Seed Protein Topology: Molecular Profiles

John A. Rothfus* and James W. Hagemann

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

Corn, wheat, and soybean seed proteins were examined for potentially useful regularity, size, and polarity. Volume and amphiphilicity profiles evidence analogies to spider silk and human collagen, but seed proteins are 28% larger and 3.6 times more polar per residue. Wheat proteins are more uniform than silk; corn zeins less polar than collagen. A soybean cell wall protein is the most uniform (±5 Å^3); a corn cell wall peptide is the smallest (92 ± 9 Å^3) and least polar (2 ± 18) per residue. Selected parts appear to be more useful than whole molecules. The α subunit of soybean β^- conglycinin concentrates half its acidic residues in the N-terminal third of the molecule and 90% of its hydrophobic residues in the remainder. Corn glutelin and wheat gliadins contain lengthy (7.6–16.3 kDa) repetitive segments. Signal peptides share similar volumes with the proteins but are more than twice as hydrophobic.

INTRODUCTION

Technological change constantly challenges agriculture to examine anew the chemical functionality and utilization of annual commodities. The development of biofuels based on fermentable polysaccharides and vegetable oils thus focuses attention on coproducts whose value could have significant impact on biofuel economics.

Large quantities of seed polysaccharides and proteins are relatively unaffected by starch fermentation or oil extraction. Proteins from spent grains and seed meals consequently find wide use in feed or, after further processing, in food. Probably less than 25% of production enters nonfood use, where added value might exceed that of feed.

At one time, seed proteins enjoyed brief commercial acceptance in textiles (Rebenfeld, 1988). More often, due to their ability to form tough dried films, they are employed in adhesives and coatings (Pocius, 1991). Sensitivity to moisture and structural complexity, however, tend to complicate their utilization. Most of these natural polymers remain ill-suited to tasks currently performed by petrochemical polymers even though their natural functions may be analogous to applications of petroleum-based engineering materials.

Physical properties of industrial materials arise as much from weakly bonded interactions of regular molecular surfaces as from covalent structure. Seed protein preparations lack such uniformity, and there is little information on how they should be modified to achieve specific polymeric properties. Nonetheless, recent experience with self-assembling proteinaceous systems from mammalian (Ghosh and Stroud, 1991) and microbial (McGrath et al., 1992; Zhang et al., 1993) sources demonstrates control of polymeric behavior by peptide structure selection. Furthermore, genetic programs attest to the manageability of seed components and promise to ultimately deliver desired structures directly from nature.

Accordingly, it is important to characterize seed proteins in terms of the surfaces they can contribute to macromolecular interactions and to understand how these surfaces can be regulated by physical, chemical, and genetic means. Fortunately, the necessary detailed chemical knowledge of plant protein structure is accumulating rapidly. Herein, we report some interesting features of molecular volume and amphiphilic character already apparent in structures representative of seed proteins in corn, wheat, and soybean.

* Author whom correspondence should be addressed.

EXPERIMENTAL PROCEDURES

Protein volume and amphiphilic character profiles were computed from amino acid sequence data by a moving window analysis as described in detail by Rose et al. (1985). Hydrated molecular volumes and amphiphilicities are summations of individual amino acid dimensions (Chen and Bendedouch, 1986) and amphipathic characteristics (Eisenberg et al., 1984) tabulated by others. Computed volumes include an additional 25 Å^3 to account for N-terminal amino and C-terminal carboxyl structures (Reynolds and McCaslin, 1985), but molecular amphilicities, which are unitless relative terms, remain unmodified for terminal charge. "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, 110; Arg, 176.

Representative protein sequences, generally from nucleic acid sequence data and without signal peptides, were scanned at a window width of seven residues (i.e., index residue ± three residues). Larger or smaller window widths diminished profile details. N- and C-terminal windows were necessarily narrower, but characteristics averaged per residue in these regions were not unusually different. Signal peptide sequences, where known, were analyzed independently.

Seed proteins that were examined are listed in Table 1 along with designations by which they are indentified in this work.

A general frame of reference was established by comparable analysis of like-sized polypeptide sequences that were generated by random selection of residues from an unweighted amino acid population.

RESULTS AND DISCUSSION

Unlike hydrocarbon polymers, whose uniform structures and regular surfaces promote self-association, proteins often contain uneven distributions of residues with conflicting polarity and bulk. These internal contradictions present a variety of surfaces to attached intramolecular segments and, externally, to other molecules. Nevertheless, amino acid residues range in size from glycine, ca. 66 Å^3, to tryptophan, ca. 238 Å^3 (Chen and Bendedouch, 1986). This range and the average per residue volume for amino acids, ca. 149 ± 35 Å^3, are similar to volumes calculated from data of Reynolds and McCaslin (1985) for monomeric

Table 1. Proteins

designation	protein	residues	reference
	Corn		
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)
MHIGLY	Gly-rich protein (CHEM2-GRP)	155	Didierjean et al. (1992)
	Wheat		
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)
	Soybean		
SG2AGLY	G2 acidic glycinin	278	Nielsen et al. (1989)
SG2BGLY	G2 basic glycinin	185	Nielsen et al. (1989)
SABCNGLY	β-conglycinin, α subunit	583	Sebastiani et al. (1990)
SHIPRO	Pro-rich protein (SbPRP2)	203	Hong et al. (1990)
	Miscellaneous		
SILK2	*N. clavipes* Spidroin 2	627	Hinman and Lewis (1992)
BCASEINB	bovine βA^2-casein	209	Ribadeau-Dumas et al. (1972)
HCOLA1X	human α1(X)collagen	680	Thomas et al. (1991)
RBISCOL	spinach ribulose bisphosphate carboxylase, large subunit	475	Zurawski et al. (1981)
JBURASE	jack bean urease	840	Riddles et al. (1991)

units of polyethylene, 55 Å^3; polypropylene, 85 Å^3; and polystyrene, 190 Å^3.

Furthermore, proteins are already known to bind together readily in adhesives, and certain seed protein mixtures, especially corn zein and wheat gluten, can be reconstituted into fibers and highly extensible gas-impervious films. Such transformations commonly enhance texture and other qualities of processed foods. Presumably, if seed proteins were more like hydrocarbon polymers, they might satisfy much broader markets. Thus, our attention focused on smaller than average residue volumes and negative amphiphilicities. Interestingly, these features are more apparent in seed protein signal peptides than in the associated proteins that compose normal mature seed.

Signal Peptides. Eleven of the seed proteins examined have signal peptide sequences. In this respect, they differ from nonexcreted proteins such as chloroplast rubisco (McIntosh et al., 1980) or jack bean urease (Riddles et al., 1991), which remain within specialized cells or organelles. The basic subunit of glycinin (SG2BGLY) also lacks a comparable signal sequence because it arises from post-translational editing of a larger transported glycinin gene product (Nielsen et al., 1989).

All of the examined signal peptide sequences are essentially the same length (Table 2), ca. 20 residues, which is not inconsistent with signal peptides of other transported proteins (Wickner, 1979). With an average per residue volume of 139 ± 14 Å^3, the signal peptides are nearly 10% smaller than random sequences and about the same size per residue as the mature seed proteins. Several, however, exhibit slightly more variation in volume than do their mature counterparts (compare Tables 2 and 3).

More important in terms of hydrocarbon polymer characteristics and potential chemical stability, the seed signal peptides are all substantially more hydrophobic than most comparably sized segments of the mature proteins. They are also noticeably deficient in residues with carboxylic or amide side chains that can enhance acid instability. Only 13 different segments, representing 150 of some 3500 residues in the mature seed proteins, are more hydrophobic (−28/residue) than the signal peptides (−20/residue). Interestingly, these especially hydrophobic segments tend to occur near the center or in C-terminal portions of the seed proteins, which is analogous to the distribution of hydrophobicity in the signal peptides. Relatively simple composition and concentrated hydrophobicity make the signal peptides and hydrophobic segments of the mature proteins logical candidates for films, fibers, and gels if they can be collected or their sequences replicated in plant or microbial host molecules.

Signal peptide hydrophobicity concentrates to essentially the same extent near the center of each peptide (Figure 1) even though they are composed of different amino acids. Profiles of the corn and wheat signal peptides thus appear remarkably similar from a functional standpoint even though Soll and co-workers (Rafalski et al., 1984) have reported essentially no homology between polynucleotide sequences for wheat gliadin and zein signal peptides.

The WLWGTENA signal peptide is both least polar and smallest near its center, unlike most of the others, in which volume declines near the C terminus (Figure 1). Though many of the profiles are similar, none are identical. The soybean cell wall protein signal peptide presents an especially interesting contrast to the others, which emphasizes the presumably quite different biological role of the mature SHIPRO that occurs more in the seed coat than in the cotyledon (Hong et al., 1990).

Mature Proteins. Molecular volumes and amphiphilicities given in Table 3 show the extent to which mature seed proteins differ from signal peptides and from each other. Among the sequences examined, most are about 5% smaller overall than random sequences that have the same degrees of polymerization.

When compared in terms of average residue volume and volume variation, none of the seed proteins are as small as SILK2, but several, including the wheat proteins and SHIPRO, appear to be more uniform. Unfortunately, none of the seed proteins combines small uniform volume with hydrophobic character to the extent seen in SILK2 (Hinman and Lewis, 1992) and HCOLA1X (Thomas et al., 1991). Silk is more hydrophobic and even smaller than collagen, but throughout its central 463-residue collagenous sequence the human connective tissue component is more uniform than the insect dragline protein. Only MGTEL approaches the per residue dimensions of these proven useful proteins, yet it is still 19% larger and 12% more polar per residue than collagen.

That another silk molecule (Xu and Lewis, 1990) examined in this work was much smaller (101 ± 11 Å^3) and much less polar (−4 ± 12) on average raises a question about tolerances within which the topological properties of plant proteins should be regulated to ensure useful materials. Both silk and collagen, not unlike plastics, are

Table 2. Computed Signal Peptide Volumes and Amphiphilicities

signal peptide	residues	vol[a] (Å)	% of random[b]	vol/residue[c] (Å)	amphiphilicity[d]	% of random[b]	amphiphilicity/residue[c]
MZEIN15	20	2754	91	138 ± 26	-474	-133	-23 ± 16
MZEIN19	21	2793	88	133 ± 12	-461	-123	-22 ± 12
MZEIN22	21	2856	90	136 ± 12	-366	-98	-17 ± 17
MGTEL	19	2516	88	132 ± 18	-356	-105	-19 ± 17
WABGLI	20	2878	96	144 ± 14	-357	-100	-18 ± 22
WGGLI	19	2696	94	142 ± 15	-465	-137	-25 ± 12
WLWGTENA	23	3200	92	139 ± 17	-343	-84	-15 ± 18
WHWGTEN	21	2936	93	140 ± 11	-416	-111	-20 ± 25
SG2AGLY	18	2598	96	144 ± 9	-440	-137	-24 ± 10
SABCNGLY	22	3220	97	146 ± 14	-283	-72	-13 ± 26
SHIPRO	22	3022	91	137 ± 7	-551	-140	-25 ± 13
RANDOM[e]				151 ± 12			18 ± 18

[a] Unadjusted for N- and C-terminal atoms (25 Å^3). [b] Percent of volume/amphiphilicity for a like number of random residues. [c] Mean residue volume/amphiphilicity ± average deviation from mean. [d] Unadjusted for N- and C-terminal polarity. [e] Averages from 100 sequences generated by random selection of residues.

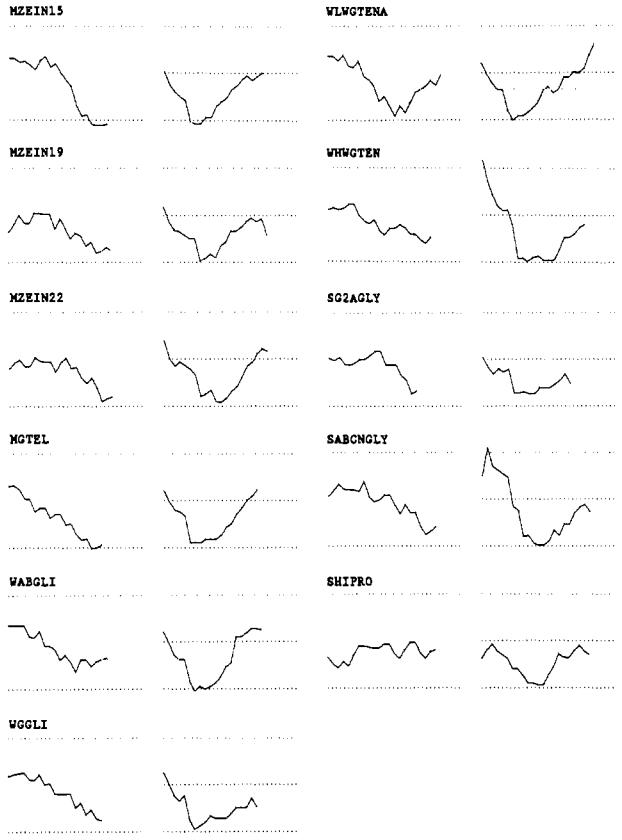

Figure 1. Volume and amphiphilic profiles of signal peptides from selected seed proteins. Properties are displayed, left to right, from N terminus to C terminus. Volumes are on the left in each pair; dotted lines indicate 100 (lower) or 200 Å^3 (upper). Amphiphilicities, on the right in each pair, are relative; dotted lines indicate zero (center), -50 (lower), and +50 (upper).

derived from soluble/dispersed precursors. Novel applications of seed proteins could easily involve analogous transformations from fluid to static systems. Proper mixtures of bulky and polar sequences with small uniform and nonpolar segments could prove advantageous in fabricating goods from plant proteins. Accordingly, it is interesting that lengthy noncollagenous sequences, which flank repetitive collagenous regions in prepropeptides of human α1(V) collagen (Takahara et al., 1991), are essentially the same size per residue as MGTEL but 70–80% more polar.

The silk and collagen molecules, which aggregate readily (Thomas et al., 1991), represent higher degrees of amino acid polymerization than seen thus far in the corn proteins. MGTEL molecules joined chemically to produce higher molecular weights would be less dense and somewhat more polar than either silk or collagen, but they might produce analogous aggregates.

Among the larger seed proteins, WHWGTEN exhibits essentially the same uniform residue volume as MGTEL and even less polarity variation. Consistent with its importance in breadmaking (Ewart, 1990; Miles et al., 1991; Halford et al., 1992), this seed protein can easily form long viscous strands through cross-linkable cysteine residues at each end of the molecule. (In this regard, it is interesting that cysteine residues in the other wheat and corn proteins are distributed quite differently.) WHWGTEN's relatively high average residue amphiphilicity, however, suggests that solid materials made from this protein would hydrate some 3 times more easily than comparable materials from MGTEL.

Chemical modifications and choice of application can, of course, accommodate shortcomings in the native proteins. By such handling, caseins and soy proteins have become common ingredients of formulated adhesives for interior-grade plywood (Pocius, 1991), where they seldom encounter extreme moisture. The volume and amphiphilicities of JBURASE, an unusually large seed protein, or RBISCOL from green leaves, one of the world's most abundant proteins, suggest they would perform equally well given similar advantage.

Plant proteins other than those in seed endosperm or cotyledons certainly deserve consideration, especially if they can be suited to novel applications. Two examples of unusually repetitive protein structures, MHIGLY and SHIPRO, are included here for comparison. One of these, SHIPRO, is remarkably uniform with respect to both volume and amphiphilicity. Though treated here as a high-proline sequence, this protein is analogous in both structure and occurrence to a group of hydroxyproline-rich glycoproteins known as extensins (Cassab and Varner, 1988). The gene for this soybean protein is highly expressed in the root and seed coat (Hong et al., 1990). Quite likely it is hydroxylated and glycosylated like other extensins if it survives in the mature seed. Hydroxylation, of course, would increase the protein's volume and amphiphilicity from levels given in Table 2. Although its uniform structure would not change markedly, a hydroxylated SHIPRO probably would find stable conformations quite different from those preferred by the nonhydroxylated version.

Conclusions drawn from the quantitative analyses of whole proteins are borne out and visualized easily in detailed profiles (Figures 2–5). SHIPRO (Figure 5) is

Table 3. Computed Protein Volumes and Amphiphilicities

protein	vol[a] (Å)	% of random[b]	vol/residue[c] (Å^3)	amphiphilicity[d]	% of random[b]	amphiphilicity/residue[c]
MZEIN15	22193	92	139 ± 16	1076	38	7 ± 17
MZEIN19	32389	98	148 ± 11	496	13	2 ± 12
MZEIN22	35256	96	144 ± 11	658	15	3 ± 15
MGTEL	28011	91	137 ± 10	1752	48	9 ± 18
MHIGLY	18819	81	121 ± 21	2495	90	16 ± 22
WABGLI	43819	99	149 ± 9	6580	126	22 ± 18
WGGLI	40940	98	148 ± 8	4833	98	15 ± 15
WLWGTENA	41957	98	148 ± 9	5123	101	18 ± 16
WHWGTEN	83899	90	135 ± 9	15909	144	26 ± 11
SG2AGLY	39341	94	141 ± 14	7668	155	28 ± 26
SG2BGLY	25915	93	140 ± 11	2395	73	13 ± 20
SABCNGLY	84838	97	146 ± 10	17619	170	30 ± 26
SHIPRO	30354	99	149 ± 5	4182	116	21 ± 10
SILK2	6772	72	108 ± 11	213	2	1 ± 12
BCASEINB	30548	97	146 ± 9	2233	60	11 ± 17
HCOLA1X	53266	76	115 ± 8	3644	44	8 ± 13
RBISCOL	66453	93	140 ± 12	6070	72	13 ± 18
JBURASE	115074	88	137 ± 11	9045	59	11 ± 17
RANDOM[e]			151 ± 12			18 ± 18

[a] Includes 25 Å^3 for N- and C-terminal atoms. [b] Percent of volume/amphiphilicity for a like number of random residues. [c] Mean residue volume/amphiphilicity ± average deviation from mean. [d] Unadjusted for N- and C-terminal polarity. [e] Averages from 100 sequences generated by random selection of residues.

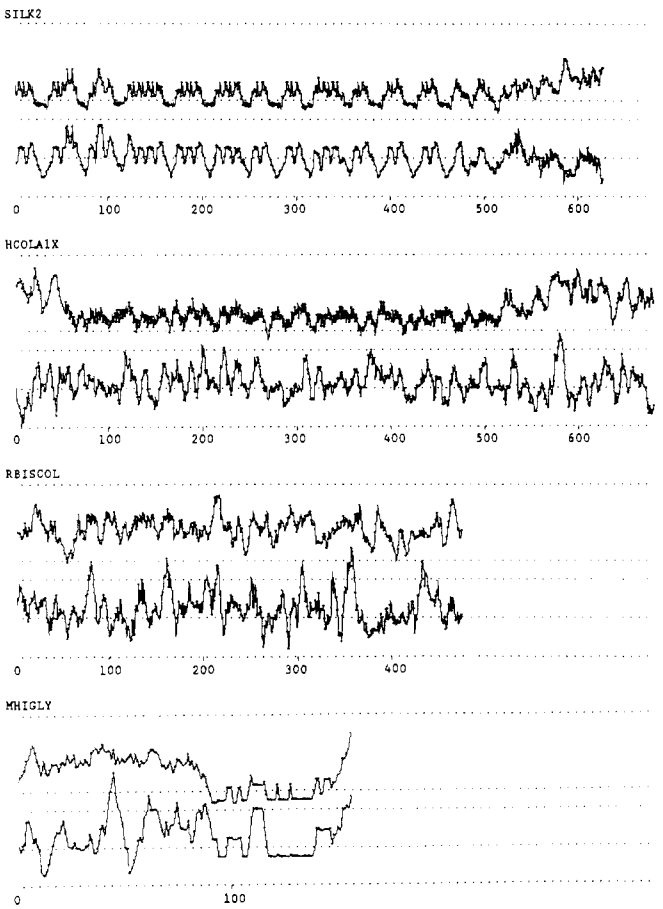

Figure 2. Volume and amphiphilic profiles of selected nonseed proteins. Upper curve in each pair is volume; lower, amphiphilicity. Properties are displayed as in Figure 1. Tick marks indicate 100 residues. HCOLA1X profile includes N-terminal and C-terminal noncollagenous sequences of 56 (includes 18-residue signal peptide) and 161 residues, respectively.

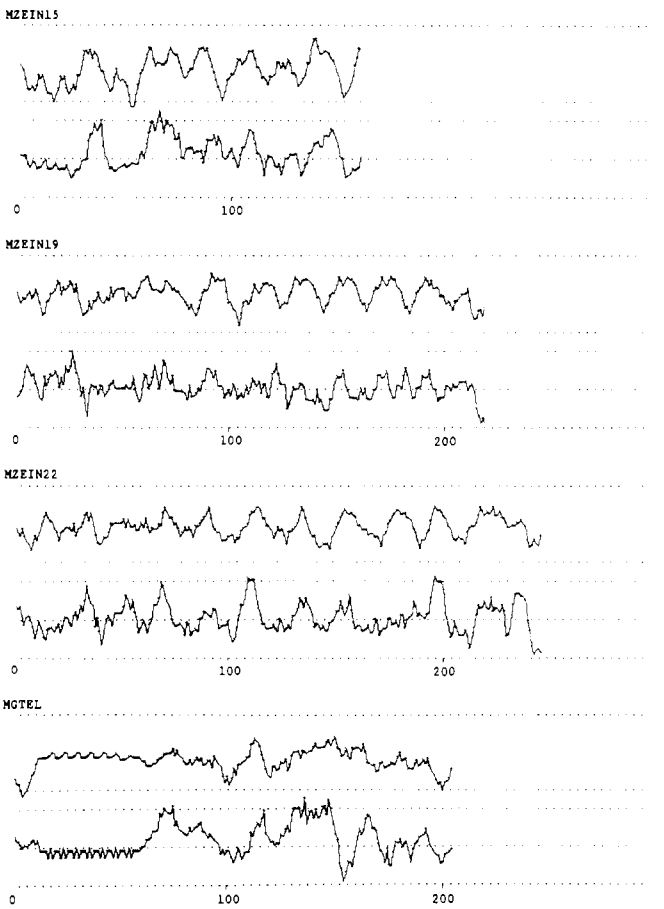

Figure 3. Volume and amphiphilic profiles of selected corn proteins.

indeed quite uniform. Silk (SILK2) and collagen (HCOLA1X) (Figure 2) are both relatively small and uniform throughout the bulk of their sequences, especially compared to RBISCOL (Figure 2), in which polarity fluctuates substantially and rather regularly throughout the molecule.

A general correspondence between comparisons allowed by these profiles and those based on molecular values (Table 3) demonstrates the utility of computations. The profile for MHIGLY (Figure 2), however, shows how whole protein computations can deceive. This protein combines quite different segments. For the complete protein (residues 1–155), per residue volumes and amphiphilicities average 121 ± 21 and 16 ± 22, respectively. For residues

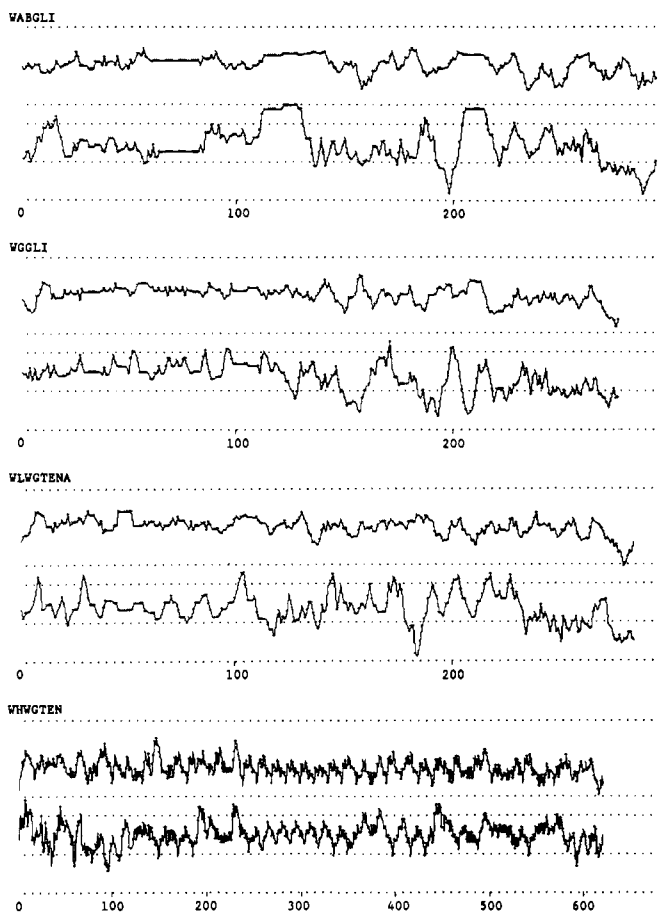

Figure 4. Volume and amphiphilic profiles of selected wheat proteins.

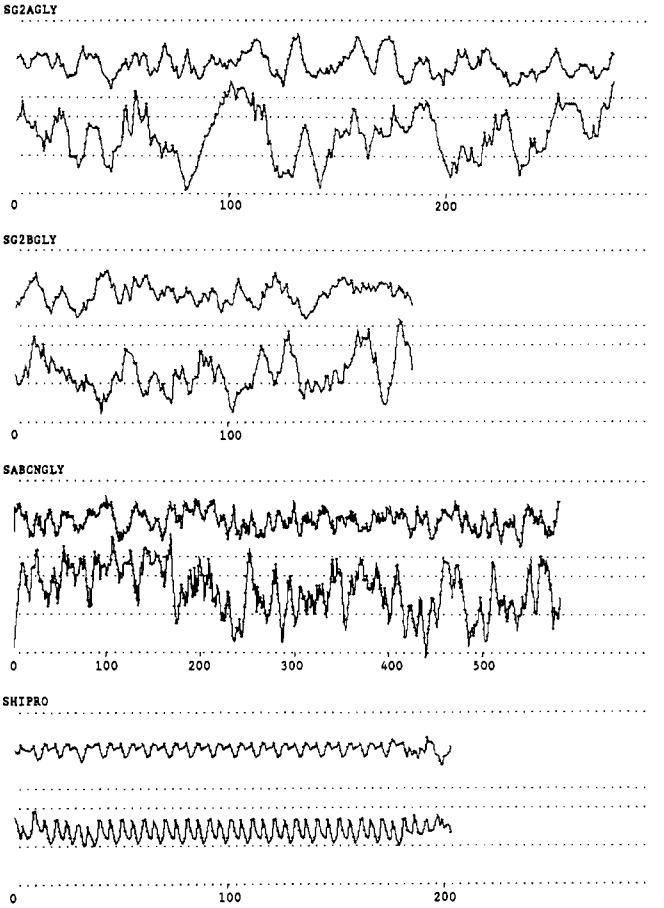

Figure 5. Volume and amphiphilic profiles of selected soybean proteins.

1–90, however, the same properties are 138 ± 8 and 22 ± 22 and for residues 90–140, 92 ± 9 and 2 ± 18.

Uneven distributions and nonrandom concentrations of size and polar or nonpolar functionalities are also apparent in other profiles. With some of these proteins, substantial change to smaller more uniform volume and reduced amphiphilicity would be accomplished if segments of the proteins were utilized rather than the whole native molecules. For example, the N-terminal quarter of SABCNGLY (Figure 5) is extremely polar. Less than a third of the molecule contains more than half of the acid residues, and concentrations of aspartic and glutamic acid are unusually high near residue 150.

With respect to the presence of acidic amino acids in SABCNGLY, it is interesting to compare the amphiphilicity of SABCNGLY with that of BCASEINB (Table 2) and then recall (a) that casein glues are essentially calcium salts of proteins that contain rather high concentrations of dicarboxylic acids and (b) that soy protein makes better plywood glue than casein (Pocius, 1991).

Removal of the polar portion from SABCNGLY, perhaps by two-phase hydrolysis, would leave three-fourths of the molecule with over 90% of its hydrophobic residues. Polarity would be reduced by one-third, and solubility should be altered drastically without marked effect on volume characteristics.

Structural "editing" of the seed proteins by specific enzymic hydrolysis similar to work already undertaken by Popineau and co-workers (Masson et al., 1989; Popineau et al., 1990) could likewise produce beneficial changes. Digestion of the wheat gliadins, WABGLI and WGGLI (Figure 4), and corn glutelin, MGTEL (Figure 3), with trypsin to specifically break the peptide chains at lysine or arginine residues should yield highly repetitive segments of different size but generally more uniform dimension and polarity. The largest fragment, residues 1–138 from WGGLI, would have a per residue volume of 152 ± 5 and amphiphilicity of 29 ± 8. These values represent about a 40% improvement in uniformity. In another tryptic peptide, residues 3–89 from WABGLI, the volume would also be 152 ± 5, but amphiphilicity would fall to 20 ± 10. The smallest fragment, residues 1–72 from MGTEL, would have a per residue volume of 137 ± 10 and very low amphiphilicity, 2 ± 11. This fragment's per residue volume would be unchanged from the original MGTEL average, but polarity and variation of polarity would be reduced some 80 and 40%, respectively. In addition, this fragment from MGTEL would have cysteine residues at each end, which should allow for formation of linear disulfide-linked-proteins analogous to those that can form from WHWGTEN molecules.

To the extent that the examined sequences are representative of seed proteins, it is obvious that conditions which reduce their polarity and enhance uniform molecular association must be found if they are to serve material needs currently satisfied by petrochemical polymers. It also appears that the challenge to make useful polymers from such commodities is as much a challenge to recognize and secure useful parts from them as it is to identify and enrich single protein constituents.

LITERATURE CITED

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.

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, 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. *MGG, Mol. Gen. Genet.* **1989**, *216*, 81–90.

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.

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.

Ewart, J. A. D. Comments on Recent Hypotheses for Glutenin. *Food Chem.* **1990**, *38*, 159–169.

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.

Ghosh, P.; Stroud, R. M. Ion Channels Formed by a Highly Charged Peptide. *Biochemistry* **1991**, *30*, 3551–3557.

Halford, N. G.; Field, J. M.; Blair, H.; Urwin, P.; Moore, K.; Robert, L.; Thompson, R.; Flavell, R. B.; Tatham, A. S.; Shewry, P. R. Analysis of HMW Glutenin Subunits Encoded by Chromosome 1A of Bread Wheat (*Triticum aestivum L.*) Indicates Quantitative Effects on Grain Quality. *Theor. Appl. Genet.* **1992**, *83*, 373–378.

Hinman, M. B.; Lewis, R. V. Isolation of a Clone Encoding a Second Dragline Silk Fibroin. *J. Biol. Chem.* **1992**, *267*, 19320–19324.

Hong, J. C.; Nagao, R. T.; Key, J. L. Characterization of a Proline-rich Cell Wall Protein Gene Family of Soybean. *J. Biol. Chem.* **1990**, *265*, 2470–2475.

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.

Masson, P.; Popineau, Y.; Pineau, F. Limited Hydrolysis of Gamma-gliadin by Pepsin: Fractionation and Partial Characterization of Large Polypeptides. *Lebensm. Wiss. Technol.* **1989**, *22*, 157–163.

McGrath, K. P.; Fournier, M. J.; Mason, T. L.; Tirrell, D. A. Genetically Directed Syntheses of New Polymeric Materials. Expression of Artificial Genes Encoding Proteins with Repeating $-(AlaGly)_3ProGluGly-$ Elements. *J. Am. Chem. Soc.* **1992**, *114*, 727–733.

McIntosh, L.; Poulsen, C.; Bogorad, L. Chloroplast Gene Sequence for the Large Subunit of Ribulose Bisphosphatecarboxylase of Maize. *Nature* **1980**, *288*, 556–560.

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.

Nielsen, N. C.; Dickinson, C. D.; Cho, T-J.; Thanh, V. H.; Scallon, B. J.; Fischer, R. L.; Sims, T. L.; Drews, G. N.; Goldberg, R. B. Characterization of the Glycinin Gene Family in Soybean. *Plant Cell* **1989**, *1*, 313–328.

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

Popineau, Y.; Masson, P.; Pineau, F.; Guary, J. C. Limited Hydrolysis of a Gamma-type Gliadin by Chymotrypsin: Isolation of Specific Sequence Domains. *Lebensm. Wiss. Technol.* **1990**, *23*, 474–480.

Rafalski, J. A.; Scheets, K.; Metzler, M.; Peterson, D. M.; Hedgcoth, C.; Soll, D. G. Developmentally Regulated Plant Genes: the Nucleotide Sequence of a Wheat Gliadin Genomic Clone. *EMBO J.* **1984**, *3*, 1409–1415.

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

Reynolds, J. A.; McCaslin, D. R. Determination of Protein Molecular Weight in Complexes with Detergent without Knowledge of Binding. *Methods Enzymol.* **1985**, *117*, 41–53.

Ribadeau-Dumas, B.; Brignon, G.; Grosclaude, F.; Mercier, J-C. Primary Structure of Bovine β-Casein. *Eur. J. Biochem.* **1972**, *25*, 505–514.

Riddles, P. W.; Whan, V.; Blakeley, R. L.; Zerner, B. Cloning and Sequencing of a Jack Bean Urease-encoding cDNA. *Gene* **1991**, *108*, 265–267.

Rose, G. D.; Gierasch, L. M.; Smith, J. A. Turns in Peptides and Proteins. *Adv. Protein Chem.* **1985**, *37*, 1–109.

Sebastiani, F. L.; Farrell, L. B.; Schuler, M. A.; Beachy, R. N. Complete Sequence of a cDNA of α Subunit of Soybean β-Conglycinin. *Plant Mol. Biol.* **1990**, *15*, 197–201.

Takahara, K.; Sato, Y.; Okazawa, K.; Okamoto, N.; Noda, A.; Yaoi, Y.; Kato, I. Complete Primary Structure of Human Collagen α1(V) Chain. *J. Biol. Chem.* **1991**, *266*, 13124–13129.

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.

Wickner, W. The Assembly of Proteins into Biological Membranes: the Membrane Trigger Hypothesis. *Annu. Rev. Biochem.* **1979**, *14*, 23–45.

Xu, M.; Lewis, R. V. Structures 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.

Zurawski, G.; Perrot, B.; Bottomley, W.; Whitfield, P. R. The Structure of the Gene for the Large Subunit of Ribulose 1,5-bis-Phosphate Carboxylase from Chloroplast DNA. *Nucleic Acids Res.* **1981**, *9*, 3251–3270.

Received for review October 4, 1993. Accepted January 25, 1994.[⊗] The mention of firm names or trade products does not imply that they are endorsed or recommended by the U.S. Department of Agriculture over other firms or similar products not mentioned.

[⊗] Abstract published in *Advance ACS Abstracts*, March 1, 1994.