10, and 15, whole gliadin; 3, PE-α_8; 4, PE-α_9; 5, PE-α_{10}; 6, PE-α_{11}; 7, PE-α_{12}; 8, PE-β_5; 9, CM-γ_1; 11, CM-γ_2; 12, PE-γ_2; 13, CM-γ_3; and 14, PE-γ_3.

Table III shows, in part, the gc data used to extend the sequence of α_8-gliadin beyond the 25 residues previously reported for α_2-gliadin (7).

α_9-*Gliadin.* As with α_8, results indicated a single protein sequence, except that slight tryptophan microheterogeneity was again observed at position 7. No identification was possible at position 14, but serine is proposed, as for α_8. In contrast to α_8, the glutamic acid identification at residue 20 was very good. No identifications were possible for residues 26–29, but glutamine is the probable

TABLE II
N-Terminal Sequences of Ponca Gliadins[a]

Gliadin	Residue Number						
	1	5	10	15	20	25	30
α_8	V-R-V-P-V-P-Q-L-Q-P-Q-N-P- -Q-Q-Q-P-Q̲-E̲-Q̲-V-P-L-V-Q-Z-P-Z- W						
α_9	V-R-V-P-V-P-Q-L-Q-P-Q-N-P- -Q-Q-Q-P-Q-E-Q-V-P-L-V - - - - -Q̲- W						
α_{10}	V-R-V-P-V-P-Q-L-Q-P-Q-N-P-S-Q-Q-Q-P-Q-E-Q-V-P-L-V-Z̲-Z̲- c						
α_{11}	V-R-V-P-V-P-Q-L-Q-P-Q-N-P-S̲-Q-Q-Q-P-Q-E-Q-V-P-L-V-Q̲-Z̲- W						
α_{12}	V-R-V-P-V-P-Q-L-Q-P-Q-N-P-S̲-Q-Q-Q-P-Q̲-E-Q-V-P-L-V-Q̲-Q̲- W						
β_5	V-R-V-P-V-P-Q-L-Q-P-Q-N-P- -Q-Q-Q-P-Q-E-Q-V-						
γ_1	V-R-V-P-V-P-Q-L-Q-P-Q-N-P-S̲-Q-Q-Q-P-Q-Z-Q-V-P-L-V-Z̲-Z̲-I̲-Z̲-						
γ_2	N-I-G-V-D-P-W-G-Q-V-Q-W-L-P-Q-Q-Q̲-V-P-Q-L-Z-Q̲-P̲- -Q̲- P Q V Q L I V Q r L̲ Q̲ Q Q̲ R̲ vlq Q						
γ_3	N-M-G-V-D-P-W-G-Q-V-Q-W-V-P-Q-Q̲-L-Q-p-Q̲-Q̲-Q̲- -Z̲- - -v-Q̲- P Q Q V Q Q L P Q Q̲ H q p vlq q v						
$\alpha_2{}^b$	V-R-V-P-V-P-Q-L-Q-P-Q-N-P-S-Q-Q-Q-P-Q-E-Q-V-P-L-V- a m						
γ^c	V-I-V-Q-V-R-Q-L-Q-V-Q-Q-						

[a] Data are presented using the standard one-letter notation for amino acids: A, Alanine; C, cysteine; D, aspartic acid; E, glutamic acid; G, glycine; H, histidine; I, isoleucine; L, leucine; M, methionine; N, asparagine; P, proline; Q, glutamine; R, arginine; S, serine; V, valine; W, tryptophan; and Z, glutamic acid or glutamine. More than one amino acid in a position indicates sample heterogeneity. Lower-case letters denote amino acids present in minor amounts, while underlined abbreviations indicate identifications that, although probably correct, did not totally meet the stringent requirements for unambiguous assignments as outlined previously.
[b] Isolated from Golden Seal flour and sequence determined by Kasarda *et al.* (7).
[c] Isolated from Kolibri wheat and sequence determined by Patey *et al.* (26).

amino acid at residue 30.

α_{10}-*Gliadin.* Excellent assignments were obtained through position 24. No tryptophan was noted at residue 7. However, cysteine (as PE cysteine) was identified at position 25 along with valine. Amino acids at positions 26–27 are probably glutamic acid or glutamine.

α_{11}-*Gliadin.* One sequence was obtained, except that slight tryptophan microheterogeneity at position 7 was noted. Identifications at positions 14 (serine), 26 (glutamine), and 27 (glutamine or glutamic acid) were tentative.

α_{12}-*Gliadin.* Results indicate a single amino acid sequence with slight tryptophan microheterogeneity as residue 7. Assignments of serine at position 14 and glutamine at positions 12, 26, and 27 were tentative.

β_5-*Gliadin.* One sequence was obtained through residue 22. No identification was possible at residue 14 (which, however, by comparison to the α-gliadins, may be serine).

γ_1-*Gliadin.* A single sequence was obtained. Identifications at position 14 (serine), positions 26, 27, and 29 (glutamic acid or glutamine), and position 28 (isoleucine) are tentative.

γ_2-*Gliadin.* Sequence analysis of γ_2-gliadin indicated heterogeneity not apparent by starch gel electrophoresis or chromatography: two major proteins, plus at least three minor ones, make up this fraction. In spite of this, position assignments could be made with certainty for the first 16 residues, using the same criteria described previously for sequence assignments in single polypeptide chains. In γ_2-gliadin, individual amino acids were identified at 9 of the first 24 positions.

γ_3-*Gliadin.* γ_3-Gliadin is heterogeneous to approximately the same extent as γ_2-gliadin. Definite assignments were achieved through residue 15, and 10 single assignments could be made in the first 22 positions.

DISCUSSION

Comparison of α_{8-12}-, β_5-, and γ_1-Gliadins

Automated Edman degradation reveals a high degree of homology among α_{8-12}, β_5-, and γ_1-gliadins; except for slight heterogeneity at positions 7 and 25

TABLE III
Assignment of Residues 26–29 in α_8-Gliadin

Residue	nmol in Fraction					Assignment
	V	P	L	Q	E	
22	57	26	21	7	6	V
23	28	63	22	9	5	P
24	21	34	52	7	4	L
25	38	34	39	8	2	V
26	29	19	26	26	13	Q
27	23	18	21	17	24	Z
28	17	30	24	14	17	P
29	12	24	24	17	25	Z
30	12	22	16	10	6	...

and occasional uncertainty about the identities of residues 14 and 23–27, these proteins are all identical for the first 27 residues (Table II). The differences between these proteins must be few in number and must occur well past the amino-terminal region. This explains how a mixture of α-gliadins may be successfully sequenced (8). The sequence determined for these proteins is identical, for the first 25 residues, to that of α_2-gliadin from a commercial hard red winter wheat flour and, for 10 residues, to whole α-gliadin from the variety Scout 66 (7). Our inability to significantly extend the amino-terminal sequence previously obtained for α_2-gliadin (7) demonstrates the inherent difficulty of gliadin sequence determination.

Comparison of our data to the N-terminal sequence of α_2-gliadin (7) suggests that there is relatively little sequence difference between gliadin proteins from different varieties. However, some heterogeneity was observed at positions 1 and 5 of α_2-gliadin, but not in Scout α-gliadin; α_2-gliadin was isolated from a commercial flour which may have been derived from more than one variety.[5] Thus minor differences in a gliadin sequence may occur between some varieties. Ponca α-gliadin is aggregable like α_2-gliadin (25), so this heterogeneity difference probably does not affect the ability of only certain α-gliadins to aggregate.

Comparison of γ-Gliadins

γ_{2-3}-Gliadins differ considerably in amino-terminal sequence from α_{8-12}-, β_5-, and γ_1-gliadins (Table II). Both γ_2- and γ_3-gliadin are heterogeneous, but considerable sequence homology must occur within each fraction, as revealed by the considerable number of single amino acid identifications achieved. In addition, γ_2- and γ_3-gliadin are homologous to each other: 12 of their first 22 residues are identical. Definite differences also occur, however, between γ_2- and γ_3-gliadins: γ_2 has only isoleucine at position 2, while γ_3 has only methionine. Other major differences include residues 4 and 7 (glutamine present only in γ_3), residue 5 (glutamine only in γ_2), residues 8, 13, and 21 (leucine only in γ_2), and residue 18 (glutamine and histidine only in γ_3).

The tryptophan residues at positions 7 and 12 of γ_{2-3}-gliadins are probably the only ones in the molecules: assuming 2.0 tryptophans per molecule, the data of Table I indicate approximate mol wt of 36,500 and 42,600 for γ_2- and γ_3-gliadins, respectively, in good agreement with mol wt calculated from previous amino acid compositional data (9) or by SDS-PAGE (2).

Comparison of Ponca and Kolibri γ-gliadins (Table II) suggests that major differences in γ-gliadin sequence occur between varieties. Kolibri γ-gliadin (26) is said to be a mixture of γ_1- and γ_2-gliadins with identical amino-terminal sequences; isolation data (27) suggest that it may be similar to Ponca γ_{2-3}-gliadin at 7 out of 12 positions; it is identical to Ponca γ_2-gliadin at 4 positions, partially the same at 3 positions, and different for 5 residues; and it compares to Ponca γ_3-gliadin by being identical at 3 positions, partially the same 5 times, and having 4 positions which are different. Thus, although Ponca and Kolibri γ-gliadins are homologous, large sequence differences do occur between the two varieties.

The number of γ-gliadins may also differ between varieties. Kolibri γ-gliadin

[5]D. D. Kasarda, personal communication.

is apparently homogeneous, while Ponca probably contains at least five major γ-gliadins plus additional minor ones. Chinese Spring contains two major γ-gliadins and 4 or 5 minor ones (1), while a flour milled from two wheat varieties (28) contains at least nine major γ-gliadins; thus, γ-gliadins of the two varieties may be different. Further research is necessary to reveal the relation between γ-gliadin sequence or number and the properties or quality of a variety.

Homology of All Gliadin Proteins

Although γ_{2-3}-gliadins do differ in sequence from α_{8-12}-, β_5-, and γ_1-gliadins, they also exhibit homology to this group of proteins. In the first 22 residues, γ_2-gliadin is identical to the α-gliadin sequence at four positions, partially the same at nine positions, and totally different at nine positions. γ_3-Gliadin is identical to the α-gliadins at six positions, partially the same at seven positions, and different at nine positions. Tryptophan is a major constituent at residue 7 of γ_{2-3}-gliadins, but a minor at one position 7 of some α-gliadin species; the possibility of contamination of the α-gliadins with γ-gliadins is, however, remote considering their wide resolution by ion-exchange chromatography (Fig. 1).

If these nine gliadin proteins we have examined are representative of all low-mol wt gliadins, as we believe, then all gliadins in a variety indeed are homologous. Two gliadin subgroups may be differentiated, however: one contains α-, β_5-, and γ_1-gliadins, as well as others with the same amino-terminal sequence, while the other group contains γ_{2-3}-gliadins and any other closely related proteins.

Other studies also confirm the homology and sequence differences we have observed. Peptide maps of Ponca α-, β_5-, γ_1-, and γ_3-gliadins are similar, but that of γ_3-gliadin is least like the others (3). Amino acid analyses (ref. 9 and Table I) also indicate much similarity among α-, β_5-, and γ_1-gliadins, but suggest that γ_{2-3}-gliadins differ from α_{8-12}-, β-, and γ-gliadins. γ_{2-3}-Gliadins have also been shown to differ from other low-mol wt gliadins in mol wt, as determined by amino acid analysis (9) and ultracentrifugation (29), and in precipitation behavior (28). The higher mol wt of γ_{2-3}-gliadins must result from additions to their sequences in other than their amino-terminal regions.

The sequences of γ_{2-3}-gliadins suggest that they may also have different functional properties from all other gliadin proteins, and as such are worthy of further study. Meredith (28) noted that γ_{2-3}-gliadins are a considerable part of the gluten structure, and that they may seriously modify the properties of glutenin. If so, wheat quality could be markedly affected by differences among γ-gliadins.

Origin of Gliadin Proteins

The homology we observed suggests that all low-mol wt gliadins originated from one or two common ancestral proteins through hybridization and mutation. This is supported by knowledge of wheat's evolution and of how gliadin synthesis is coded.

Wheat originated from three separate genomes, or chromosome sets, all of which had a common precursor. The diploid species *Triticum boeoticum* (AA) hybridized with a BB species (whose exact identity is not certain) to form tetraploid (AABB) wheat, similar to durum; this later crossed with the DD diploid species *Aegilops squarrosa* to produce *Tr. aestivum* (AABBDD),

common hexaploid wheat. The protein synthetic capability must have greatly increased from combining these three similar sets of genes into hexaploid wheat. Perhaps more importantly, gene duplication and nonlethal mutations at all ploidy levels have further increased the number of proteins synthesized. Both events may explain the very large number of gliadins present in wheat (30).

In Chinese Spring and presumably in other wheat varieties, γ-gliadins are coded by chromosomes 1B and 1A, and most α-gliadins are coded by group 6 chromosomes; all other gliadins are also coded by group 1 or 6 chromosomes (1,31,32). When considered with the homology we have noted, this suggests that, in wheat's ancestral species, a group 6 gene coded a protein like α_{8-12}-, β_5-, and γ_1-gliadins, while a group 1 gene coded a protein similar to γ_{2-3}-gliadins. Homology of γ_{2-3}-gliadins to the other gliadin proteins further suggests that the group 1 and 6 genes themselves arose through duplication and mutation of a single gene in a still more remote ancestor.

Electrophoretic Studies of Gliadin Proteins

There is little or no relation between the electrophoretic behavior of gliadin proteins and their sequence homology. SGE separates proteins in a gel matrix based on charge and mass differences, and is useful for identifying individual polypeptides; but it greatly separates some proteins that have now been shown to be closely related (such as α_{8-12}-, β_5-, and γ_1-gliadins) and barely separates others that are more distantly related (such as γ_{1-3}-gliadins), so it reveals little about gliadin protein relations. SDS-PAGE, which separates proteins primarily on the basis of their molecular size, also separates some proteins which are closely related (such as α-, β_5-, and γ_1-gliadins) and fails to differentiate proteins which are less closely related (such as α and γ_2) (Fig. 2). Thus electrophoretic methods do not adequately reveal gliadin homology, and all classification schemes based on electrophoresis may be useful for identification, but are largely artificial.

It is also interesting to note how closely gliadins' SGE mobilities relate to their amino acid compositions (Table I). The content of glutamic and aspartic acids in gliadins is relatively constant; furthermore, these residues exist mainly as amides. Thus, basic amino acid contents primarily determine electrophoretic mobility at acidic pH, since molecular sizes are similar. γ_{1-3}-Gliadins contain 3.6–4.4% basic residues; β_5 contains 4.9% basic amino acids; and α_{8-12}-gliadins have 5.2–5.9% basic residues. Thus, SGE electrophoretic mobility is a direct function of basic amino acid content.

Relation of Gliadin Proteins to Celiac Disease

There is considerable evidence that the toxic principle of gliadin in celiac disease is α-gliadin (5). However, other gliadin proteins may also produce an immunological response characteristic of celiac disease (6). Hemmings et al. (33) have further shown that β- and γ-gliadins behave similarly to α-gliadin in their transmission across the gut, suggesting that β- and γ-gliadins may also be toxic in celiac disease. Our data suggest that some β- and γ-gliadins may be practically identical to α-gliadin proteins. It will be interesting to see if they are as toxic as α-gliadins in celiac disease.

Literature Cited

1. WRIGLEY, C. W., and SHEPHERD, K. W. Electrofocusing of grain proteins from wheat genotypes. Ann. N.Y. Acad. Sci. 209: 154 (1973).
2. BIETZ, J. A., and WALL, J. S. Wheat gluten subunits: Molecular weights determined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Cereal Chem. 49: 416 (1972).
3. BIETZ, J. A., HUEBNER, F. R., and ROTHFUS, J. A. Chromatographic comparisons of peptic digests of individual gliadin proteins. Cereal Chem. 47: 393 (1970).
4. KASARDA, D. D. Celiac disease: Malabsorption of nutrients induced by a toxic factor in gluten. In: Protein nutritional quality of foods and feeds, Part 2: Quality factors—plant breeding, composition, processing, and antinutrients, ed. by M. Friedman, p. 565. Marcel Dekker, Inc.: New York (1975).
5. EVANS, D. J., and PATEY, A. L. Chemistry of wheat proteins and the nature of the damaging substances. Clin. Gastroenterol. 3: 199 (1974).
6. BECKWITH, A. C., and HEINER, D. C. An immunological study of wheat gluten proteins and derivatives. Arch. Biochem. Biophys. 117: 239 (1966).
7. KASARDA, D. D., DA ROZA, D. A., and OHMS, J. I. N-terminal sequence of α_2-gliadin. Biochim. Biophys. Acta 351: 290 (1974).
8. KASARDA, D. D. Wheat proteins, amino acid sequencing, and celiac disease. Proc. 9th Nat. Conf. Wheat Util. Res., U.S. Dep. Agric. ARS-NC-40: 175 (1976).
9. HUEBNER, F. R., ROTHFUS, J. A., and WALL, J. S. Isolation and chemical comparison of different gamma-gliadins from hard red winter wheat flour. Cereal Chem. 44: 221 (1967).
10. HUEBNER, F. R., and ROTHFUS, J. A. Gliadin proteins from different varieties of wheats. Cereal Chem. 45: 242 (1968).
11. CRESTFIELD, A. M., MOORE, S., and STEIN, W. M. The preparation and enzymatic hydrolysis of reduced and S-carboxymethylated proteins. J. Biol. Chem. 238: 622 (1963).
12. FRIEDMAN, M., KRULL, L. H., and CAVINS, J. F. The chromatographic determination of cystine and cysteine residues in proteins as S-β-(4-pyridylethyl)cysteine. J. Biol. Chem. 245: 3868 (1970).
13. KOENIG, R., STEGEMANN, H., FRANCKSEN, M., and PAUL, M. L. Protein subunits in the potato virus X group. Determination of the molecular weights by polyacrylamide electrophoresis. Biochim. Biophys. Acta 207: 184 (1970).
14. BIETZ, J. A., SHEPHERD, K. W., and WALL, J. S. Single-kernel analysis of glutenin: Use in wheat genetics and breeding. Cereal Chem. 52: 513 (1975).
15. LIU, T.-Y., and CHANG, Y. M. Hydrolysis of proteins with p-toluenesulfonic acid. Determination of tryptophan. J. Biol. Chem. 246: 2842 (1971).
16. Beckman Model 890C Sequencer Instruction Manual S-IM-2B (1973).
17. BENSON, J. V., Jr., and PATTERSON, J. A. Accelerated automatic chromatographic analysis of amino acids on a spherical resin. Anal. Chem. 37: 1108 (1965).
18. CAVINS, J. F., and FRIEDMAN, M. Automatic integration and computation of amino acid analyses. Cereal Chem. 45: 172 (1968).
19. EDMAN, P., and HENSCHEN, A. Sequence determination. In: Protein sequence determination: A sourcebook of methods and techniques, ed. by S. B. Needleman, p. 232. Springer-Verlag: Berlin, Heidelberg, New York (1975).
20. Beckman Insts., Inc.: New Quadrol Protein Programs. In Sequence No. 8: 8 (1975).
21. NIALL, H. D. Automated Edman degradation: The protein sequenator. Methods Enzymol. 27: 942 (1973).
22. SUMMERS, M. R., SMYTHERS, G. W., and OROSZLAN, S. Thin-layer chromatography of sub-nanomole amounts of phenylthiohydantoin (PTH) amino acids on polyamide sheets. Anal. Biochem. 53: 624 (1973).
23. PISANO, J. J. Analysis of amino acid phenylthiohydantoins by gas chromatography and high performance liquid chromatography. In: Protein sequence determination: A sourcebook of methods and techniques, ed. by S. B. Needleman, p. 280. Springer-Verlag: Berlin, Heidelberg, New York (1975).

24. HAMAUZU, Z., TOYOMASU, T., and YONEZAWA, D. Molecular weight determination of gliadin fractions in gel filtration by SDS-polyacrylamide gel electrophoresis and sedimentation equilibrium. Agric. Biol. Chem. 38: 2445 (1974).
25. PLATT, S. G., KASARDA, D. D., and QUALSET, C. O. Varietal relationships of the α-gliadin proteins in wheat. J. Sci. Food Agric. 25: 1555 (1974).
26. PATEY, A. L., EVANS, D. J., TIPLADY, R., BYFIELD, P. G. H., and MATTHEWS, E. W. Sequence comparison of γ-gliadin and coeliac-toxic α-gliadin. Lancet 2(7937): 718 (1975).
27. PATEY, A. L., and EVANS, D. J. Large-scale preparation of gliadin proteins. J. Sci. Food Agric. 24: 1229 (1973).
28. MEREDITH, P. On the solubility of gliadinlike proteins. IV. Characterization of some gamma-gliadins. Cereal Chem. 44: 436 (1967).
29. SEXSON, K. R., and WU, Y. V. Molecular weights of wheat γ_1- and γ_3-gliadins in various solvents. Biochim. Biophys. Acta 263: 651 (1972).
30. KASARDA, D. D., BERNARDIN, J. E., and NIMMO, C. C. Wheat proteins, p. 158. In: Advances in Cereal Science and Technology, Vol. 1, ed. by Y. Pomeranz. Amer. Assoc. Cereal Chem.: St. Paul, Minn. (1976).
31. SHEPHERD, K. W. Chromosomal control of endosperm proteins in wheat and rye. 3rd Int. Wheat Genet. Symp., ed. by K. W. Finlay and K. W. Shepherd, p. 86. Plenum Publishing Corp.: New York (1968).
32. KASARDA, D. D., BERNARDIN, J. E., and QUALSET, C. O. Relationship of gliadin protein components to chromosomes in hexaploid wheats (*Triticum aestivum* L.). Proc. Nat. Acad. Sci. USA 73: 3646 (1976).
33. HEMMINGS, C., HEMMINGS, W. A., and PATEY, A. L. A fate of oral doses of beta- and gamma-gliadin in adult rats. IRCS Libr. Compend. 4: 152 (1976).