Potential β-Sheet Surfaces of Corn and Wheat Proteins

John A. Rothfus*

Biomaterials Processing Research, National Center for Agricultural Utilization Research, Agricultural Research Service, U.S. Department of Agriculture, Peoria, Illinois 61604

Comparison of computed profiles for β-sheet conformations of zein, glutelin, gliadin, glutenin, and other cereal or nonplant proteins discloses unusual distributions of volume, polarity, and chemical functionality as well as inter- and intraspecific surface analogies between dissimilar amino acid sequences. Volume and polarity surfaces of major cereal proteins are more like those of collagen or fibrinogen than those of silk or hydrocarbon polymers. Potentially useful segments occur nonrandomly. A low molecular weight glutenin fragment approximates two-thirds of human γ-fibrinogen. Four others duplicate the polarity of 25−50% of collagen. Ten exceptionally asymmetric segments show potential for inclusion in film-forming preparations and adhesives. Ultrahydrophobic sequences in corn oil body oleosins approach the diameter of spider silk, match its uniformity, and exceed its hydrophobicity.

Keywords: *Maize; zein; glutelin; gliadin; glutenin; oleosin; computational chemistry; amphiphilicity; molecular volume*

INTRODUCTION

Fibrous proteins, which readily find nonfood use, consist of extended uniform structures that self-associate like hydrocarbon polymers. With chemical adaptation, less-fibrous proteins also perform well in nonfood applications. Agricultural proteins, particularly collagen and casein, were thus used extensively in adhesives, coatings, plastics, and fibers until midcentury (Salzberg, 1953), when petroleum-based polymers began to supplant natural materials. Collagen and casein are now consumed mostly in food (Schrieke and Winter, 1985), but other agricultural proteins promise to become increasingly plentiful and economical as energy and environmental concerns encourage development of botanical replacements for fossil-based fuels and chemicals. Biofuels byproduct proteins ill-suited for direct food use might serve nonfood material needs if their structures can be matched to specific tasks at reasonable cost.

Soybean−blood protein blends already provide example. They rank among the best and most moisture-resistant plywood adhesives, but some formulations based solely on soy protein are also among the least water-resistant (Pocius, 1991). Such experience and early commercial ventures that produced textile fibers from plant proteins (Rebenfeld, 1988) emphasize the importance of judicious selection from seed protein mixtures. They also suggest that careful attention should be given to favored conformations of the individual proteins and to physicochemical reactivities allowed by their covalent structures. In this regard, knowledge of especially stable native configurations like that found with wheat storage protein (Miles et al., 1991) is helpful. In addition, however, the unnatural uses to which seed proteins might be directed demand that any search for useful components also consider unnatural conformations and reactivities. Accordingly, this work compares volume and polarity distributions and reactive residue placements in seed proteins at a fundamental level, in β-sheet conformation. Exploitable similarities and potential utility can be identified by such comparisons.

* Author for correspondence [phone, (309) 685-4011, ext. 318; fax, (309) 681-6686].

PROCEDURES

Volume and amphiphilic character profiles are computed from amino acid sequence data by a moving window analysis described by Rose et al. (1985). Modifications allow residue sorting to simulate β-sheet distributions and allow analysis of residues along separate sides of each extended polypeptide chain. Each side is scanned from N-terminus to C-terminus at a window width of five residues (i.e., index residue ± two residues). Larger or smaller window widths diminish or obscure profile details.

Segment or molecular properties are summations of individual amino acid dimensions and amphiphilic characteristics. Residue numbers are inclusive, volumes are in Å^3, and amphiphilicities are in arbitrary units (au). Values assigned to alternate residues are summed separately to estimate bulk and polarity on either side of each polypeptide chain. Side A thus represents odd-numbered residues of the original protein, side B, even-numbered. Summations, which remain unmodified for terminal structure and charge, are expressed on a per-residue basis to facilitate comparisons between molecules or segments of different size.

Uniformity of volume and amphiphilicity on each side of a molecule or segment is estimated in terms of property dispersions, i.e., the standard deviation of volume or amphiphilicity for a specific segment or side. For example, homopolypeptides have volume and amphiphilicity dispersions of zero, while peptides with equal quantities of each of the 20 common amino acids on either side give volume and amphiphilicity dispersions of 42 and 62, respectively. Certain alternating sequences, e.g., poly(glycyltryptophan) or poly(isoleucylarginine), produce the greatest volume (86 Å^3) or amphiphilicity (125 au) dispersions for their entire sequences while maintaining zero dispersions on either side of the peptide backbone.

Residue volumes were those of Chen and Bendedouch (1986) with values as follows: Ile, 169; Phe, 203; Val, 142; Leu, 168; Trp, 238; Met, 171; Ala, 92; Gly, 66; Cys, 106; Tyr, 204; Pro, 129; Thr, 122; Ser, 99; His, 167; Glu, 141; Asn, 135; Gln, 161; Asp, 114; Lys, 176; Arg, 181.

"Consensus" hydrophobicities of Eisenberg et al. (1984) were multiplied by −100 to facilitate comparisons and associate positive amphiphilicity with enhanced polarity. Values were as follows: Ile, −73; Phe, −61; Val, −54; Leu, −53; Trp, −37; Met, −26; Ala, −25; Gly, −16; Cys, −4; Tyr, −2; Pro, 7; Thr, 18; Ser, 26; His, 40; Glu, 62; Asn, 64; Gln, 69; Asp, 72; Lys,

S0021-8561(95)00477-8

Table 1. Proteins

designation	protein	residues	reference
	Corn—Seed		
MZEIN15	15 kDa zein (cZ15A3)	160	Marks et al., 1985
MZEIN19	19 kDa zein (cZ19D1)	219	Marks et al., 1985
MZEIN22	22 kDa zein (cZ22B1)	245	Marks et al., 1985
MGTEL	28 kDa glutelin-2 (Zc2)	204	Boronat et al., 1986
MOLSN16	oil body oleosin	147	Tzen et al., 1992
MOLSN18	oil body oleosin	187	Tzen et al., 1992
	Corn—Nonseed		
MHIPRO	Pro-rich protein	267[a]	Stiefel et al., 1988
MHPFOB	Pro-rich hydrophobic protein	301[a]	Jose-Estanyol et al., 1992
MHIGLY	Gly-rich protein (CHEM2-GRP)	155[a]	Didierjean et al., 1992
	Wheat—Seed		
WABGLI	α/β-type gliadin	293	Garcia-Maroto et al., 1990
WGGLI	γ-gliadin	276	Bartels et al., 1986
WLWGTENA	low MW glutenin (LMWG-1D1)	284	Colot et al., 1989
WHWGTEN	high MW glutenin (Glu-D1-2b)	621	Anderson et al., 1989
	Wheat—Nonseed		
WHIPRO	Pro-rich protein	378[a]	Raines et al., 1991
	Nonseed		
SILK2	*N. clavipes* Spidroin 2	627	Hinman and Lewis, 1992
HCOLA1X	human α1(X)collagen	680[b]	Thomas et al., 1991
HGFBGN	human fibrinogen, γ-chain	411	Watt et al., 1978

[a] Includes signal peptide. [b] Includes N-terminal and C-terminal noncollagenous sequences of 56 and 161 residues, respectively.

110; Arg, 176. A general frame of reference was established by analyses of 50 100-residue random sequences and 5 1000-residue random sequences generated from a randomized amino acid population weighted according to the frequency of occurrence of amino acid residues in known protein sequences (McCaldon and Argos, 1988). For comparison purposes, several nonseed and nonplant proteins were also examined. Proteins discussed in this work are listed in Table 1 along with designations by which they are identified.

RESULTS AND DISCUSSION

With backbone chains of carbon and nitrogen atoms, the simplest polypeptides resemble hydrocarbon polymers. Their monomeric units range from 66 to 238 Å^3 and are about the size of monomeric units in common plastics such as poly(ethylene) (55 Å^3) or poly(styrene) (190 Å^3). Thus, silk, in which small amino acids predominate, readily assumes a stable β-sheet conformation with its backbone atoms arranged in zigzag chains analogous to solid alkanes. Depending on composition, sequence, and hydrogen bond structure, other proteins adopt three-dimensional configurations that are far more complex but also more susceptible to change with variation in their environments. Knowledge of parameters that control conformations of seed proteins and the kinds of surfaces they present is thus fundamental to adapting them to nonfood uses. Comparisons of β-sheet conformations allow insight into inherent features that affect how their surface properties arise.

Random Sequences. Arithmetic averages suggest that a completely random polypeptide sequence should have an average per-residue volume of 149 Å^3, which is slightly larger than the average monomeric volume for common hydrocarbon polymers (ca. 123 Å^3). Amphiphilicities would average 14.6 au/residue. This is much more polar than −9 au/methylene, which might be expected for poly(ethylene) if an average of the amphiphilicity differences between isoleucine or leucine and valine is an accurate estimate of methylene polarity.

Sequences constructed randomly from an amino acid population that reflects natural abundance are somewhat different from arithmetic means. Analyses of such sequences, representing 10 000 residues, led to average per-residue volume and amphiphilicity values as follows: side A, 141 ± 3 Å^3, 16.4 ± 5.9 au; side B, 142 ± 3 Å^3, 13.4 ± 4.3 au. These values are generally consistent with the natural predominance of small and polar amino acids (McCaldon and Argos, 1988), the range of amino acid volumes from 66 to 238 Å^3, and the fact that amino acid amphiphilicities span a range from −73 to 176 au. Volume dispersion in the random sequences averaged 39 ± 1 Å^3, amphiphilicity dispersion, 66 ± 3 au.

As anticipated, there was little or no variation in volume or volume dispersion from side to side among most random sequences. Amphiphilicity averages, however, suggested substantial variation. Furthermore, distributions of specific amino acids to either side of random β-sheet sequences averaged ±3.5 residues from one-half the total for each type, not zero or ±0.5, as would be expected from equal distribution. The asymmetry implied by these unexpected deviations was most apparent in segments from two of five 1000-residue random polypeptides. Each deviant segment represented about 18% of the parent sequence. One was noticeably different due to asymmetric polarity (side A, 21.3 au; side B, −1.9 au), the other, due to reduced volume and polarity, 136 Å^3 and 3.5 au.

Natural replication of seed proteins under the influence of nucleic acid templates is hardly susceptible to random selection. However, in the absence of adequate data, examination of random sequences provides a context within which to differentiate between sequences and recognize features that natural selection has enriched beyond random levels. Accordingly, polypeptides that differ by more 7 Å^3 and 13 au/residue probably arise by other than random selection and therefore may exhibit significantly different properties. Likewise, dispersion differences exceeding 5 Å^3 and 10 au would suggest significant differences in surface uniformity, provided clusters of extreme properties are recognized and profiles justify comparisons.

Fortunately, nonrandomness is often obvious in natural protein sequences. Throughout most proteins, distributions of amino acid types to either side of a β-sheet deviate only ±1.5 residues from one-half the total for

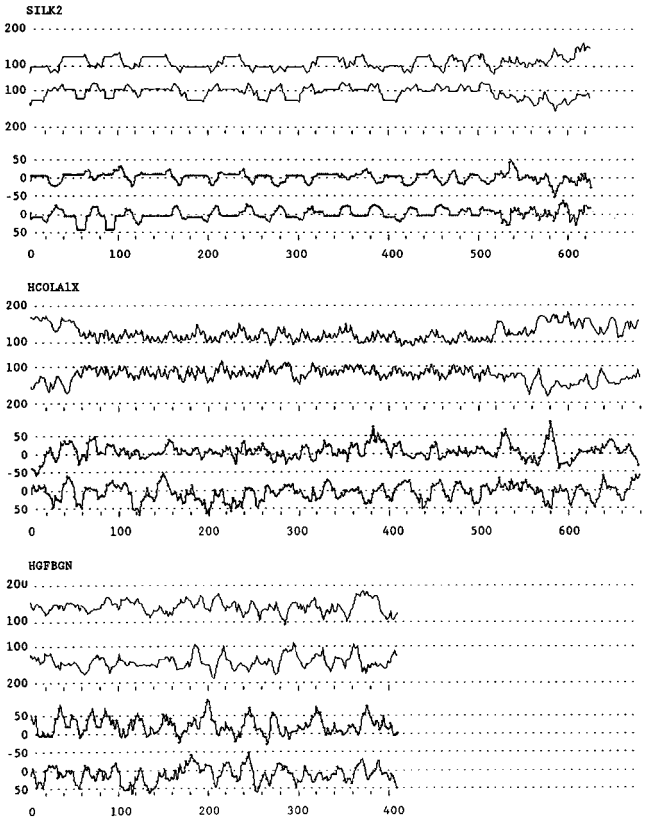

Figure 1. Volume and amphiphilicity profiles of selected nonplant proteins. Upper pair of curves traces volume in Å^3, lower, amphiphilicity in arbitrary units. Upper curve in each pair represents side A of the peptide chain, lower, side B. Properties are displayed, left to right, from N-terminus to C-terminus. Tick marks indicate 20 residues. HCOLA1X profiles include N-terminal and C-terminal noncollagenous sequences of 56 and 161 residues, respectively.

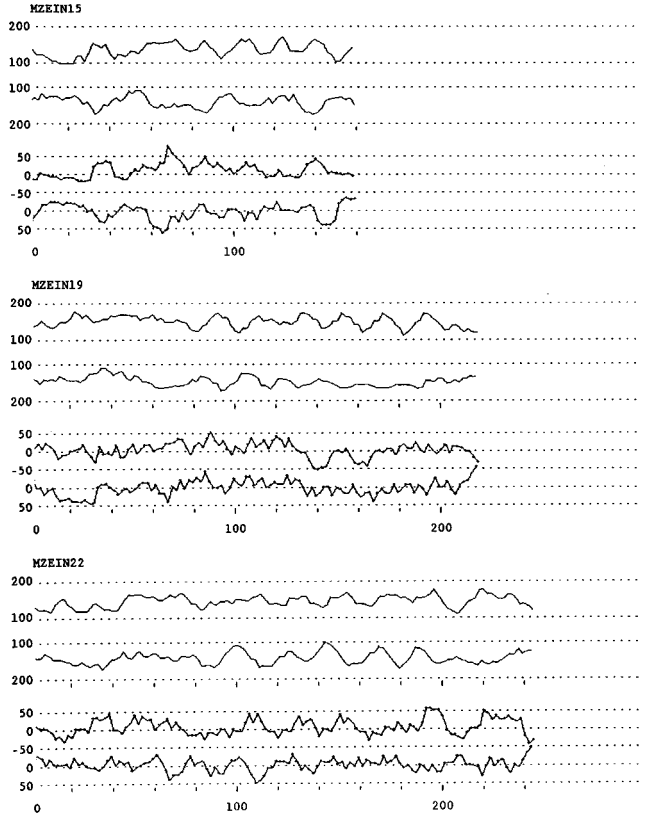

Figure 2. Volume and amphiphilicity profiles of corn zein proteins. Properties are displayed as in Figure 1.

the type, and sequences that segregate the majority of a specific amino acid type to a single side are easily recognized.

Nonplant Proteins. Compared to random sequences, spider silk (Figure 1, SILK2) is remarkably narrow and apolar. Its uniform architecture appears ideal for self-association and flexibility. Regular segments of reduced volume and polarity along the SILK2 profiles coincide with repeated short sequences of poly(alanine), which are suspected (Xu and Lewis, 1990) to undergo reversible helix formation that imparts elasticity to silk fibers. From the standpoint of chemical reactivity, it is interesting, though unobvious in the profiles, that two-thirds of the protein's tyrosine residues and nearly 60% of its serine residues align along side A of the β-sheet.

Human collagen (Figure 1, HCOLA1X) is also smaller in diameter and more uniform in volume than random sequences, but its amphiphilicity profile evidences surfaces more polar and varied than those of silk. Properties along opposite sides of the collagen backbone are frequently asymmetric. Residues 200−500 contain several examples of asymmetric polarity, in which polar and nonpolar residues alternate. Such extreme disparity between sides is known in at least one natural sequence and several synthetic peptides to cause spontaneous generation of membranous macrostructure (Zhang et al., 1993). Analogous to silk, collagen concentrates 60−70% of its tyrosine, histidine, and serine residues on side A of the β-sheet.

The γ-chain of human fibrinogen (Figure 1, HGFBGN) exhibits little of the uniformity seen in silk or collagen even though it embodies elements of the durable solid fibrin (Doolittle, 1992). Like HCOLA1X, HGFBGN contains short segments in which polarity differs substantially from one side to the other, which could favor aggregation during initial stages of clot formation. The fact that complete conversion of fibrinogen to fibrin also involves both intermolecular association and covalent bonding (Chen and Doolittle, 1970) demonstrates the importance of multiple mechanisms in the formation of useful materials from proteins.

Corn Seed Proteins. Throughout β-sheet profiles for corn zeins (Figure 2), properties along one side are frequently reflected on the other. In this respect, zeins resemble silk (Figure 1, SILK2), and they are even more uniform than the insect protein in terms of volume. However, they are larger in cross section than silk and are more like collagen in terms of polarity and amphiphilic uniformity. Whole zeins average 143 Å^3 and 3.8 au/residue.

Argos et al. (1982) attributed periodicity in MZEIN22 to the presence of nine homologous repeat units composed of relatively hydrophobic helical segments joined through hydrophilic glutamine-rich coils. Curiously, similar periodicity is apparent in MZEIN15, in which Pedersen et al. (1986) found neither repeated sequences nor sequence homology to MZEIN19 or MZEIN22. It is easy to presume that such surface analogies allow zeins to fulfill similar functions even though constituted by dissimilar primary structures.

Considering that evolutionary amino acid insertions or deletions may have separated critical elements (Shewry and Tatham, 1990) or inverted the sides of sequences, it is logical that added analogies can be found by comparing profiles in different orientations. Super-

Table 2. Analogous Segments and β-Sheet Surfaces in Corn Proteins[a]

segment		residues	volumes (Å^3)		dispersions[b] (Å^3)		amphiphilicities (au)		dispersions (au)	
MZEIN15	MZEIN19									
66−127A[c]	71−132A	62	143	146	33	33	20.6	19.8	52	57
65−140B	25−100B	76	147	143	36	33	5.8	4.9	52	64
	MZEIN22									
1−24	1−24	24	126	138	43	31	−10.0	−10.3	30	44
28−47	61−80	20	139	148	42	30	6.9	6.3	53	53
87−126	127−166	40	140	142	35	31	6.0	4.8	51	49
122−160	167−205	39	142	146	42	33	3.4	4.6	46	56
3−135	67−199	133	136	144	40	32	6.1	5.3	50	52
	MGTEL									
53−83	126−156	31	143	148	33	33	27.3	27.8	59	68
MZEIN19	MZEIN22									
163−210	21−68	48	148	142	32	33	3.9	4.3	53	55
72−109	123−160	38	144	143	35	32	1.9	1.9	52	49
2−35	208−241	34	145	149	36	32	9.1	7.9	62	65
35−209	6−180	175	149	143	32	32	1.4	1.0	57	51
MGTEL	MOLSN16									
117−151	108−142	35	142	135	32	38	33.1	32.3	58	70

[a] Volumes and amphiphilicities are per-residue. [b] Standard deviation of residue volume or amphiphilicity for segment. See text. [c] Inclusive residue numbers. A = odd-numbered residues; B = even-numbered.

imposing profiles in Figure 2 without regard to sequence orientation identifies numerous identical or nearly identical segments of less than 20 residues. Frequently, similar segments span 30−100 residues. Less often, analogies persist throughout nearly complete sequences. Similarities can be striking even when profiles fail to fit completely within limits established by random sequence analysis. Hierarchal preference favoring amphiphilicity > volume > uniformity allows selection of several segments with nearly identical amphiphilicities but imperfect volume or uniformity matches. Quantitative comparisons affirm the similarities and dissimilarities seen in profiles.

As shown in Table 2, N-terminal sequences of MZEIN15 and MZEIN22 are both hydrophobic to the same extent even though the MZEIN15 segment has a smaller volume and is somewhat less uniform. Three additional segments, which overlap the homologous repeat units described by Argos et al. (1982), match equally well. It appears that more than 75% of MZEIN15 could have functional counterparts in MZEIN22. At least 25% of MZEIN15, residues 87−126, matches MZEIN22, residues 127−166, quite well in terms of polarity, volume, and uniformity.

All three zeins have analogous domains near the center of their sequences and probably share whatever functionality these central segments confer. Similar skewed polar domains flanked by nonpolar segments appear frequently in the corn profiles. For example, as shown in Table 2, MZEIN15(53−83) compares nicely with MGTEL(126−156). The function of such segments is obscure.

MZEIN15 and MZEIN22, like SILK2, carry a majority of their tyrosine residues on one side of the β-sheet backbone: 64% and 86%, respectively. In addition, MZEIN15 distributes 60% of its hydrophilic amide and hydroxy residues along side A, while 65% of its most hydrophobic residues are on side B. This gives the molecule a potentially useful functional sidedness. In MZEIN22, amides and hydroxy residues are essentially the same on either side, but amphiphilic sidedness is enhanced by six acidic and basic residues located on side A. Asymmetry is different in MZEIN19. Volume and amphiphilic sidedness are substantial in specific segments, but the molecule is internally balanced overall because the levels of different residue types are essentially the same on either side of the backbone.

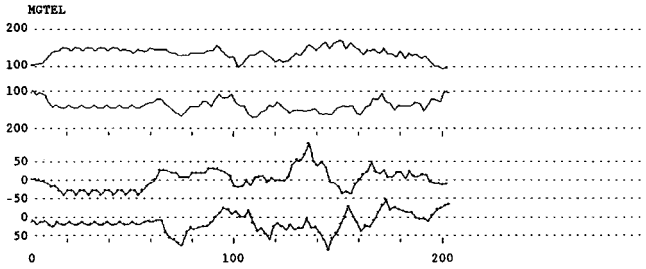

Figure 3. Volume and amphiphilicity profiles of corn glutelin.

Unfortunately, major seed proteins seldom sustain sidedness and uniformity simultaneously for more than 30−40 residues. The uniformly asymmetric and repetitive N-terminal domain of MGTEL (Figure 3) is an interesting exception. Polar asymmetry in this portion of the molecule generates a per-residue difference of 50 au between sides throughout a 50-residue sequence, one-fourth of the molecule. By comparison, sides of a membrane-forming 16-residue peptide in yeast zuotin (Zhang et al., 1993) differ by 111 au, and the sides of segments in a marine mussel adhesion protein analog (Williams et al., 1989) differ by 53 au. The yeast peptide spontaneously forms "feltlike" membranous macrostructures in the presence of salt. It seems likely that a glutelin fragment, with one-half the polarity difference sustained for 3 times the length, might behave similarly. Conveniently, MGTEL contains a potential site for cleavage by trypsin at residue 72. Digestion at this point would release the N-terminal peptide with 10 of the parent protein's 16 histidine residues. Nine of these 10 histidines distribute to the same side of the β-sheet. In addition, cysteine residues at either end of the peptide should allow access to novel materials by chemical construction of disulfide-linked extensions and networks.

Similar prospects for exploitation in biomaterials also make seed oleosins appealing. Represented here by MOLSN16 and MOLSN18 (Figure 4), oleosins are a relatively minor group of proteins associated with seed oil bodies. Huang and co-workers (Tzen et al., 1993) estimate that oleosins may account for as much as 7% of total protein in rapeseed. Levels in corn would be proportionately lower depending upon oil content.

All oleosins contain a genetically conserved and extremely hydrophobic central segment, generally 70−

Table 3. Analogous Segments and β-Sheet Surfaces in Wheat Proteins

segment		residues	volumes (Å³)		dispersions (Å³)		amphiphilicities (au)		dispersions (au)	
WABGLI	WGGLI									
22−106	98−182	85	153	151	28	31	21.6	19.7	47	55
10−250B	10−250A	241	150	151	27	31	24.7	24.0	52	48
	WLWGTENA									
11−285B	3−277B	275	149	149	27	31	21.5	23.4	53	58
	WHWGTEN									
5−189	271−455	185	152	133	27	42	26.4	27.5	52	41
2−111B	58−167B	110	149	141	27	42	18.3	18.4	49	49
194−293	180−279	100	146	137	32	35	18.6	19.8	53	57
WGGLI	WLWGTENA									
5−95B	5−95B	91	150	154	28	30	25.5	24.8	43	53
73−152	7−86	80	151	155	27	29	24.8	23.5	50	52
100−180	120−200	81	150	148	31	31	18.8	18.4	55	63
128−242A	92−206A	115	146	144	35	30	17.7	16.4	51	58
WHWGTEN	WHWGTEN									
18−565A	18−565B	548	135	136	42	43	26.6	24.7	47	46

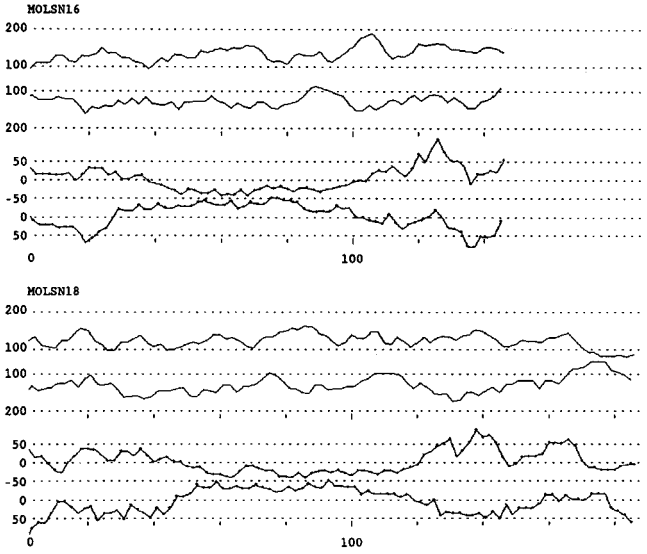

Figure 4. Volume and amphiphilicity profiles of corn oleosin proteins.

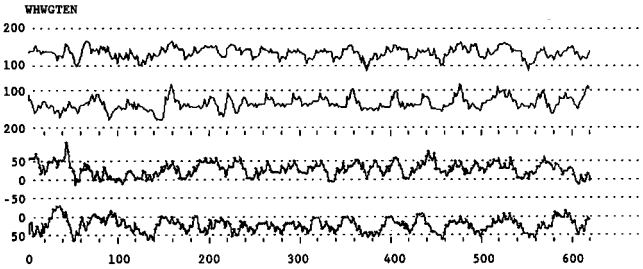

Figure 5. Volume and amphiphilicity profiles of wheat high molecular weight glutenin.

80 residues in length, which is thought to form a β-structure loop that protrudes into the triglyceride-rich matrix of the oil body, while polar portions of the protein combine with phospholipids to constitute the lipid body membrane (Tzen et al., 1992). Basic residues generally flank this central domain, which should allow for selective release of the extremely hydrophobic fragment, provided the parent protein can be dispersed adequately for enzymic digestion.

Central fragments from the oleosins, like whole corn proteins, approximate collagen in terms of per-residue volume (ca. 130 Å³), but their hydrophobicity (ca. −25 au) is unmatched by other seed proteins. A 70-residue fragment from the central hydrophobic domain of jack bean urease (Takishima et al., 1988) is slightly smaller per-residue (126 Å³) and produces a strikingly similar profile, but it contains five acidic and five amide residues, enough to impart polarity (−7.7 au) more like that of silk (0.3 au).

Near its C-terminus, MOLSN16 also contains a 31-residue segment of skewed polarity analogous to domains observed near the center of MGTEL and throughout the zeins. MOLSN16(108−138), for example, compares to MGTEL(117−147) (Table 2). The occurrence of analogous surfaces involving extremes of polarity in these proteins suggests some common fundamental function that might, for example, be critical to cellular order during seed development. Obviously, much remains to be learned about structures that contribute to higher order interactions among cereal proteins.

In this regard, it is noteworthy that bulky regions of certain corn volume profiles interdigitate nicely with those of other sequences, as if the molecules might pack efficiently. This is obvious in the zein profiles (Figure 2) but less apparent in the glutelin (Figure 3) and oleosin (Figure 4) profiles. Nonetheless, side A of MGTEL(125−165) meshes with side B of MOLSN16-(100−140), and side B of MGTEL(50−100) fits on side A of MOLSN16(7−57).

Wheat Seed Proteins. Wheat proteins, composed of repeated sequences (Kasarda et al., 1984; Anderson and Greene, 1989) rich in glutamine and proline, are volumewise among the most uniform of major seed proteins. Unfortunately, they are also quite hydrophilic and bulky. Though not identical, the major wheat proteins share analogous segments and surfaces (Table 3).

WHWGTEN is exceptionally uniform with nearly palindromic profiles (Figure 5). A 2-fold inversion that places the C-terminus over the N-terminus and side A over side B superimposes much of the WHWGTEN profile on itself. Consistent with this visual match, per-residue values for the bulk of the molecule (Table 3) indicate analogy between the two sides of WHWGTEN.

WHWGTEN's per-residue volume is among the lowest for seed proteins, but the molecule is also among the most hydrophilic. Accordingly, it is hardly surprising that these long uniform molecules of WHWGTEN associate readily to form the highly extensible hydrated films that are critical to breadmaking.

In WABGLI (Figure 6), the five domains defined for α-gliadin by Kasarda et al. (1984) are as follows: I, repeating, residues 1−109; II, poly(glutamine), 110−133; III, unique, 134−202; IV, poly(glutamine), 203−217; V, unique, 218−293. The poly(glutamine) domains

Table 4. Analogous Segments in Corn and Wheat Proteins

segment		residues	volumes (Å^3)		dispersions (Å^3)		amphiphilicities (au)		dispersions (au)	
MZEIN15	WABGLI									
59−119	126−186	61	146	152	32	29	18.9	18.2	59	58
	WLWGTENA									
60−130	119−189	71	146	148	32	30	16.2	16.8	57	64
	WHWGTEN									
60−120	41−101	61	145	140	32	39	19.5	20.0	59	62
MZEIN22	WGGLI									
91−159	154−222	69	143	149	31	34	6.2	7.7	51	56
MGTEL	WLWGTENA									
75−160	105−190	86	140	149	34	28	16.3	16.8	55	63
55−179	159−283	125	139	143	33	34	15.3	15.0	54	58
	WHWGTEN									
63−162	12−111	100	140	139	33	40	18.4	18.6	55	63
MOLSN18	WHWGTEN									
6−45	97−136	40	130	129	47	43	23.1	21.6	58	45
120−156	221−257	37	135	138	38	44	31.7	31.8	71	48
20−51	109−140	32	130	133	42	45	24.9	25.7	56	48
119−148	2−31	30	140	138	36	32	32.4	32.8	76	67

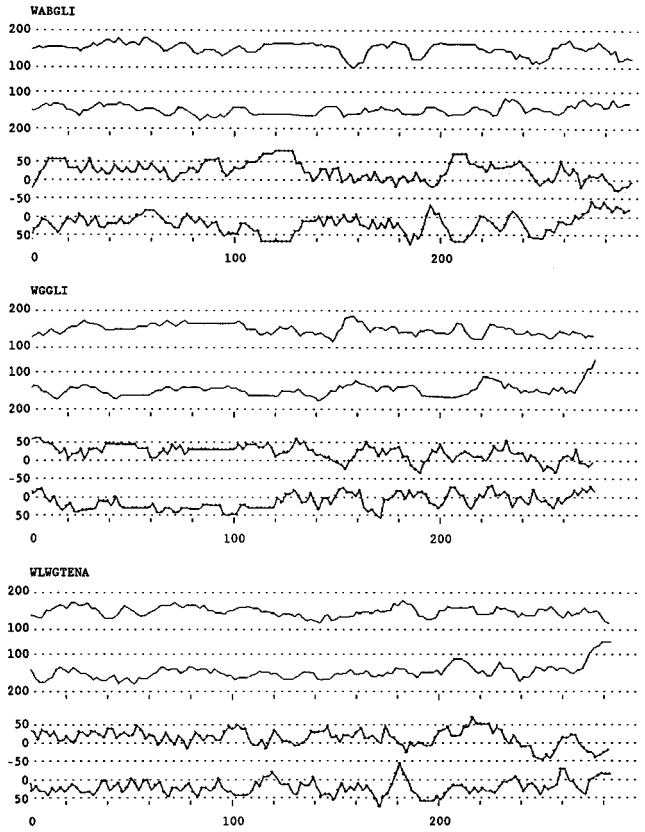

Figure 6. Volume and amphiphilicity profiles of wheat gliadin and low molecular weight glutenin proteins.

produce obvious volume and amphiphilic symmetry in the WABGLI profiles. Domain V is asymmetric in ways that leave both sides of the segment nearly equal overall. Domain III, the most hydrophobic segment of WABGLI, is also quite sided due to concentrations of small residues at either end and polar residues near its C-terminus.

The repetitive domain, WABGLI(1−109), compares in an interesting way with analogous repetitive portions of WGGLI and MGTEL. WABGLI(1−109) accounts for 31% of the molecule's most hydrophobic residues, 34% of its amide residues, 75% of its proline, and 60% of its tyrosine and phenylalanine. Polar and nonpolar residues alternate such that amphiphilic sidedness is not extreme (side A, 28.3 au; side B, 17.4 au), and tyrosine and phenylalanine residues are distributed about equally on either side.

The N-terminal half of WGGLI is similar, containing 34% of the molecule's most hydrophobic residues, 66% of its amide residues, 78% of its proline, and 86% of its phenylalanine. Higher levels of amide residues make this segment polar overall but not significantly sided (side A, 34.0 au; side B, 23.6 au).

The analogous repetitive domain, MGTEL(1−88), likewise accounts for a total 60% of the molecule's hydrophobic and amide residues, 75% of its proline, and 80% of its histidine. In the MGTEL segment, however, all of the most hydrophobic residues are on side A, and 12 of 13 histidine residues and a majority of the amide residues are on side B, which imparts an extreme polar (side A, −12.0 au; side B, 24.6 au) and chemical sidedness not seen in its wheat counterparts.

In each of these proteins, basic residues are located such that digestion with trypsin should leave their repetitive domains intact. WGGLI would yield the longest fragment, a bulky (16 kDa, 21 000 Å^3) 138-residue hydrophilic polymer (29 au) that averages 1 phenylalanine/11 residues. With a single cysteine penultimate to its C-terminus, this peptide should easily form a disulfide-bonded dimer. Two C-terminal lysine residues and two side-chain carboxyl groups near N-termini could give this dimer interesting sequestering properties. Such a dimer also presents intriguing possibilities for ionic bonded or condensation cyclizations and polymerizations. Unusual ionic character, perhaps useful in surfactants, could be achieved by oxidizing the single cysteine residue near the C-terminus of the WGGLI fragment. The WGGLI fragment also contains 10 of the 25 serine and threonine residues in WGGLI, which would allow glycosylation or esterification.

Other lengthy profiles shared by corn and wheat are identified in Tables 4 and 5. Among those most surprising are three involving more than one-half the molecule in MGTEL, MZEIN19, and MZEIN22 and those in which profiles for opposite ends of MOLSN18 overlay more than one segment of the WHWGTEN profile. It appears that portions of WHWGTEN might engage in ordered interactions in wheat analogous to those through which the polar ends of MOLSN18 form oil body membranes in corn.

When compared in profile and quantitatively (Table 6), none of the seed proteins are as small or nonpolar as silk fibroin (SILK2). Only the 70−80-residue central segments from corn oleosins, which average 128 Å^3 and −25.5 au with dispersions of 40 Å^3 and 31 au, exceed

Table 5. Analogous β-Sheet Surfaces in Corn and Wheat Proteins

segment		residues	volumes (Å^3)		dispersions (Å^3)		amphiphilicities (au)		dispersions (au)	
MZEIN15	WABGLI									
62−152B	5−95B	91	146	149	35	28	12.1	11.6	61	43
	WGGLI									
70−160A	94−184B	91	141	149	37	30	15.1	14.5	48	60
62−152A	150−240A	91	142	150	36	36	16.8	14.4	50	53
62−152B	150−240A	91	146	150	35	36	12.1	14.4	61	53
MZEIN19	WABGLI									
14−78B	16−80B	65	140	146	31	27	9.9	9.7	70	43
136−175B	141−180B	40	156	150	27	27	14.9	13.4	54	55
	WLWGTENA									
20−135A	63−178A	116	153	146	32	29	12.2	14.4	60	54
	WHWGTEN									
45−130A	75−160A	86	150	131	32	49	14.7	14.0	56	43
MZEIN22	WABGLI									
29−103A	130−204A	75	145	148	30	35	6.6	7.0	58	58
99−163A	140−204A	65	147	146	30	37	5.9	4.9	52	57
	WGGLI									
39−193A	120−274B	155	147	145	30	33	3.6	3.0	51	59
	WLWGTENA									
60−120A	219−279A	61	147	146	28	30	1.7	1.1	53	49
MGTEL	WABGLI									
9−65B	9−65B	57	140	141	19	25	18.3	13.1	18	41
122−162B	165−205B	41	152	151	30	23	27.4	26.0	63	64
	WGGLI									
65−125A	210−270A	61	131	142	31	32	10.1	10.1	33	53

Table 6. Volume and Amphiphilicity Distributions in β-Sheet Proteins

	volumes (Å^3)		dispersions (Å^3)		amphiphilicities (au)		dispersions (au)	
segment	side A	side B	side A	side B	side A	side B	side A	side B
				Corn				
MZEIN15	135	142	40	40	9.8	4.4	47	52
MZEIN19	149	147	35	32	2.5	2.0	57	59
MZEIN22	145	143	33	33	6.2	−2.1	58	50
MGTEL	135	139	31	33	1.1	16.0	45	49
MOLSN16	135	126	43	41	4.6	−3.2	62	52
MOLSN18	124	129	40	44	2.7	4.8	55	63
MHIPRO	137	136	28	28	20.8	21.5	38	38
MHPFOB	134	138	35	33	4.9	4.8	50	51
MHIGLY	124	119	52	49	21.3	10.8	70	57
				Wheat				
WABGLI	150	149	31	28	24.1	20.8	53	54
WGGLI	149	148	31	31	22.2	12.8	50	55
WLWGTENA	148	147	30	33	13.9	22.2	56	57
WHWGTEN	135	135	41	43	27.2	24.0	47	48
WHIPRO	131	156	18	22	11.7	54.3	31	55
				Nonseed				
SILK2	110	106	42	39	0.3	0.4	34	37
HCOLA1X	125	124	44	43	5.0	7.3	49	54
HGFBGN	141	142	41	43	21.9	12.0	61	58

silk's hydrophobicity and match its uniformity. Interestingly, several seed proteins present volume surfaces that are more uniform overall than those of silk.

Several seed proteins also exhibit surprising analogy to lengthy segments of human collagen and fibrinogen (Table 7). More than one-half the peptide sequence in MZEIN22 or WGGLI is as uniformly amphiphilic as that of HCOLA1X. These large fragments and one from WLWGTENA compare even better to HGFBGN, well enough to call for experimental evaluation of effects that small differences in volume and uniformity might have on comparable physical properties.

Nonseed Proteins. The uniformity of WHWGTEN is matched and that of SILK2 is rivaled in profile by another class of proteins common to vegetative portions of crop plants. Figure 7 gives profiles for high-proline proteins of corn (MHIPRO) and wheat (WHIPRO), in which the sequence Pro-Pro-Thr-Tyr-Thr-Pro is repeated 13 times in corn and the sequence Met-Pro-Lys-Pro-Glu-Pro-Lys-Pro-Glu-Pro-Lys-Pro-Glu-Pro- is repeated 14 times in wheat.

MHIPRO generates profiles reminiscent of those for SILK2. Both sequences are highly symmetrical and uniform even though the plant molecule is much more polar than the insect polymer. In contrast, WHIPRO is exceedingly asymmetric. Its sequence, which alternates proline with other types of amino acids, generates per-residue differences of 31 Å^3 and 61 au between sides A and B throughout most of the molecule. Such differences are also seen in MZEIN19, MGTEL, and HGFBGN but only in much shorter segments (10−30 residues).

Treated here as high-proline sequences, MHIPRO and WHIPRO are analogous in both structure and occurrence to extensins, a group of hydroxyproline-rich glycoproteins associated with rapidly growing plant tissue. They are thought to play structural roles in cell walls (Cassab and Varner, 1988; Keller, 1993). Quite likely, MHIPRO and WHIPRO are hydroxylated and glycosylated like other extensins if they survive in mature plants. Such modifications would, of course, increase volumes and amphiphilicities and change profiles from

Table 7. Analogous Segments and β-Sheet Surfaces in Seed and Human Proteins

segment		residues	volumes (Å^3)		dispersions (Å^3)		amphiphilicities (au)		dispersions (au)	
MZEIN15	HGFBGN									
19−110B	114−205B	92	143	143	39	40	10.1	10.3	55	60
MZEIN19	HGFBGN									
26−133A	97−204B	108	152	142	32	39	12.5	12.6	61	63
MZEIN22	HCOLA1X									
90−204	60−174	115	144	115	31	41	5.2	5.5	53	49
9−238	283−512	230	144	114	32	41	4.0	6.0	54	47
	HGFBGN									
23−237A	191−405B	215	146	143	31	47	9.9	9.1	59	56
WABGLI	HGFBGN									
3−95	12−104	93	152	144	27	33	20.8	20.2	47	60
126−197	304−375	72	151	139	30	49	18.3	17.2	60	51
WGGLI	HCOLA1X									
119−270	129−280	152	146	116	32	42	8.7	9.5	56	54
138−258	302−422	121	146	114	33	42	8.4	9.4	56	48
	HGFBGN									
95−265	200−370	171	147	139	31	47	13.5	13.5	55	55
115−233A	76−194A	119	146	145	35	35	18.5	19.3	50	62
WLWGTENA	HGFBGN									
2−276	31−305	275	148	143	30	39	18.6	17.8	57	61

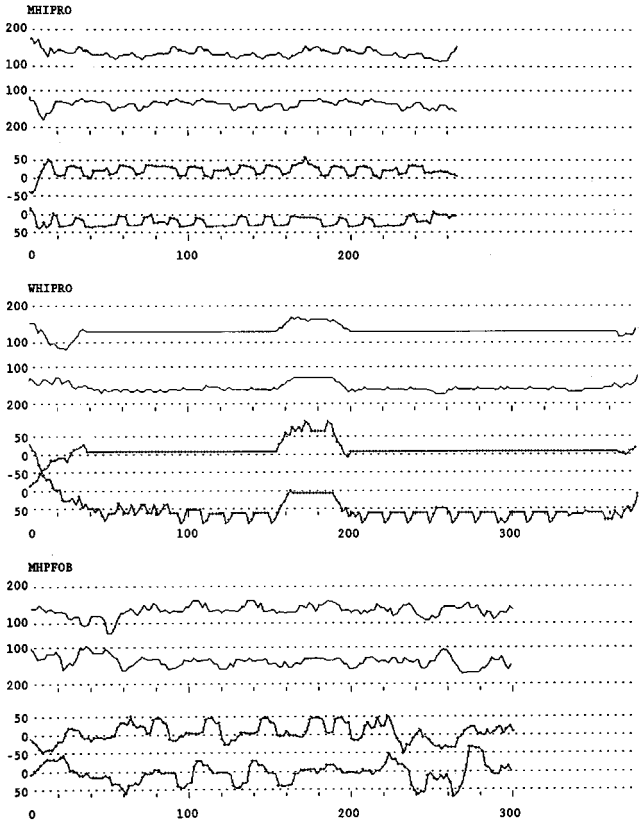

Figure 7. Volume and amphiphilicity profiles of corn and wheat proline-rich proteins.

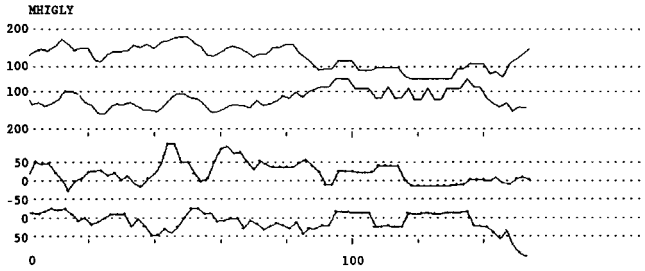

Figure 8. Volume and amphiphilicity profiles of corn glycine-rich protein.

those illustrated here. However, they probably would not destroy the unique uniformity of these sequences, which should be considered among agriculture's stock of novel and potentially useful structures.

MHPFOB, the product of an embryo-specific gene in corn, is an especially hydrophobic analog of MHIPRO, which makes it appealing for use in composite materials with hydrocarbon polymers. In plant cells, MHPFOB may fulfill dual biological functions, i.e., one role similar to that of other proline-rich proteins and another more important to nutrient storage and protection (Jose-Estanyol et al., 1992). Considering that MHPFOB contains rather high levels of cysteine that could form disulfide bonds and stabilize multisubunit forms, it is interesting to note that a rotation of the MHPFOB profiles to superposition N-termini and sides A over sides B produces bimolecular profiles that approximate the monomolecular profiles of MHIPRO. In the absence of information on MHPFOB's function, it is tempting to conjecture that a dimer of MHPFOB might accomplish tasks typical of molecules like MHIPRO while the monomer of MHPFOB would be better suited to other roles.

Similar manipulation of profiles for MHIGLY (Figure 8), which exemplifies another group of structural proteins (Keller, 1993), can lead to additional conjecture about how such molecules function in plant cell walls, but its occurrence in plants is the principal reason for mention here. If hydrolyzed carefully, MHIGLY might yield sizable fragments containing ca. 80% glycine and 20% arginine or tyrosine, which could be restructured chemically into novel materials.

In terms of new materials from seed proteins, it is especially encouraging that segments of the plant proteins often meet or exceed volume, polarity, and uniformity characteristic of well-known polymeric materials. In certain domains, their sidedness (Table 8) rivals the tacticity of some petrochemical polymers and approaches the polar sidedness seen in marine mussel adhesion protein (Williams et al., 1989).

Isotactic poly(propylene), for example, distributes side-chain volume and polarity such that one side of the polymer differs from the other by about 26 Å^3 and 9 au/monomeric unit. This sidedness leads to a 30 °C increase in melting point, reduced hydrocarbon solubility, and higher yield strength compared to syndiotactic poly(propylene) (Lieberman and Barbe, 1988).

Side to side volume differences approaching that of poly(propylene) occur in MZEIN15(3−33), MZEIN19-(32−70), and WLWGTENA(120−159). WHIPRO is

Table 8. Asymmetric β-Sheet Protein Segments

segment	volumes (Å^3)		dispersions (Å^3)		amphiphilicities (au)		dispersions (au)	
	side A	side B	side A	side B	side A	side B	side A	side B
Corn								
MZEIN15(3−33)	112	133	43	48	−12.1	−15.9	27	20
MZEIN19(1−31)	152	144	34	34	5.4	25.8	47	74
MZEIN19(32−70)	158	135	29	33	−0.8	0.6	67	59
MZEIN19(71−136)	148	148	32	33	19.1	−11.9	57	51
MZEIN19(137−147)	144	139	41	27	−48.5	18.8	18	50
MZEIN19(148−158)	172	171	16	26	−5.8	4.8	61	52
MZEIN19(159−170)	128	156	38	24	−35.7	25.2	27	59
MZEIN22(25−65)	142	146	31	36	16.0	−7.8	63	49
MZEIN22(66−100)	145	142	31	33	−3.5	8.4	49	48
MZEIN22(188−245)	147	143	38	33	14.2	−1.7	70	53
MGTEL(1−65)	138	135	22	26	−22.1	17.0	32	19
MOLSN16(108−147)	144	122	36	38	41.5	21.3	72	60
MHPFOB(27−69)	113	129	46	41	10.5	23.5	40	37
MHPFOB(70−229)	140	139	29	26	14.3	3.3	54	43
MHPFOB(230−301)	131	144	37	36	−8.5	7.6	44	72
MHIGLY(1−89)	146	128	42	40	33.5	6.3	80	51
Wheat								
WABGLI(15−70)	160	147	26	27	33.0	10.1	47	44
WGGLI(121−276)	145	145	34	33	13.2	3.0	52	59
WLWGTENA(120−159)	135	159	29	24	14.0	20.0	57	67
WLWGTENA(240−284)	143	132	31	42	−16.5	9.1	42	63
WHIPRO(91−157)	129	160	0	18	7.0	72.0	0	47
WHIPRO(164−191)	163	129	16	0	73.4	7.0	46	0
WHIPRO(196−365)	129	162	0	17	7.0	70.0	0	47
Nonseed								
HCOLA1X(100−300)	116	117	41	43	2.8	14.5	44	60
HGFBGN(170−205)	149	139	36	49	38.3	−14.2	60	55

volume-sided throughout its entire sequence, which is essentially the length of 570 dp, 24 000 kDa, poly(propylene). Perhaps new materials quite atypical of those associated with plant seeds would be possible if such segments or whole proteins were isolated and joined with greater selectivity than is currently practicable.

Polarity differences of 9 au or greater between sides occur frequently in corn proteins. Such asymmetry is especially obvious in MZEIN19(71−170), MGTEL(1−65), MHGLY(1−89), and throughout MHPFOB. Several shorter segments of 30−40 residues, such as MZEIN22-(25−65 and -(66−100), also exhibit obvious asymmetry. Often, as in MZEIN19 and MZEIN22, polarity reverses in contiguous asymmetric segments.

Polar disparity averaging 65 au between sides throughout 70% of WHIPRO is the largest observed thus far for any plant protein segment longer than 40 residues. MGTEL(1−65), second at 39 au, is still significantly more asymmetric than other seed proteins, in which the difference between sides is generally less than 25 au. In this context, differences between sides of HCOLA1X-(100−300), 11.7 au, and HGFBGN(170−205), 52.5 au, make interesting comparisons.

Asymmetric distribution of hydrophobic residues and hydrophilic amide residues is responsible for much of seed protein sidedness. For example, WABGLI(15−70) shows a difference of 23 au primarily due to location of 70% of the segment's amide residues on side A. Similarly, WGGLI(121−276), the C-terminal half of WGGLI, is sided due to 60% of the segment's amide residues on side A and 60% of its most hydrophobic residues on side B. In WLWGTENA, amide residues distribute evenly to either side, but the molecule is still as polarly asymmetric as poly(propylene) overall due to asymmetric concentrations of hydrophobic residues near the C-terminus and in several shorter segments. In WHWGTEN, acidic residues occur preferentially (66%) along one side of the molecule. One-half are located within 140 residues from the N-terminus, where they produce slightly larger volume (9.7 Å^3) and amphiphilicity (4.8 au) differences between sides than are seen throughout the remainder of WHWGTEN (1.7 Å^3, 2.8 au).

In general, the surface characteristics of major seed proteins are consistent with their uses. It is not surprising that zein molecules, with exceptional concentrations of hydrophobic residues and relatively uniform asymmetric segments, associate readily. They find wide use in adhesives or coatings and, for several years, were regenerated (Croston et al., 1945) commercially into textile fibers with good wet properties and resistance to acids and alkalines (Jenkins, 1956). It is intriguing to speculate how much better such products might perform if constructed from selected parts of the seed proteins. Prospects for novel polymers from botanical sources seem sufficient to justify chemical combination of specific protein fragments or the genetic magnification of sequences targeted for particular industrial tasks.

LITERATURE CITED

Anderson, O. D.; Greene, F. The characterization and comparative analysis of high-molecular-weight glutenin genes from genomes A and B of a hexaploid bread wheat. *Theor. Appl. Genet.* **1989**, *77*, 689−700.

Anderson, O. D.; Greene, F.; Yip, R. E.; Halford, N. G.; Shewry, P. R.; Malpica-Romero, J.-M. Nucleotide sequences of the two high-molecular-weight glutenin genes from the D-genome of a hexaploid bread wheat, *Triticum aestivum* L. cv Cheyenne. *Nucleic Acids Res.* **1989**, *17*, 461−462.

Argos, P.; Pedersen, K.; Marks, M. D.; Larkins, B. A. A structural model for maize zein proteins. *J. Biol. Chem.* **1982**, *257*, 9984−9990.

Bartels, D.; Altosaar, I.; Harberd, N. P.; Barker, R. F.; Thompson, R. D. Molecular analysis of γ-gliadin gene families at the complex *Gli-1* locus of bread wheat (*T. aestivum* L.). *Theor. Appl. Genet.* **1986**, *72*, 845−853.

Boronat, A.; Martinez, M. C.; Reina, M.; Puigdomenech, P.; Palau, J. Isolation and sequencing of a 28 kD glutelin-2 gene from maize. Common elements in the 5′ flanking regions among zein and glutelin genes. *Plant Sci.* **1986**, *47*, 95−102.

Cassab, G. I.; Varner, J. E. Cell wall proteins. *Annu. Rev. Plant Physiol.* **1988**, *39*, 321−353.

Chen, R.; Doolittle, R. F. Isolation, characterization, and location of a donor-acceptor unit from cross-linked fibrin. *Proc. Natl. Acad. Sci. U.S.A.* **1970**, *66*, 472−479.

Chen, S. H.; Bendedouch, D. Structure and interactions of proteins in solution studies by small-angle neutron scattering. *Methods Enzymol.* **1986**, *130*, 79−116.

Colot, V.; Bartels, D.; Thompson, R.; Flavell, R. Molecular characterization of an active wheat LMW glutenin gene and its relation to other wheat and barley prolamin genes. *Mol. Gen. Genet.* **1989**, *216*, 81−90.

Croston, C. B.; Evans, C. D.; Smith, A. K. Zein fibers....preparation by wet spinning. *Ind. Eng. Chem.* **1945**, *37*, 1194−1198.

Didierjean, L.; Frendo, P.; Burkard, G. Stress responses in maize: sequence analysis of cDNAs encoding glycine-rich proteins. *Plant Mol. Biol.* **1992**, *18*, 847−849.

Doolittle, R. F. A detailed consideration of a principal domain of vertebrate fibrinogen and its relatives. *Protein Sci.* **1992**, *1*, 1563−1577.

Eisenberg, D.; Weiss, R. M.; Terwilliger, T. C. The hydrophobic moment detects periodicity in protein hydrophobicity. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 140−144.

Garcia-Maroto, F.; Marana, C.; Garcia-Olmedo, F.; Carbonero, P. Nucleotide sequence of a cDNA encoding an α/β-type gliadin from hexaploid wheat (*Triticum aestivum*). *Plant Mol. Biol.* **1990**, *14*, 867−868.

Hinman, M. B.; Lewis, R. V. Isolation of a clone encoding a second dragline silk fibroin. *J. Biol. Chem.* **1992**, *267*, 19320−19324.

Jenkins, H. S. Protein fibers, synthetic. In *Encyclopedia of Chemical Technology*; Kirk, R. E.,Othmer, D. F., Eds.; Interscience: New York, 1956; Vol. 11, p 200.

Jose-Estanyol, M.; Ruiz-Avila, L.; Puigdomenech, P. A Maize embryo-specific gene encodes a proline-rich and hydrophobic protein. *Plant Cell* **1992**, *4*, 413−423.

Kasarda, D. D.; Okita, T. W.; Bernardin, J. E.; Baecker, P. A.; Nimmo, C. C.; Lew, E. J.-L.; Dietler, M. D.; Greene, F. C. Nucleic acid (cDNA) and amino acid sequences of α-type gliadins from wheat (*Triticum aestivum*). *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 4712−4716.

Keller, B. Structural cell wall proteins. *Plant Physiol.* **1993**, *101*, 1127−1130.

Lieberman, R. B.; Barbe, P. C. Propylene polymers. In *Encyclopedia of Polymer Science and Engineering*, 2nd ed.; Kroschwitz, J., Ed.; Wiley: New York, 1988; Vol. 13, p 477.

Marks, M. D.; Lindell, J. S.; Larkins, B. A. Nucleotide sequence analysis of zein mRNAs from maize endosperm. *J. Biol. Chem.* **1985**, *260*, 16451−16459.

McCaldon, P.; Argos, P. Oligopeptide biases in protein sequences and their use in predicting protein coding regions in nucleotide sequences. *Proteins* **1988**, *4*, 99−122.

Miles, M. J.; Carr, H. J.; McMaster, T. C.; I'Anson, K. J.; Belton, P. S.; Morris, V. J.; Field, J. M.; Shewry, P. R.; Tatham, A. S. Scanning tunneling microscopy of a wheat seed storage protein reveals details of an unusual supersecondary structure. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 68−71.

Pedersen, K.; Argos, P.; Naravana, S. V. L.; Larkins, B. A. Sequence analysis and characterization of a maize gene encoding a high-sulfur zein protein of M_r 15,000. *J. Biol. Chem.* **1986**, *261*, 6279−6284.

Pocius, A. V. Adhesives. In *Encyclopedia of Chemical Technology*, 4th ed.; Kroschwitz, J. I., Ed.; Wiley: New York, 1991; Vol. 1, p 458.

Raines, C. A.; Lloyd, J. C.; Chao, S.; John, U. P.; Murphy, G. J. P. A novel proline-rich protein from wheat. *Plant Mol. Biol.* **1991**, *16*, 663−670.

Rebenfeld, L. Fibers. In *Encyclopedia of Polymer Science and Engineering*, 2nd ed.; Kroschwitz, J. I., Ed.; Wiley: New York, 1988; Vol. 12, p 698.

Rose, G. D.; Gierasch, L. M.; Smith, J. A. Turns in peptides and proteins. *Adv. Protein Chem.* **1985**, *37*, 1−109.

Salzberg, H. K. Protein plastics. In *Encyclopedia of Chemical Technology*; Kirk, R. E., Othmer, D. F., Eds.; Interscience: New York, 1953; Vol. 11, p 221.

Schrieke, R. R.; Winter, G. Casein. In *Encyclopedia of Polymer Science and Engineering*, 2nd ed.; Kroschwitz, J., Ed.; Wiley: New York, 1985; Vol. 2, p 690.

Shewry, P. R.; Tatham, A. S. The prolamin storage proteins of cereal seeds: structure and evolution. *Biochem. J.* **1990**, *267*, 1−12.

Stiefel, V.; Perez-Grau, L.; Albericio, F.; Giralt, E.; Ruiz-Avila, L.; Ludevid, M. D.; Puigdomenech, P. Molecular cloning of cDNAs encoding a putative cell wall protein from *Zea mays* and immunological identification of related polypeptides. *Plant Mol. Biol.* **1988**, *11*, 483−493.

Takishima, K.; Suga, T.; Mamiya, G. The structure of Jack Bean urease. *Eur. J. Biochem.* **1988**, *175*, 151−165.

Thomas, J. T.; Cresswell, C. J.; Rash, B.; Nicolai, H.; Jones, T.; Solomon, E.; Grant, M. E.; Boot-Handford, R. P. The Human collagen X gene. *Biochem. J.* **1991**, *280*, 617−623.

Tzen, J. T. C.; Lie, G. C.; Huang, A. H. C. Characterization of the charged components and their topology on the surface of plant seed oil bodies. *J. Biol. Chem.* **1992**, *267*, 15626−15634.

Tzen, J. T. C.; Cao, Y.-z.; Laurent, P.; Ratnayake, C.; Huang, A. H. C. Lipids, proteins, and structure of seed oil bodies from diverse species. *Plant Physiol.* **1993**, *101*, 267−276.

Watt, K. W. K.; Takagi, T.; Doolittle, R. F. Amino acid sequence of the β chain of human fibrinogen: homology with the γ chain. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 1731−1735.

Williams, T.; Marumo, K.; Waite, J. H.; Henkens, R. W. Mussel glue protein has an open conformation. *Arch. Biochem. Biophys.* **1989**, *269*, 415−422.

Xu, M.; Lewis, R. V. Structure of a protein superfiber: spider dragline silk. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 7120−7124.

Zhang, S.; Holmes, T.; Lockshin, C.; Rich, A. Spontaneous assembly of a self-complementary oligopeptide to form a stable macroscopic membrane. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 3334−3338.

Received for review July 25, 1995. Accepted July 15, 1996.[⊗] Names are necessary to report factually on available data; however, the USDA neither guarantees nor warrants the standard of the product, and the use of the name by the USDA implies no approval of the product to the exclusion of others that may also be suitable.

JF9504778

[⊗] Abstract published in *Advance ACS Abstracts,* September 1, 1996.