Journal of Agricultural and Food Chemistry

SEPTEMBER 1990
VOLUME 38, NUMBER 9

REVIEWS

Molecular Approaches to Improving the Nutritional and Functional Properties of Plant Seeds as Food Sources: Developments and Comments

Benito O. de Lumen

Department of Nutritional Sciences, University of California, Berkeley, California 94720

During seed development, proteins, carbohydrates, lipids, vitamins, minerals, and nonnutrients are laid down, leading to the unique chemical composition of plant seeds that determines their nutritional and functional properties. Understanding the molecular genetic mechanisms by which photosynthetic assimilates are partitioned into these components will allow the design of genetic engineering strategies to improve the usefulness of seeds as food sources. Utilization problems of legumes and cereals and their possible causes are summarized in Table I.

The application of recombinant DNA techniques and related methodologies to plants has opened up the potential to improve agronomic characters, food processing traits, and food quality properties of plants as food sources. The rapid accumulation of knowledge on gene expression processes in plants in the past 5 years has given us an understanding of structure, function, and organization of plant genes that has brought about some applications. My purposes in this review are to discuss the rationale for improving the nutritional and functional properties of plant seeds through genetic engineering and point out recent developments in molecular biology of seed development that I think are relevant to this goal. Since this is not meant to be a comprehensive review, the reader is referred to more complete reviews on biochemistry and molecular biology of seed proteins written in the past 5 years (Shewry and Miflin, 1984; Higgins, 1984; Croy and Gatehouse, 1985; Gatehouse et al., 1986; Casey et al., 1986; Goldberg, 1986; Goldberg et al., 1989).

In the past 5 years, there has been a rapid accumulation of knowledge on gene expression processes in higher plants (Schell, 1987; Kuhlemeir et al., 1987; Goldberg, 1988; Goldberg et al., 1989). Much of this information has been obtained from gene systems that are intensively studied because they code for abundant gene products, are inducible, and are involved in unique plant processes that are economically important.

Table I. Utilization Problems of Legumes and Cereals and Possible Causes

problems	causes
legumes	
incomplete protein	limiting amounts of methionine and cysteine
incomplete availability of amino acids	unknown
low protein digestibility	nature of protein (?), protease inhibitors (?)
antinutritional factors	protease inhibitors, hemagglutinins, goitrogens, etc.
flatulence	raffinose oligosaccharides, dietary fiber (?), undigestible starch (?)
undesirable "beany flavor"	lipoxygenases
cereals	
low protein content	
incomplete protein	limiting amounts of lysine and, in some cereals, tryptophan, threonine

PLANT SEEDS CONTRIBUTE SIGNIFICANTLY TO THE HUMAN DIET

Seeds of cereals and legumes and their byproducts make up a major part of the human diet. Rice is the staple food of half of the world's population and is the single largest source of protein in the world (Juliano, 1985). Cereals contribute 50% of the per capita energy intake worldwide and 65% in developing countries and Asian centrally planned economies (FAO, 1980). Cereals account for about 45% of the daily per capita protein supply in the world and approximately 63% in developing countries. Although the worldwide production of grain legumes is small relative to cereals, its relative contribution to human nutrition is

0021-8561/90/1438-1779$02.50/0

greater than its relative production. Approximately 700 million people use legumes as an essential part of their diet (Abbott, 1966). Legumes make up 9% of the total plant food production in developing countries and serve as the major source of proteins in the diet of a number of countries (Toenniessen, 1985).

Animal products which have replaced plant foods in human diets in the developed market economies of the world are derived from farm animals. Nonruminants and intensively reared ruminants consume seeds and their byproducts supplemented with protein concentrates processed from oil seeds. In addition, the consumption of vegetable oils and fats, mostly derived from oilseeds, accounts for 4–10% of the daily per capita energy supply in the world (Duffus and Slaughter, 1980). Soybean is the largest source of vegetable oil in the world (American Soybean Association, 1982). Thus, the total contribution of seeds to the human diet, as energy and protein sources, can be as high as 85% of total calories.

The protein content of cereals varies from 7% to 14%, while that of legumes from 20% to 40%. Lysine is the first nutritionally limiting essential amino acid in most cereals; tryptophan is the second limiting amino acid in maize and threonine in other cereals (Eggum and Beames, 1983). The high content of prolamin (alcohol-soluble fraction) in cereals accounts for the low content of essential amino acids. Not only are prolamins very low in lysine, they are also low in threonine, valine, and isoleucine. Rice and oats have a better balance of essential amino acids than other cereals due to a lower content of prolamins (Shewry and Miflin, 1984; Doll, 1984). Legumes are uniformly limiting in methionine and cysteine because their major storage proteins, the globulins (salt-soluble fraction), are low in these amino acids (Gupta, 1983; Duffus and Slaughter, 1980).

IS THERE A NEED TO GENETICALLY ENGINEER LEGUMES AND CEREALS FOR IMPROVED NUTRITIONAL PROPERTIES?

In practice, cereals and legumes are eaten with other foods. The overall protein quality of cereal–legume mixtures is better than that of either protein source alone due to the complementary nature of their amino acid profiles. Cereals are low in lysine and relatively high in methionine and cysteine, while legumes are high in lysine and low in methionine and cysteine. Supplementation of cereals and legumes with crystalline amino acids improves their value to support growth of animals and increases the efficiency of utilization of dietary protein (Bressani and Elias, 1968). Methionine is used as a supplement in the formulations of poultry and swine rations of the highly competitive feed industry where soybean is the main protein source (Abbott, 1966). Little practical use of supplementation for planning human diets has been made, although clinical studies with human beings have demonstrated their effectiveness (Bresanni and Elias, 1968).

Since cereals are the major sources of protein worldwide, increasing the protein content alone will improve the protein intake of a large number of people, especially those in developing countries, where protein–calorie malnutrition is prevalent. For instance, an increase of 3% in rice protein without a decrease in grain yield will add an additional 12 million tons of protein to the diet of rice-eating people, based on 411 million tons of total rice production in 1981–1982 (Juliano, 1985). Improving the protein quality of the rice grain will bring additional significant benefits because the N digestibility and N retention of rice diets are low (Roxas et al., 1979).

Because the protein contents of legumes are already high, improving the protein quality is the more useful goal. Why is there a need for a nutritionally balanced protein in legumes when supplementation and complementation work? Several reasons can be given. First, in supplemented foods and feeds, leaching of free methionine during processing leads to losses and subsequent bacterial fermentation, resulting in objectionable odors and flavors. Acceptability is perhaps the biggest problem in supplementation of human diets with crystalline methionine. Addition of minute amounts of methionine to the diets of nontechnical, village societies may also be a problem. Second, a large number of people who are dependent on root crops as their main sources of energy will benefit from "high protein quality legumes" because the protein content of root crops is extremely low and their amino acid profile is not complementary with that of legumes. For instance, cassava, a major root crop, is the basic food for 200–400 million inhabitants of the tropics (Odigboh, 1983). Finally, the nutritionally complete protein of a genetically engineered bean could be added to the list of advantages of legumes as food such as "no cholesterol", dietary fiber source, possible beneficial effects of plant proteins on blood lipid status (Bodwell, 1987), and the capability of legumes to produce the greatest amount of protein per unit of land (Butz, 1973). These characteristics could be used as marketing attributes to encourage more people to eat legumes as sources of protein.

There are reports emphasizing that protein from plant sources meets the amino acid requirements of adults and that protein from some plant sources meets those of young children (Bodwell, 1987). However, recent reassessment of the requirements for essential amino acids in healthy adults led to the conclusion that the requirements for lysine, leucine, valine, and threonine are likely to be 2–3 times higher than those reported in the 1985 WHO/FAO/UNU report (FAO/WHO/UNU, 1985; National Research Council, 1988).

FUNCTIONAL PROPERTIES OF PROTEINS

Functional properties of proteins are broadly defined as those properties other than nutritional attributes that affect its utilization. These include sensory properties (flavor, odor, color, texture), water interacting properties (hydration, dispersibility, solubility), surface active properties (emulsification, foaming, adsorption), and rheological properties (gelation, texturization) (Nakai and Powrie, 1981). The overall functional property of a food system is the result of composite properties of individual protein components as they interact with one another and with nonprotein components. Clearly, the functionality of proteins is very important in foods, yet very little is known about its molecular basis. Computer modeling of structure–function relationships in proteins, mainly enzymes, is increasingly being used to design proteins rationally for specific functions (Blundell and Sternberg, 1985; Lin, 1986; Hol, 1987). Such an approach could be used to determine how site-specific mutations would affect functional properties of proteins. Since only minute amounts of proteins are currently available from molecular techniques, microsystems are needed to test the effects of structural changes on protein functionalities. Such information, combined with a knowledge of the critical sequences in seed proteins needed for posttranslational processing and packaging, would allow the rational design of proteins with specific functional attributes without adverse effects on seed biology.

MOLECULAR BIOLOGY OF SEED PROTEINS

The seed proteins are classified into two major classes, namely, storage proteins and the "housekeeping proteins". The storage proteins, which make up a considerable part of the total protein, provide the germinating seed with carbon, nitrogen, sulfur, and other nutrients, while the housekeeping proteins are necessary for the normal metabolism of the seed.

Because of their economic importance, the storage proteins of legumes have been intensively studied. Logically, the seed proteins became the focus of studies on plant gene expression because they are abundant gene products that accumulate rapidly at midmaturation stage of seed development and because much is known about their biochemical and chemical properties as food proteins. The seed proteins also offer to basic plant biologists a model for temporal and tissue-specific regulation during seed development. Consequently, the first plant gene cloned and sequenced completely was that of the French bean phaseolin and established the presence of introns in plant genes (Slightom et al., 1983; Sun et al., 1981).

1. Seed Development Begins after a Double-Fertilization Process Unique to Angiosperms (Raven et al., 1987). One of the two sperms in the pollen tube unites with the egg cell in the ovule to form the diploid zygote that develops into the embryo. The other sperm unites with the two polar nuclei to form a triploid endosperm which provides nourishment to the developing seed as well as the germinating seedling. In legumes, the endosperm is absorbed by the developing embryo while the fleshy food-storing cotyledons develop. Thus, in legumes, the seed differentiates into two organ systems, the axis and the cotyledons, which have different developmental pathways. The axis develops into the roots, the stem, and the leaves, while the cotyledon is a terminally differentiated organ that provides the nutrients to the germinating seedling. In most legumes, the cotyledon makes up at least 80% of the mature seed weight. In cereals, the starchy endosperm constitutes the major part of the seed.

2. The Seed Undergoes Major and Distinct Physiological Processes during Development (Goldberg et al., 1989; Spencer and Higgins, 1982). The early stages are characterized by cell division followed by a phase of cell expansion in which reserve synthesis and deposition occur. Starch, proteins, and lipids are synthesized at the cell expansion stage in which DNA content also increases. The last stage is characterized by a decrease in RNA and protein synthesis, loss of water, and finally dormancy.

3. The Synthesis of a Number of Seed Storage Proteins Is Characterized by Posttranslational Processing Including Glycosylation, Endoproteolytic Processing, and Intracellular Transport (Higgins, 1984; Muntz, 1989). The primary translation products have "signal sequences" that are cleaved as the proteins enter the lumen of the endoplasmic reticulum. Once inside the endoplasmic reticulum, the polypeptides may undergo proteolytic cleavage and, in most cases, are transferred via the Golgi apparatus into the membrane-bound protein bodies. Inside the protein bodies, the proteins aggregate, evidently determined by the primary structure of the protein. A structural model for the 19- and 22-kDa zein proteins has been proposed (Argos et al., 1982), implying that certain critical sequences need to be preserved for proper packaging in addition to targeting signals for intracellular sorting.

4. The Major Storage Proteins in Legumes and Cereals Are Coded by Several Homologous Multigene Families That Vary in Size, Organization, and Chromosomal Location (Casey et al., 1986; Goldberg et al., 1989; Osborn, 1988; Elliston et al., 1988; Gatehouse et al., 1988; Kreis et al., 1985; Higgins, 1984). In peas, the 50-kDa subunit of vicilin (7S) is coded by at least 3 families with up to 6 members within each family, indicating that as many as 18 genes code for this subunit. At least four homologous genes code for the 11S proteins (legumin) of pea (Croy et al., 1982). The legumin (glycinin) subunits of soybean are coded by five genes belonging to two subfamilies (Nielsen et al., 1989). The three genes belonging to one subfamily (Gy1, Gy2, Gy3) occupy two distinct domains in the soybean genome. Two subfamilies of legumin-like proteins are also found in pea (Domoney and Casey, 1985) and the bean *Vicia faba* (Baumlein et al., 1986). The 7S soybean proteins (β-conglycinin) are coded by 2 distinct genes with at least 15 members clustered in six regions in the soybean genome (Harada et al., 1989). Similarly, the 7S proteins (phaseolin), which are the major storage proteins in *Phaseolus vulgaris*, are encoded by two unique gene families made up of seven to nine members (Slightom et al., 1985; Talbot et al., 1984).

The prolamin (zein) family in maize is coded by 3–10 genes for each family, and since there may be 10 families, there could be as many as 30–100 zein genes in maize (Burr et al., 1982; Hagen and Rubinstein, 1981; Pedersen et al., 1982; Viotti et al., 1982; Wienand and Feix, 1980). The prolamin genes characterized to date do not contain introns, unlike other plant genes (Kreis et al., 1985). Glutelin, the major rice protein, is encoded by a small multigene family (Takaiwa et al., 1987a). In contrast with the maize prolamin genes, the rice glutelin gene has three introns (Takaiwa et al., 1987b).

The genes for the 7S proteins from different species of legumes show remarkable homologies, suggesting a common ancestor (Casey et al., 1986). Likewise, the genes coding for the 11S proteins from different species have strong homologies (Nielsen et al., 1989; Casey et al., 1986). The rice glutelin precursor sequence also shows homologies with leguminous 11S proteins (Takaiwa et al., 1986). The 5′-flanking region of rice glutelin gene contains a sequence similar to that of the legumin box (Gatehouse et al., 1986; Takaiwa et al., 1987b). Comparison of gene sequences not only suggests common evolutionary origins but reveals sequences that may be useful in genetic engineering such as putative seed-specific promoters, quantitative-control sequences, and variable protein sequences where gene modifications can conceivably be made without adverse effects on seed biology.

The number and complexity of seed storage protein genes present a problem if mutations are to be made to modify the amino acid composition of the proteins they code for. Ideally, the most strongly expressed gene would be modified to bring about a significant change in the overall amino acid composition of the seed. However, it is possible to enhance seed-specific expression significantly by inserting a seed-specific DNA sequence into a strong constitutive promoter such as cauliflower mosaic virus (CaMV) 35S promoter (Chen et al., 1988).

5. The Differential Expression of Seed Protein Genes Is Controlled Temporally during Seed Development and Spatially within the Plant and within the Seed. Using RNA blotting and hybridization with labeled cDNA, detailed analyses of seed mRNAs have been carried out during legume seed development. In general, each mRNA shows a characteristic period during which it increases and then declines, although all the seed mRNA overlap and are maximal at some stage within the

period of rapid protein accumulation (Gatehouse et al., 1986; Higgins, 1984; Goldberg et al., 1981). In addition, the seed mRNAs are either nondetectable or present at levels several orders of magnitude lower in mature plant organs (Goldberg et al., 1981; Walling et al., 1986). Within the organ system of the seed, differential expression of seed protein genes is further observed. For instance, lectin mRNA is present at lower concentration in the embryonic axis than in the cotyledon (Walling et al., 1986). In situ hybridization with labeled antisense mRNA reveals a finer degree of differential expression in that seed mRNAs are localized within specific cells in the cotyledon and axis but not in the seed coat (Barker et al., 1988).

That seed protein gene expression is controlled by transcriptional and posttranscriptional processes is concluded from comparing accumulation of mRNA and rates of transcription. Transcription rates are obtained by in vitro synthesis of RNA from isolated nuclei, and mRNA accumulation is obtained by RNA dot blots (Walling et al., 1986; Gatehouse et al., 1986; Chappell and Chrispeels, 1986; Evans et al., 1984; Beach et al., 1985). Features of transcription rates that support transcriptional control include observations that fluctuations in transcription rates parallel the accumulation and decay of mRNA and that transcription rates are much lower or absent in plant where mRNA levels are very low or absent.

The evidence that posttranscriptional processes play a role in controlling gene expression is based on the general observation that relative rates of transcription do not reflect the relative accumulation of mRNA (Beach et al., 1985; Boston et al., 1986; Chappell et al., 1986; Walling et al., 1986). For example, 22-kDa zein genes are transcribed at about twice the rate of 19-kDa zein genes, yet the relative amount of 19-kDa mRNA is almost 3 times that of the 22-kDa transcript (Boston et al., 1986). Mechanisms of posttranscriptional control are not known, but efficiency of transport of mRNA through the nuclear membrane and stability of mRNA in the cytoplasm have been suggested (Goldberg et al., 1989).

6. DNA Sequences That Control Seed-Specific and Quantitative Expression Have Been Identified in Seed Protein Genes. These DNA regions are being determined experimentally by analyzing the expression of seed protein genes in transgenic plants and conceptually by comparing seed protein gene sequences and identifying common, conserved sequences. A conserved region of 28 bases (approximately 50% G+C) in legumin genes is present in pea, soybean, and *V. faba* and is now called the "legumin box" (Gatehouse et al., 1986). A similar "vicilin box" of conserved sequences in vicilin gene family is found in pea, *V. faba*, and *P. vulgaris* (Gatehouse et al., 1986). The vicilin box is not homologous to the legumin box and is probably an example of a regulatory sequence specific to a gene family. A consensus sequence of 5′ T(A,C)AACACA(A,C)T(A,C) 3′ is proposed as a candidate seed-specific cis-control element because it is found in the 5′ regions of five soybean seed proteins (Goldberg, 1986). A "universal seed-specific DNA sequence" remains to be identified and tested by mutations and functional analysis in transgenic plants.

Expression of seed protein genes in transgenic tobacco and petunia plants has revealed DNA sequences that control seed-specific expression (Chen et al., 1986, 1988; Barker et al., 1988; Bray et al., 1987; Okamuro et al., 1986; Goldberg et al., 1989). Only 159 nucleotides contiguous to the transcription initiation site are needed for a low-level (5% of maximum) expression of the α-subunit of β-conglycinin gene in transgenic petunia plants (Chen et al., 1986). An additional 98 nucleotides increased the expression to maximum, equal to that of the complete gene. Analysis of the enhancing sequence of 127 nucleotides (−131 to −257) 5′ to the gene revealed 4 repeats of a 6-base-pair (G+C)-rich sequence (AG[A,C]CCCA), suggesting the critical role of the (G+C)-rich repeats in determining the level of expression of the β-conglycinin gene (Chen et al., 1986). Interestingly, insertion of a longer version of this sequence (−78 to −257) into the 5′ and 3′ ends of a constitutive promoter (CaMV 35S promoter) conferred seed-specific enhancement of 25–40-fold (Chen et al., 1988). This demonstrates that it is possible to enhance the expression of a weakly expressed seed protein gene significantly by inserting key DNA sequences in its 5′-flanking region into a CaMV 35S promoter. Significant increases in the expression of foreign genes in plants in a non-tissue-specific manner has also been achieved by duplicating the CaMV 35S promoter in tandem (Kay et al., 1987).

7. DNA Binding Proteins That Interact with Seed Protein Gene Regulatory Regions Have Been Identified in Plants. A model proposed to explain differential expression of seed protein genes during seed development is that each gene has cis-control elements that interact with cellular factors (trans elements) to activate transcription (Goldberg, 1986). These transacting factors may be coded by regulatory genes. In higher organisms, sequence-specific DNA binding proteins are directly involved in the regulation of mRNA transcription initiation (Dynan and Tjian, 1985; Kadonaga et al., 1986). A DNA binding protein that interacts with the 5′-flanking region of the soybean lectin gene and that of two other seed protein genes has been identified (Jofuku et al., 1987; Goldberg et al., 1989). The DNA binding protein activity increases and decreases during seed development, parallels the transcription rate, and is associated with a 60-kDa protein (Jofuku et al., 1987). A similar DNA binding protein has been identified in corn endosperm and interacts with the 15-bp DNA sequence found in the 5′ region of all zein genes sequenced so far (Maier et al., 1987). The physiological significance of these DNA binding proteins remains to be established. Regulatory regions in other plant genes have been shown to interact with nuclear proteins (Allen et al., 1989; Bustos et al., 1989; Jordano et al., 1989; Riggs et al., 1989).

GENETIC ENGINEERING STRATEGIES TO INCREASE THE METHIONINE CONTENT IN LEGUMES

Modification of seed proteins, if it is to be considered a success, must not have any adverse effect on seed development and germination. Several approaches are possible to increase the level of methionine, the limiting essential amino acid in legumes.

One strategy is to introduce methionine residues or methionine-rich peptides into nonconserved regions of storage proteins. This is made difficult by the large number of multigene families coding for seed storage proteins. A pertinent question is, how many single-base mutations are needed to increase the methionine content of soybean, for example, from an average of 1.2% to 2.2%, the nutritionally adequate level (FAO, 1957, 1973)? One approach is to make mutations in one subunit of glycinin, A2B1a, which is well characterized (Nielsen, 1984). Taking into account that soybean seed has 40% protein, glycinin makes up approximately 42% of the total protein, and A2B1a is one of the six subunits of glycinin, the methionine content of A2B1a subunit has to be increased from 1.8% to 14.3%. This would require 72 single-base mutations on non-

essential amino acids and essential amino acids that are abundant. This is likely to alter the processing and transport of this subunit. Alternatively, methionine-rich sequences could be inserted in certain restriction sites.

Another strategy is to transfer genes coding for methionine-rich proteins (MRP) from other species. Complementary DNA (cDNA) clones for MRPs from Brazil nut (Altenbach et al., 1987) and sunflower seed (Lilley et al., 1989) had been isolated and sequenced. The Brazil nut MRP gene, when transformed into tobacco, resulted in seed-specific increases in methionine of up to 30% in transgenic plants (Altenbach et al., 1989). A variation on this theme is to introduce synthetic genes coding for a protein with a high level of essential amino acids (Jaynes et al., 1986; Yang et al., 1989).

For the reason that increasing the level of endogenous proteins would be least likely to be deleterious to the seed, our approach is to increase the level of endogenous MRP. We identified a MRP in soybean (Kho and de Lumen, 1988) with a method that we developed for identifying methionine-containing proteins and quantitating their methionine contents (de Lumen and Kho, 1987). Further resolution using two-dimensional electrophoresis identified a 10.8-kDa protein that has 12.10% methionine for which we are now cloning cDNA (George and de Lumen, 1990). The level of this MRP has to be increased about 17-fold from its level of 0.6% of the total protein to bring the overall methionine content of soybean seed to that of the FAO reference protein.

MOLECULAR BIOLOGY OF PROTEINASE INHIBITORS

The structure, chemistry, and mechanism of action of serine protease inhibitors are well understood (Ryan, 1981). The inability of raw soybean to support the growth of rats (Liener and Kakade, 1969) and the lethal effect of raw winged bean (de Lumen and Chan, 1986) are attributed to antinutritional factors, primarily proteinase inhibitors and hemagglutinins (Higuchi et al., 1983). Since their adverse effects can be removed by proper processing, the practical significance of proteinase inhibitors in human diets in terms of their native function is questionable. In fact, the presence of denatured proteinase inhibitors in human diet may be beneficial because they are sources of sulfur amino acids and may have anticarcinogenic effects in the lower intestines (Ryan, 1989).

Thus, the molecular biology of proteinase inhibitors has focused on their function as defense compounds against predators and as models of seed proteins that are differentially expressed during development. The genes and/or cDNAs for several proteinase inhibitors have been isolated (Jofuku and Goldberg, 1989; Perez-Grau and Goldberg, 1989; Ryan, 1989; Hilder et al., 1987; Lee et al., 1986; Cleveland et al., 1987; Williamson et al., 1987; Thornburg et al., 1987; Hammond et al., 1984; Walling et al., 1986; Joudrier et al., 1987; Sanchez-Serrano, 1987; Graham et al., 1986). The defense capability of foreign proteinase inhibitors against predators in transgenic plants was demonstrated by Hilder et al. (1987).

Transformation of a cowpea trypsin inhibitor DNA construct into tobacco resulted in a high constitutive expression in the leaves and a significant resistance to tobacco budworm. Other proteinase inhibitors isolated from potato and tomato are similarly being evaluated (Ryan, 1989). Investigations on the Kunitz trypsin inhibitor gene family demonstrate their usefulness as models of seed proteins genes that are differentially expressed during embryogenesis and are also represented in leaf, stem, and root mRNA populations. Studies on Kunitz trypsin inhibitor genes suggest that each gene contains cis elements that control its expression in the developing seed and in the mature plant (Jofuku and Goldberg, 1989) and that cell-specific expression program is determined very early during embryogenesis (Perez-Grau and Goldberg, 1989).

MOLECULAR BIOLOGY OF CARBOHYDRATE SYNTHESIS

1. Starch Biosynthesis. Starch is the major component in cereals and legumes except in certain oilseeds. Cloning the genes for key enzymes in the biosynthetic pathways and studying the regulation of their expression during seed development may provide information useful for molecular genetic strategies to improve functional properties of carbohydrates.

A comprehensive discussion of starch biosynthesis in plants is covered in reviews (Preiss, 1982; Juliano, 1985; Manners, 1985). The formation of new glucosidic bonds in a growing amylose or amylopectin molecule is catalyzed by starch synthase via the transfer of a glucosyl group from ADP- or UDPglucose. The nucleotide sugars are synthesized either through the pyrophosphorylase reaction or the reversal of the sucrose synthase reaction. The current thinking is that ADPglucose pyrophosphorylase is a key regulatory enzyme in the starch biosynthetic pathway in the leaf (Preiss, 1982). However, the sucrose synthase reaction may play a very important role in sucrose–starch conversion in the developing maize and rice seed (Lee and Su, 1982; Chourey and Nelson, 1976).

The availability of mutants that are defective in starch biosynthesis has allowed extensive genetic and biochemical studies of starch biosynthesis in maize seed. These mutations lead to an altered level of starch or its quality (Creech, 1968). Four mutants in maize have been identified with their enzyme products: (a) shrunken (sh), associated with decreased sucrose synthase and a collapsed phenotype due to a reduced level of starch in the endosperm (Chourey and Nelson, 1976); (b) shrunken 2 (sh-2) and brittle 2 (bt-2), associated with reduction in endosperm starch content and a complete loss of ADPglucose pyrophosphorylase (Tsai and Nelson, 1966); and (c) waxy (wx), associated with an altered starch. The wx mutation produces starch with 100% amylopectin compared with 25% amylose and 75% amylopectin in wild type (Chourey, 1982) and starch granules that have much reduced UDPglucose starch synthase compared with wild type (Nelson and Rines, 1962).

Similar endosperm mutations like amylose extender (ae), shrunken (sh), sugary (su), dull (du), waxy (wx), white-core, and floury have been induced in rices by using chemical mutagens and irradiations (Okuno, 1976; Satoh and Omura, 1981; Yano et al., 1985). Since many of the cooking and eating qualities of rice are determined by the ratio of amylose to amylopectin, molecular analysis of these mutations might provide useful information for manipulation of amylose/amylopectin ratios through genetic engineering.

The cDNA for rice ADPglucose pyrophosphorylase has been cloned (Krishnan et al., 1986). Studies by this group showed that there are at least two tissue-specific forms of ADPglucose pyrophosphorylase encoded by distinct mRNA in the leaf and seed of wheat and rice.

The genetic expression of sucrose synthase in maize has been investigated intensively after it was shown that mutants deficient in sucrose synthase had significantly reduced levels of starch compared with wild type (Chourey and Nelson, 1976; Chourey et al., 1986; Taliercio and Chourey, 1989; Gupta et al., 1988; McCarty et al., 1986).

In maize, there are two tissue-specific isozymes of sucrose synthase encoded by two sets of genes which are differentially expressed in endosperm and embryo tissues (McCarty et al., 1986; Chourey, 1981). Both sets of enzymes have been cloned and characterized (Sheldon et al., 1983; McCarty et al., 1986; Gupta et al., 1988).

Strategies to increase the protein content of cereals, especially rice, will have to take into consideration the inverse relationship between protein content and grain yield (Doll, 1984; Juliano, 1985). In a given rice variety, grain yield and protein content may be increased up to a certain threshold, beyond which further increases in protein result in a sacrifice in grain yield. Since the decrease in grain yield is due completely to reduction in starch (Doll, 1984), understanding the genetic relationships between starch biosynthesis and storage protein biosynthesis in the developing cereal grain may provide useful information to increase protein content without adverse effect on yield. Competition for photoassimilates between the carbohydrate and protein biosynthetic pathways may be a simple explanation. Nitrogen translocation may also be a factor, since high-protein rice varieties seem to have a more efficient nitrogen transport from the leaves to the developing grain, as reflected in increased free amino acids in the culm sap and developing grain, rather than increased absorption of nitrogen from the soil (Cagampang et al., 1971; Perez et al., 1973). Rice mutants defective in starch biosynthesis that have reduced starch and increased protein and lipid contents might be useful models to study the starch–protein relationships (Juliano, 1989).

2. Biosynthesis of Raffinose Oligosaccharides. The raffinose family oligosaccharides, which are α-D-galactosides of sucrose and include raffinose, stachyose, and verbascose, are second only to sucrose in abundance in legume seeds (Dey, 1985). They accumulate at the late stage of seed maturation as the seed desiccates toward dormancy and have been proposed to play roles in response of plants to cold, seed viability, and flatulence in humans upon consumption of beans (Saravitz et al., 1987; Kandler and Hopf, 1980; Calloway and Murphy, 1968; Rackis, 1981; Ovcharov and Koshelev, 1974)). The raffinose oligosaccharides are synthesized by the sequential transfer of galactosyl units to sucrose via galactinol catalyzed by specific transferases (Dey, 1985). Since the biosynthetic pathway of these oligosaccharides is known, understanding the regulation of genes for key enzymes in the pathway would be useful in manipulating the levels of raffinose oligosaccharides in legume seeds to clarify their roles in plants and in flatulence formation (Castillo et al., 1990).

MOLECULAR BIOLOGY OF LIPID METABOLISM

Legumes are the world's major source of vegetable oil, soybean being the single largest source (Hymowitz, 1987). The nutritional properties, food applications, and industrial uses of vegetable oils are largely dependent on their fatty acid composition. Thus, a major objective of oilseed plant breeding programs is to obtain cultivars with fatty acid compositions that are suited to certain market needs.

The application of techniques of molecular biology to plant lipid metabolism is very much in its infancy. Mutants of *Arabidopsis thaliana* that are deficient in fatty acid desaturation are being used to study the regulation and functional significance of lipid unsaturation (Somerville et al., 1987; Hugly et al., 1989; Kunst et al., 1989). Since acyl carrier protein (ACP) plays a central role in plant lipid metabolism and is well characterized, one approach is to study its protein and gene structure, which may reveal control mechanisms that regulate fatty acid biosynthesis in plants (Ohlrogge et al., 1987; Slabas et al., 1987; Guerra et al., 1987; Elhussein et al., 1987). Another approach is suggested by the differences in fatty acid composition in seed storage lipids and those found in membrane phospholipids in other organs of plants. Studies are being undertaken to determine if oilseeds have mechanisms to target unusual fatty acids to storage triglycerides, where they will have less deleterious effects, and to exclude them from phospholipids in the membranes, where their presence can disrupt membrane structure (Battey and Ohlrogge, 1989). Such information will give insights into the design of genetic engineering techniques to modify fatty acid composition without disrupting normal plant metabolism.

Lipoxygenases are a group of polyunsaturated fatty acid oxidases that are thought to play important roles in the generation of undesirable flavors and aroma in legume seed protein and oil products (Kinsella, 1979). Mutants lacking various isoenzymes of lipoxygenases have been identified in soybean (Hildebrand and Hymowitz, 1981; Kitamura et al., 1983; Kitamura, 1984). cDNA clones for soybean lipoxygenases have been isolated and used as probes to demonstrate that lipoxygenase null mutants accumulate reduced levels of lipoxygenases transcripts (Start et al., 1986). Soybean genomic clones have been isolated and characterized and are being transformed into soybean and tobacco to obtain transformants with up and down regulated lipoxygenase activity. These are being correlated with generation of volatile compounds and with pest resistance to better understand the role of lipoxygenases in plant metabolism (Hildebrand et al., 1987).

TRANSFORMATION TECHNIQUES

A key to any genetic engineering strategy requires the introduction of the desirable gene or set of genes into the genome of the target plant. Very rapid progress has been achieved in this area in the past 4 years, and the reader is referred to several reviews on plant transformation systems for detailed discussion (Uchimiya et al., 1989; Gobel and Lorz, 1988; Weising et al., 1988; Gasser and Fraley, 1989; Schell, 1987; Klee et al., 1987; Cocking and Davey, 1987; Fraley et al., 1986).

Currently used gene-transfer techniques in plants can be broadly classified into one group that uses a delivery vehicle for DNA transfer and another group that involves the direct transfer of DNA. The most commonly used vector is *Agrobacterium tumefaciens*, which causes tumorous crown galls on infected species. The virulence of *Agrobacterium* is due to tumor-inducing (Ti) plasmids which contain two regions necessary for the transfer and integration of the bacterial genes: a transferred DNA (T-DNA) domain that is transferred to the plant and a virulence (Vir) domain that catalyzes the transfer but which is not itself transferred. Molecular analysis of the Ti plasmid revealed that the T-DNA and the Vir region do not have to be in the same piece of DNA (Fraley et al., 1986), leading to the design of intermediate vectors that can be multiplied in *Escherichia coli* and manipulated before introduction into *Agrobacterium* cells and infection of plants. A typical vector contains the border sequences that define the limits of the DNA transferred to the plant, a plant selectable marker such as kanamycin resistance, an origin of replication that allows multiplication in *Agrobacterium*, another origin of replication for *E. coli*, and an *E. coli* selectable marker such as spectinomycin resistance (Gasser and Fraley, 1989). The piece of DNA transferred to the plants bounded by the border sequences contains the foreign gene and the selectable marker. The

intermediate vector is introduced into *Agrobacterium* cells containing the plasmid with the Vir region and inactivated pathogenesis genes (disarmed Ti). Upon infection of the host, the Vir DNA acts in trans to mobilize the transfer of the DNA within the border sequences and the phenotype conferred by the kanamycin resistance gene is used to screen transformed cells during plant regeneration. The most commonly used infection procedure is exposure of leaf disks to *Agrobacterium* cells (Horsch et al., 1985).

An important limitation of *Agrobacterium*-mediated gene transfer is that the major cereal and legume crops have not been successfully transformed because they are poor hosts and because of the inability to regenerate whole plants from leaf disks. Recent efforts to deliver DNA directly into the genome of these plants have been reported. Such techniques include facilitated uptake of DNA by protoplasts (Fromm et al., 1986; Uchimiya et al., 1986), DNA application into reproductive organs (Luo and Wu, 1988; de la Pena et al., 1987), microinjection into cells of immature embryos (Neuhaus et al., 1987), and rehydration of desiccated embryos (Topfer et al., 1989). Of practical value is the stable transformation of soybean by DNA-coated gold particles introduced into meristems by electric discharge acceleration (McCabe et al., 1988; Christou et al., 1988). The introduced genes are inherited in a Mendelian manner (Christou et al., 1989). Of interest also is the regeneration of fertile rice plants from transformed protoplasts (Shimamoto et al., 1989), since easy and reproducible production of transgenic cereals has not been reported previously.

SUMMARY

In summary, rapid developments in our knowledge of plant gene structure and organization and their regulation bring us closer to actual applications of genetic engineering to improve the nutritional properties of legumes and cereals. Improving the digestibility of proteins and starch, reducing flatulence, altering fatty acid composition, manipulating the amylose/amylopectin ratios in rice starch, and minimizing formation of undesirable flavors and aroma will have to be long-term objectives because of the very limited knowledge we have of the molecular biology of carbohydrate and lipid metabolism in the developing seed. To design seed proteins with specific functional attributes without adverse effects on seed biology requires knowledge of the molecular basis of functional properties and of critical features of seed protein structures that are essential in posttranslational processing and packaging.

With the rapid progress being made in gene cloning and transformation methodologies in plants, our knowledge of basic plant metabolic processes can be the limiting factor in the application of genetic engineering to improve the usefulness of plant seeds as food sources. Any genetic changes must not have any adverse effects on normal seed biology.

ACKNOWLEDGMENT

I thank the following for reading the manuscript and for their comments and criticisms: Kenneth Carpenter, Pilar Reyes, Ann George, Bill Odegard, and Alice Saliba.

LITERATURE CITED

Abbott, J. C. Protein Supplies and Prospects: The Problem. In *World Protein Resources*; Gould, R. F., Ed.; American Chemical Society: Washington, DC, 1966; pp 1–15.

Allen, R. D.; Bernier, F.; Lessard, P. A.; Beachy, R. N. Nuclear Factors that Interact with a Soybean β-conglycinin Enhancer. *Plant Cell* **1989**, *1*, 623–631.

Altenbach, S. B.; Pearson, K. W.; Leung, F. W.; Sun, S. S. M. Cloning and Sequence Analysis of a cDNA Encoding a Brazil Nut Protein Exceptionally Rich in Methionine. *Plant Mol. Biol.* **1987**, *8*, 239–250.

Altenbach, S. B.; Pearson, K. W.; Meeker, G.; Staracci, L. C.; Sun, S. M. Enhancement of the Methionine Content of Seed Proteins by the Expression of a Chimeric Gene Encoding a Methionine-rich Protein in Transgenic Plants. *Plant Mol. Biol.* **1989**, *13*, 513–522.

American Soybean Association. *Soya Bluebook*; St. Louis, MO, 1982.

Argos, P.; Pedersen, K.; Marks, M. D.; Larkins, B. A. A Structural Model for Maize Zein Proteins. *J. Biol. Chem.* **1982**, *257*, 9984–9990.

Barker, S. J.; Harada, J. J.; Goldberg, R. B. Cellular Localization of Soybean Storage Protein mRNA in Transformed Tobacco Seeds. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 458–462.

Battey, J. F.; Ohlrogge, J. B. A Comparison of the Metabolic Fate of Fatty Acids of Different Chain Lengths in Developing Oilseeds. *Plant Physiol.* **1989**, *90*, 835–840.

Baumlein, H.; Wobus, U.; Pustell, J.; Kafatos, F. C. The Legumin Gene Family: Structure of a B Type Gene of Vicia faba and a Possible Legumin Gene Specific Regulatory Element. *Nucleic Acids Res.* **1986**, *14*, 2707–2720.

Beach, L. R.; Spencer, D.; Randall, P. J.; Higgins, T. J. V. Transcriptional and Posttranscriptional Regulation of Storage Protein Gene Expression in Sulfur Deficient Pea Seeds. *Nucleic Acids Res.* **1985**, *13*, 999–1013.

Blundell, T.; Sternberg, M. J. Computer Aided Design in Protein Engineering. *Trends Biotechnol.* **1985**, *3*, 229–231.

Bodwell, C. E. Nutritional Implications of Cereals, Legumes and Their Products. In *Cereals and Legumes in the Food Supply*; Dupont, J., Osman, E. M., Eds.; Iowa State University Press: Ames, IA, 1987; pp 259–275.

Boston, R. S.; Kodrycki, R.; Larkins, B. A. Transcriptional and Posttranscriptional Regulation of Maize Zein Genes. In *Molecular Biology of Seed Storage Proteins and Lectins*; Shannon, L. M., Chrispeels, M. J., Eds.; American Society of Plant Physiologists: Rockville, MD, 1986.

Bray, E. A.; Naito, S.; Pan, N. S.; Anderson, E.; Dube, P.; Beachy, R. N. Expression of the β-subunit of β-conglycinin in Seeds of Transgenic Plants. *Planta* **1987**, *172*, 364–370.

Bressani, R.; Elias, L. G. Processed Vegetable Protein Mixtures for Human Consumption in Developing Countries. *Adv. Food Res.* **1968**, *16*, 1–103.

Burr, B.; Burr, F. A.; St. John, T. P.; Thomas, M.; Davis, R. W. Zein Storage Gene Family of Maize: An Assessment of Heterogeneity With Cloned mRNA Sequences. *J. Mol. Biol.* **1982**, *154*, 33–49.

Bustos, M. M.; Guiltinan, M. J.; Jordano, J.; Begum, D.; Kalkan, F. A.; Hall, T. C. Regulation of B-glucuronidase Expression in Transgenic Tobacco Plants by an A/T-rich,cis-acting Sequence Found Upstream of a French Bean B-phaseolin Gene. *Plant Cell* **1989**, *1*, 839–853.

Butz, E. L. World Protein Markets—A Supplier's View. Plenary Session of World Protein Markets, World Soybean Conference Nov 11–14, 1973, Munich, Germany.

Cagampang, G. B.; Cruz, L. J.; Juliano, B. O. Free Amino acids in the Bleeding Sap and Developing Grain of the Rice Plant. *Cereal Chem.* **1971**, *48*, 533–539.

Calloway, D. H.; Murphy, E. L. The Use of Expired Air to Measure Intestinal Gas Formation. *Ann. N.Y. Acad. Sci.* **1968**, *150*, 82–85.

Casey, R.; Downey, C.; Ellis, N. Legume Storage Proteins and Their Genes. *Oxford Surv. Plant Mol. Cell Biol.* **1986**, *3*, 1–95.

Castillo, E. M.; de Lumen, B. O.; Reyes, P. S.; de Lumen, H. Z. Raffinose Synthase and Galactinol Synthase in Developing Seeds and Leaves of Legumes. *J. Agric. Food Chem.* **1990**, *38*, 351–355.

Chappell, J.; Chrispeels, M. J. Transcriptional and Posttranscriptional Control of Phaseolin and Phytohemagglutinin Gene Expression in Developing Cotyledons of Phaseolus vulgaris. *Plant Physiol.* **1986**, *81*, 50–54.

Chen, Z. L.; Schuller, M. A.; Beachy, R. N. Functional Analysis of Regulatory Elements in a Plant Embryo-Specific Gene. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 8560–8564.

Chen, Z. L.; Pan, N. S.; Beachy, R. N. A DNA Sequence Element That Confers Seed-Specific Enhancement to a Constitutive Promoter. *EMBO J.* **1988**, *7*, 297–302.

Chourey, P. S. Genetic Control of Sucrose Synthetase in Maize Endosperm. *MGG, Mol. Gen. Genet.* **1981**, *184*, 372–376.

Chourey, P. S. Starch Mutants and Their Protein Products. In *Maize for Biological Research*; Sheridan, W. F., Ed.; Plant Molecular Biology Association: Charlottesville, VA, 1982; pp 129–133.

Chourey, P. S.; Nelson, O. E. The enzymatic Deficiency Conditioned by the Shrunken-1 Mutations in Maize. *Biochem. Genet.* **1976**, *14*, 1041–1055.

Chourey, P. S.; Lantham, M. D.; Still, P. E. Expression of Two Sucrose Synthetase Genes in Endosperm and Seedlings of Maize: Evidence of Tissue Specific Polymerization of Protomers. *MGG, Mol. Gen. Genet.* **1986**, *203*, 251–255.

Christou, P.; McCabe, D. E.; Swain, W. F. Stable Transformation of Soybean Callus by DNA-coated Gold Particles. *Plant Physiol.* **1988**, *87*, 671–674.

Christou, P.; Swain, W. F.; Yang, N. S.; McCabe, D. E. Inheritance and Expression of Foreign Genes in Transgenic Soybean Plants. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 7500–7504.

Cleveland, T. E.; Thornburg, R. W.; Ryan, C. A. Molecular Characterization of a Wound-inducible Inhibitor I Gene from Potato and Processing of its mRNA and Protein. *Plant Mol. Biol.* **1987**, *8*, 119–207.

Cocking, E. C.; Davey, M. R. Gene Transfer in Cereals. *Science* **1987**, *236*, 1259–1262.

Creech, R. G. Carbohydrate Synthesis in Maize. *Adv. Agron.* **1968**, *20*, 275–322.

Croy, R. R. D.; Gatehouse, J. A. Genetic Engineering of Seed Proteins: Current and Potential Applications. In *Plant Genetic Engineering*; Dodds, J. H., Ed.; Cambridge University Press: Cambridge, U.K., 1985; pp 143–268.

Croy, R. R. D.; Lycett, G. W.; Gatehouse, J. A.; Yarwood, J. N.; Boulter, D. Cloning, Characterization and Sequence Analysis of cDNAs Encoding Pea Storage Protein Precursors-Legumin. *Nature* **1982**, *295*, 76–79.

De la Pena, A.; Lorz, H.; Schell, J. Transgenic Rye Plants Obtained by Injecting DNA Into Young Floral Tillers. *Nature* **1987**, *325*, 274–275.

de Lumen, B. O.; Chan, J. Antinutritional and Biochemical Properties of Winged Bean Trypsin Inhibitors. In *Nutritional and Toxicological Significance of Enzyme Inhibitors in Foods*; Friedman, M., Ed.; Plenum Press: New York, 1986; pp 413–427.

de Lumen, B. O.; Kho, C. J. Identification of Methionine Containing Proteins and Quantitation of their Methionine Contents. *J. Agric. Food Chem.* **1987**, *35*, 688–691.

Dey, P. M. D-Galactoside Containing Oligosaccharides. In *Biochemistry of storage Carbohydrates in Green Plants*; Dey, P. M., Dixon, R. A., Eds.; Academic Press: New York, 1985; pp 53–129.

Doll, H. Nutritional Aspects of Cereal Proteins and Approaches to Overcome Their Deficiencies. *Philos. Trans. R. Soc. London* **1984**, *B304*, 373–380.

Domoney, C.; Casey, R. Measurement of Gene Number for Seed Storage Proteins in Pisum. *Nucleic Acids Res.* **1985**, *13*, 687–699.

Duffus, C. M.; Slaughter, J. C. *Seeds and Their Uses*; Wiley: New York, 1980.

Dynan, W. S.; Tjian, R. Control of Eukaryotic mRNA Synthesis by Sequence Specific DNA Binding Proteins. *Nature* **1985**, *316*, 774–778.

Eggum, B. O.; Beames, R. M. The Nutritive Value of Seed Proteins. In *Seed Proteins: Biochemistry, Genetics, Nutritive Value*; Gottschalk, W., Muller, H. P., Eds.; Nijhoff: The Hague, 1983; pp 499–531.

Elhussein, S. A.; Miernyk, J. A.; Ohlrogge, J. B. A Preliminary Characterization of Plant Holo-Acyl Carrier Protein Synthase. In *The Metabolism, Structure, and Function of Plant Lipids*; Stumpf, P. K., Mudd, J. B., Ness, W. D., Eds.; Plenum Press: New York, 1987; pp 709–713.

Elliston, K. O.; Imran, S.; Messing, J. Cloned Genes of Storage Proteins. *Plant Mol. Biol. Rep.* **1988**, *6*, 22–26.

Evans, I. M.; Gatehouse, J. A.; Croy, R. R. D.; Boulter, D. Regulation of Transcription of Storage Protein mRNA in Nuclei Isolated from Developing Pea Cotyledons. *Planta* **1984**, *160*, 559–568.

FAO. *Protein Requirements*; FAO Nutrition Studies 16; FAO: Rome, 1957.

FAO. *Food Balance Sheets 1975–1977*; Food and Agriculture Organization of the UN, 1980.

FAO/WHO. Energy Requirements and Protein Requirements. *WHO Tech. Rep. Ser.* **1973**, *No. 522*.

FAO/WHO/UNU. Energy and Protein Requirements. WHO *Tech. Rep. Ser.* **1985**, *No. 724*.

Fraley, R. T.; Rogers, S. G.; Horsch, R. B. Genetic Transformation in Higher Plants. *CRC Crit. Rev. Plant Sci.* **1986**, *4*, 1–46.

Fromm, M.; Taylor, L. P.; Talbot, V. Stable Transformation of Maize After Gene Transfer by Electroporation. *Nature* **1986**, *319*, 791–793.

Gasser, C. S.; Fraley, R. T. Genetically Engineering Plants for Crop Improvement. *Science* **1989**, *236*, 1293–1299.

Gatehouse, J. A.; Evans, I. M.; Croy, R. R. D.; Boulter, D. Differential Expression of Genes During Seed Development. *Philos. Trans. R. Soc. London* **1986**, *B314*, 367–384.

George, A.; de Lumen, B. O. A Novel Methionine Rich Protein in Soybean Seed: Identification, Amino Acid Composition and N-Terminus Sequence. *J. Agric. Food Chem.* **1990**, submitted for publication.

Gobel, E.; Lorz, H. Genetic Manipulation of Cereals. *Oxford Surv. Plant Mol. Cell Biol.* **1988**, *5*, 1–22.

Goldberg, R. B. Regulation of Plant Gene Expression. *Philos. Trans. R. Soc. London* **1986**, *B314*, 343–353.

Goldberg, R. B. Plants: Novel Developmental Processes. *Science* **1988**, *240*, 1460–1466.

Goldberg, R. B.; Hoschek, G.; Tam, S. H.; Ditta, G. S.; Breidenbach, R. W. Abundance, Diversity and Regulation of mRNA Sequence Sets in Soybean Embryogenesis. *Dev. Biol.* **1981**, *83*, 201–217.

Goldberg, R. B.; Barker, S.; Perez-Grau, L. Regulation of Gene Expression During Plant Embryogenesis. *Cell* **1989**, *56*, 149–160.

Graham, J.; Hall, G.; Pearce, G.; Ryan, C. A. Regulation of Synthesis of Proteinase Inhibitors I and II mRNAs in Leaves of Wounded Tomato Plants. *Planta* **1986**, *169*, 399–405.

Guerra, D. J.; Ohlrogge, J. B.; Frentzen, M. A Possible Differential Role for Plant Acyl Carrier Protein Isoforms in Higher Plants. In *The Metabolism, Structure and Function of Plant Lipids*; Stumpf, P. K., Mudd, J. B., Ness, W. D., Eds.; Plenum Press: New York, 1987; pp 705–708.

Gupta, M.; Chourey, P. S.; Burr, P.; Still, P. cDNAs of Two Non-allelic Sucrose Synthase Genes in Maize: Cloning Expression, Characterization and Molecular Mapping of the Sucrose Synthase-2 Gene. *Plant Mol. Biol.* **1988**, *10*, 215–224.

Gupta, Y. P. Nutritive Value of Food Legumes. In *Chemistry and Biochemistry of Legumes*; Arora, S. K., Ed.; Oxford and IBH Publishing: New Delhi, 1983.

Hagen, G.; Rubinstein, I. Complex Organization of Zein Genes in Maize. *Gene* **1981**, *13*, 239–249.

Hammond, R. W.; Foard, D. E.; Larkins, B. A. Molecular Cloning and Analysis of a Gene Coding for the Bowman-Birk Protease Inhibitor in Soybean. *J. Biol. Chem.* **1984**, *259*, 9883–9890.

Harada, J. J.; Barker, S. J.; Goldberg, R. B. Soybean β-conglycinin Genes are Clustered in Several DNA Regions and are Regulated by Transcriptional and Posttranscriptional Processes. *Plant Cell* **1989**, *1*, 415–425.

Higgins, T. J. V. Synthesis and Regulation of Major Seed Proteins in Seeds. *Annu. Rev. Plant Physiol.* **1984**, *35*, 191–221.

Higuchi, M.; Suga, M.; Iwai, K. Participation of Lectin in Biological Effects of Raw Winged Bean Seeds on Rats. *Agric. Biol. Chem.* **1983**, *47*, 1879–1886.

Hildebrand, D. F.; Hymowitz, T. Two Soybean Genotypes Lacking Lipoxygenase. *J. Am. Oil Chem. Soc.* **1981**, *58*, 583–586.

Hildebrand, D. F.; Altschuler, M., et al. Physiological and Transformational Analyses of Lipoxygenase. In *The Metabolism, Structure and Function of Plant Lipids*; Stumpf, P. K., Mudd, J. B., Ness, W. D., Eds.; Plenum Press: New York, 1987; pp 715–717.

Hilder, V. A.; Gatehouse, A. M. R.; Sheerman, S. E.; Barker, R. F.; Boulter, D. A Novel Mechanism of Insect Resistance Engineered into Tobacco. *Nature* **1987**, *330*, 160–163.

Hol, W. G. J. Applying Knowledge of Protein Structure and Function. *Trends Biotechnol.* **1987**, *5*, 137–143.

Horsch, R. B.; Fry, J. E., et al. A Simple and General Method for Transferring Genes Into Plants. *Science* **1985**, *227*, 1229–1231.

Hugly, S.; Kunst, L.; Browse, J.; Somerville, C. R. Enhanced Thermal Tolerance of Photosynthesis and Altered Chloroplast Ultra Structure in a Mutant of Arabidopsis Deficient in Lipid Desaturation. *Plant Physiol.* **1989**, *90*, 1134–1142.

Hymowitz, T. Increasing Yields, Food and Nutritional Quality Through Breeding: Grain Legumes. In *Cereals and Legumes in the Food Supply*; Dupont, J., Osman, E. M., Eds.; Iowa State University Press: Ames, IA, 1987; pp 15–24.

Jaynes, J.; Yang, M.; Espinoza, M.; Dodds, J. Plant Protein Improvement by Genetic Engineering: Use Of Synthetic Genes. *Trends Biotechnol.* **1986**, *4*, 314–320.

Jofuku, K. D.; Goldberg, R. B. Kunitz Trypsin Inhibitor Genes are Differentially Expressed During the Soybean Life Cycle and in Transformed Tobacco Plants. *Plant Cell* **1989**, *1*, 1079–1093.

Jofuku, K. D.; Okamuro, J. K.; Goldberg, R. B. Interaction of an Embryo DNA Binding Protein With a Soybean Lectin Gene Upstream Region. *Nature* **1987**, *328*, 734–737.

Jordano, J.; Almoguera, C.; Thomas, T. L. A Sunflower Helianthin Gene Upstream Sequence Ensemble Contains an Enhancer and Sites of Nuclear Protein Interaction. *Plant Cell* **1989**, *1*, 855–866.

Joudrier, P. E.; Foard, D. E.; Floener, L. A.; Larkins, B. A. Isolation and Sequence of cDNA Encoding the Soybean Protease Inhibitors PI IV and C-II. *Plant Mol. Biol.* **1987**, *10*, 35–42.

Juliano, B. O., Ed. *Rice: Chemistry and Technology*; American Association of Cereal Chemists: St. Paul, MN, 1985; pp 1–16.

Juliano, B. O. 1989, personal communication.

Kadonaga, J. T.; Jones, K. A.; Tjian, R. Promoter-specific Activation of RNA Polymerase II Transcription by Sp1. *Trends Biochem.* **1986**, *11*, 20–23.

Kandler, O.; Hopf, H. Occurrence, Metabolism and Function of Oligosaccharides. In *Biochemistry of Plants. Carbohydrates*; Preiss, J., Ed.; Academic Press: New York, 1980; Vol. 3, pp 221–270.

Kay, R.; Chan, A.; Daly, M.; McPherson, J. Duplication of CaMV 35S Promoter Sequences Creates a Strong Enhancer for Plant Genes. *Science* **1987**, *236*, 1299–1302.

Kho, C. J.; de Lumen, B. O. Identification and Isolation of Methionine-Cysteine Rich Proteins in Soybean Seed. *Plant Foods Human Nutr.* **1988**, *38*, 287–296.

Khrishnan, H. B.; Reeves, C. D.; Okita, T. W. ADP-Pyrophosphorylase is Encoded by Different mRNA Transcripts in Leaf and Endosperm of Cereals. *Plant Physiol.* **1986**, *81*, 642–645.

Kinsella, J. Functional Properties of Soy Proteins. *J. Am. Oil Chem. Soc.* **1979**, *56*, 242–258.

Kitamura, K. Biochemical Characterization of Lipoxygenase Lacking Mutants, L-1 Less, L-2 Less, L-3 Less Soybeans. *Agric. Biol. Chem.* **1984**, *48*, 2339–2346.

Kitamura, K.; Davies, C. S.; Kaizuma, K.; Nielsen, N. C. Genetic Analysis of a Nullallele for Lipoxygenase-3 in Soybean Seeds. *Crop Sci.* **1983**, *23*, 924–927.

Klee, H.; Horsch, R.; Rogers, S. Agrobacterium-mediated Plant Transformation and Its Further Appliction to Plant Biology. *Annu. Rev. Plant Physiol.* **1987**, *38*, 467–486.

Kreis, M.; Shewry, P. R.; Forde, B. G.; Forde, J.; Miflin, B. J. Structure and Evolution of Seed Storage Proteins and Their Genes With Particular Reference to Those of Wheat, Barley and Rye. *Oxford Surv. Plant Mol. Cell. Biol.* **1985**, *2*, 253–317.

Kuhlemeir, C.; Green, P. J.; Chua, N. H. Regulation of Gene Expression in Higher Plants. *Annu. Rev. Plant Physiol.* **1987**, *38*, 221–466.

Kunst, L.; Somerville, C. R. A Mutant of Arabidopsis Deficient in Desaturation of Palmitic Acid in Leaf Lipids. *Plant Physiol.* **1989**, *90*, 943–947.

Lee, J. L.; Brow, W. E.; Pearce, G.; Dreher, T. W.; Graham, J. S.; Ahern, K. B.; Pearson, G. D.; Ryan, C. A. Molecular Characterization and Phylogenic Studies of a Wound-Inducible Proteinase Inhibitor I Gene in Lycopersicum Species. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 7277–7281.

Lee, P. D.; Su, J. C. Sucrose-Starch Transforming System in Rice Grains—A Tracer Feeding Study. *Proc. Natl. Sci. Counc., Repus. China* **1982**, *B6*, 189–196.

Lilley, G. G.; Caldwell, J. B.; Kortt, A. A. Isolation and Primary Structure for a Novel, Methionine-rich Protein from Sunflowerseeds (Helianthus annus L.) *Proceedings of the World Congress on Vegetable Protein Utilization in Human Foods and Animal Feedstuffs*; Applewhite, T. H., Ed.; American Oil Chemists' Society: Champaign, IL, 1989; pp 497–502.

Lin, Y. L. Genetic Engineering and Process Development for Production of Food Processing Enzymes and Additives. *Food Technol.* **1986**, *40*, 104–106.

Luo, Z.; Wu, R. A Simple Method for the Transformation of Rice via the Pollen Tube Pathway. *Plant Mol. Biol. Rep.* **1988**, *6*, 165–174.

Maier, U. G.; Brown, J. W. S.; Feix, G. Binding of a Nuclear Factor to a Consensus Sequence in the 5'-Flanking Regions of Zein Genes from Maize. *EMBO J.* **1987**, *6*, 17–22.

Manners, D. J. Starch. In *Biochemistry of Storage Carbohydrates in Plants*; Dey, P. M., Dixon, R. A., Eds.; Academic Press: London, 1985; pp 149–203.

McCabe, D. E.; Swain, W. F., et al. Stable Transformation of Soybean (Glycine max) by Particle Acceleration. *Bio/Technol.* **1988**, *6*, 923–926.

McCarty, D.; Shaw, J.; Hannah, L. The Cloning, Genetic Mapping, and Expression of the Constitutive Sucrose Synthase Locus of Maize. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 9099–9103.

Muntz, K. Intracellular Protein Sorting and the Formation of Protein Reserves in Storage Tissue Cells of Plant Seeds *Biochem. Physiol. Pflanzen.* **1989**, *185*, 315–335.

Nakai, S.; Powrie, W. D. Modifications of Proteins for Functional and Nutritional Improvements. In *Cereals—A Renewable Resource, Theory and Practice*; Pomeranz, Y., Munck, L., Eds.; American Association of Cereal Chemists: St. Paul, MN, 1981; Chapter 11, pp 217–242.

National Research Council. *Quality Protein Maize*; National Academy Press: Washington, DC, 1988; p 6.

Nelson, O. E.; Rines, H. W. The Enzymatic Deficiency in Waxy Mutants of Maize. *Biochem. Biophys. Res. Commun.* **1962**, *9*, 297–300.

Neuhaus, G.; Spangenberg, G., et al. Transgenic Rapeseed Plants Obtained by the Microinjection of DNA into Microspore-derived Embryoids. *Theor. Appl. Genet.* **1987**, *74*, 30–36.

Nielsen, N. C. The Chemistry of Legume Storage Proteins. *Philos. Trans. R. Soc. London* **1984**, *B304*, 287–296.

Nielsen, N. C.; Dickinson, C. D.; Cho, T. J.; Thanh, V. H.; Scallon, B. T.; 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.

Odigboh, E. U. Cassava: Production, Processing and Utilization. In *Handbook of Tropical Foods*; Chan, H. T., Ed.; Dekker: New York, 1983; pp 145–202.

Ohlrogge, J. B.; Beremand, P. D.; Hannapel, D. J.; Guerra, D. E.; Elmore, D. E.; Kuhn, D. N. Acyl Carrier Protein as a Probe of the Molecular Biology of Plant Fatty Acid Synthesis. In *The Metabolism, Structure, and Function of Plant Lipids*; Stumpf, P. K., Mudd, J. B., Ness, W. D., Eds.; Plenum Press: New York, 1987; pp 689–695.

Okamuro, J. K.; Jofuko, K. D.; Goldberg, R. B. Soybean Lectin Gene and Flanking Non-seed Protein Genes are Developmentally Regulated in Transformed Tobacco Plants. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 8240–8244.

Okuno, K. A Low-amylose Mutant of Rice. *Div. Genet. Natl. Inst. Agric. Sci. Jpn. Annu. Rep.* **1976**, *1*, 28–29.

Ovcharove, K. G.; Kosheleve, Y. P. Sugar Content in Corn Seeds With Different Viability. *Fiziol Rast.* **1974**, *21*, 969–974.

Pedersen, K.; Deverux, J.; Wilson, D.; Sheldon, E.; Larkins, B. A. Cloning and Sequence Analysis Reveal Structural Variation Among Related Zein Genes in Maize. *Cell* **1982**, *29*, 1015–1026.

Perez, C. M.; Cagampang, G. B.; Esmana, B. V.; Monserrate, R. U.; Juliano, B. O. Protein Metabolism in Leaves and Developing Grains of Rices Differing in Grain Protein Content. *Plant Physiol.* **1973**, *51*, 537–542.

Perez-Grau, L.; Goldberg, R. B. Soybean Seed Protein Genes are Regulated Spatially during Embryogenesis. *Plant Cell* **1989**, *1*, 1095–1109.

Preiss, J. Regulation of the Biosynthesis and Degradation of Starch. *Annu. Rev. Plant Physiol.* **1982**, *33*, 431–454.

Rackis, J. Flatulence Caused by Soya and its Control Through Processing. *J. Am. Oil Chem. Soc.* **1981**, *58*, 503–509.

Raven, P. H.; Evert, R. F.; Eichhorn, S. E. *Biology of Plants*, 4th ed.; Worth Publishers: New York, 1987; pp 373–383.

Riggs, C. D.; Voelker, T. A.; Chrispeels, M. J. Cotyledon Nuclear Proteins Bind to DNA Fragments Harboring Regulatory Elements of Phytohemagglutinin Genes. *Plant Cell* **1989**, *1*, 609–621.

Roxas, B. V.; Intengan, C. Ll.; Juliano, B. O. Protein Quality of High Protein and Low Protein Milled Rices in Preschool Children. *J. Nutr.* **1979**, *109*, 832–839.

Sanchez-Serrano, J. J.; Keil, M.; O'Connor, A.; Schell, J.; Willmitzer, L. Wound-induced Expression of a Potato Proteinase Inhibitor II Gene in Transgenic Plants. *EMBO J.* **1987**, *6*, 303–306.

Saravitz, D. M.; Pharr, D. M.; Carter, T. E. Galactinol Synthase Activity and Soluble Sugars in Developing Seeds of Four Soybean Genotypes. *Plant Physiol.* **1987**, *83*, 185–189.

Satoh, H.; Omura, T. New Endosperm Mutations Induced by Chemical Mutagens in Rice, Oryza sativa. *Jpn. J. Breed.* **1981**, *31*, 316–326.

Schell, J. Transgenic Plants as Tools to Study the Molecular Organization of Plant Genes. *Science* **1987**, *237*, 1176–1183.

Sheldon, E.; Ferl, R.; Fedoroff, N.; Hannah, L. C. Isolation and Analysis of a Genomic Clone Encoding Sucrose Synthetase in Maize: Evidence for Two Introns in Sh. *MGG, Mol. Gen. Genet.* **1983**, *190*, 421–426.

Shewry, P. R.; Miflin, B. J. Seed Storage Proteins of Economically Important Cereals. *Adv. Cereal Sci. Technol.* **1984**, *7*, 1–83.

Shimamoto, K.; Terada, R.; Izawa, T.; Fujimoto, H. Fertile Transgenic Rice Plants Regenerated from Transformed Protoplasts. *Nature* **1989**, *337*, 274–276.

Slabas, A. R.; Harding, P., et al. Oil Seed Acyl Carrier Protein (ACP) Protein and Gene Structure. In *The Metabolism, Structure, and Function of Plant Lipids*; Stumpf, P. K., Mudd, J. B., Ness, W. D., Eds.; Plenum Press: New York, 1987; pp 697–700.

Slightom, J. L.; Drong, R. F.; Klassy, R. C.; Hoffmann, L. M. Nucleotide Sequences from Phaseolin cDNA Clones: the Major Storage Protein from Phaseolus vulgaris are Encoded by Two Unique Gene Families. *Nucleic Acids Res.* **1985**, *13*, 6483–6498.

Slightom, J. S.; Sun, S. M.; Hall, T. C. Complete Nucleotide Sequence of a French Bean Storage Protein Gene: Phaseolin. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 1897–1901.

Somerville, C. R.; McCourt, P.; Kunst, L.; Browse, J. Mutants of Arabidopsis Deficient in Fatty Acid Desaturation. In *The Metabolism, Structure and Function of Plant Lipids*; Stumpf, P. K., Mudd, J. B., Ness, W. D., Eds.; Plenum Press: New York, 1987; pp 669–676.

Spencer, D.; Higgins, T. J. V. Seed Maturation and Deposition of Storage Proteins. In *Molecular Biology of Plant Cells*; Smith, H., Ed.; Botanical Monographs 14; Blackwell: Oxford, U.K., 1982; pp 306–336.

Start, W. G.; Ma, Y.; Polacco, J. C., et al. Two Soybean Seed Lipoxygenase Nulls Accumulate Reduced Levels of Lipoxygenase Transcripts. *Plant Mol. Biol.* **1987**, *7*, 11–23.

Sun, S. M.; Slightom, J. L.; Hall, T. C. Intervening Sequences in a Plant Gene: Comparison of the Partial Sequence of cDNA and Genomic DNA of French Bean Phaseolin. *Nature* **1981**, *289*, 37–41.

Takaiwa, F.; Kikuchi, S.; Oono, K. The Structure of Rice Storage Glutelin Precursor Deduced from cDNA. *FEBS Lett.* **1986**, *206*, 33–35.

Takaiwa, F.; Kikuchi, S.; Oono, K. A Rice Glutelin Gene Family—A Major Type of Glutelin mRNAs Can Be Divided Into Two Classes. *MGG, Mol. Gen. Genet.* **1987a**, *208*, 15–22.

Takaiwa, F.; Ebinuma, H.; Kikuchi, S.; Oono, K. Nucleotide Sequence of a Rice Glutelin Gene. *FEBS Lett.* **1987b**, *221*, 43–47.

Talbot, D. R.; Adang, M. J.; Slightom, J. L.; Hall, T. C. Size and Organization of a Multigene Family Encoding Phaseolin, the Major Seed Storage Protein of Phaseolus vulgaris. *MGG, Mol. Gen. Genet.* **1984**, *198*, 42–49.

Taliercio, E. W.; Chourey, P. S. Post-transcriptional Control of Sucrose Synthase Expression in Anaerobic Seedlings of Maize. *Plant Physiol.* **1989**, *90*, 1359–1364.

Thornburg, R. W.; An, G.; Cleveland, T. E.; Johnson, R.; Ryan, C. A. Wound Inducible Expression of a Potato Inhibitor I—Chloramphenical Acetyl Transferase Gene Fusion in Transgenic Tobacco Plants. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 744–748.

Toenniessen, G. H. Plant Molecular Biology and the International Agricultural Research System. *Plant Mol. Biol. Rep.* **1985**, *3*, 1–11.

Topfer, R.; Gronenborn, B., et al. Uptake and Transient Expression of Chimeric Genes in Seed Derived Embryos. *Plant Cell* **1989**, *1*, 133–139.

Tsai, C. Y.; Nelson, O. E. Starch Deficient Maize Mutants Lacking ADP-glucose Pyrophosphorylase Activity. *Science* **1966**, *151*, 341–343.

Uchimiya, H.; Fushimi, T., et al. Expression of a Foreign Gene in Callus Derived from DNA-treated Protoplasts of Rice (Oryza sativa) *MGG, Mol. Gen. Genet.* **1986**, *204*, 204–207.

Uchimiya, H.; Handa, T.; Brar, D. S. Transgenic Plants. *J. Biotechnol.* **1989**, *12*, 1–20.

Viotti, A.; Abildsten, D.; Pogna, N.; Sala, E.; Pirotta, V. Multiplicity and Diversity of Cloned Zein cDNA Sequences and Their Chromosomal Localization. *EMBO J.* **1982**, *1*, 53–58.

Walling, L.; Drews, G. N.; Goldberg, R. B. Transcriptional and Posttranscriptional Regulation of Soybean Seed Protein mRNA Levels. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 2123–2127.

Weising, K.; Schell, J.; Kahl, G. Foreign Genes in Plants: Transfer, Structure, Expression and Application. *Annu. Rev. Genet.* **1988**, *22*, 421–477.

Wienand, U.; Feix, G. Zein Specific Restriction Enzyme Fragments of Maize DNA. *FEBS Lett.* **1980**, *116*, 14–16.

Williamson, M. S.; Forde, J.; Buxton, B.; Kreis, M. Nucleotide Sequence of Barley Chymotrypsin Inhibitor-2 (CL-2) and its Expression in Normal and High-Lysine Barley. *Eur. J. Biochem.* **1987**, *165*, 99–106.

Yang, M. S.; Espinoza, N. O.; Nagpala, P. G.; Dodds, J. H.; White, F. F.; Schnorr, K. L.; Jaynes, J. M. Expression of a Synthetic Gene for Improved Protein Quality in Transformed Potato Plants. *Plant Sci.* **1989**, *64*,.99–111.

Yano, M.; Okuno, K.; Kawakami, J.; Satoh, H.; Omura, T. High Amylose Mutants of Rice, Oryza sativa. *Theor. Appl. Genet.* **1985**, *69*, 253–257.

Received for review December 27, 1989. Revised manuscript received April 13, 1990. Accepted April 23, 1990.