BNL-24704

CONF-780955--2

INTERNATIONAL SYMPOSIUM ON SEED PROTEIN IMPROVEMENT
IN CEREALS AND GRAIN LEGUMES
Neuherberg, Federal Republic of Germany
4-8 September 1978
IAEA-SM-230/72

IDENTIFICATION OF ZEIN STRUCTURAL GENES IN THE MAIZE GENOME

Benjamin Burr
Department of Biology
Brookhaven National Laboratory
Upton, NY 11973, U.S.A.

ABSTRACT

Zein is the major storage protein in maize endosperm and is deficient in the amino acids lysine and tryptophan - amino acids generally limiting in cereal grains. If the two subunits of zein were each encoded by a single gene, it is possible to envisage a program in which altered gene products containing these two amino acids could be selected. Since zein constitutes such a large proportion of the total protein, substitutions of rare amino acids into zein will be reflected in the overall amino acid composition of the kernel.

At the moment we do not know how many genes for zein there are, nor do we know where they are located in the genome. In our approach to this question we have purified the messenger RNAs for the two zein polypeptide chains free from all other species of RNA. The mRNAs are polyadenylated, have a "cap" structure at their 5' terminus, and contain slightly less than 50% noncoding sequences. Their average length is 1.1 kilobases and we assume they are monocistronic. Using reverse transcriptase and DNA polymerase we prepared double-stranded DNA transcripts of the messages. These were inserted into the bacterial plasmid pMB9 and used to transfect E. coli. A number of clones carrying nearly full length inserts have been selected for study. These recombinant DNA molecules are being used as probes to accurately count the number of zein genes in the maize genome, check for amplification, and determine the chromosomal location of the genes.

DISCLAIMER

Zein accounts for 50% of the protein in the corn seed [1]. Therefore the amino acid composition of the corn kernel is to a large extent a reflection of zein amino acid composition. Amino acid residues which are absent from zein or present in small amounts depress the total level of that particular amino acid in the kernel. If the two zein proteins were the products of two genes, a mutation which caused the substitution of a rare amino acid for a common one in one of the zein polypeptide chains would be expressed in an important fraction of the endosperm protein and thus would be expected to alter the overall amino acid composition of the kernel. Before one can consider mutation breeding schemes which are aimed at altering the amino acid composition of the grain by using zein structural genes as the target, it is important to know how many zein genes there are. Our approach to this problem has been to prepare very pure, nearly full length copies of zein mRNAs which are being used in hybridization studies to measure the number of zein genes and to determine their chromosomal location.

There is evidence to suggest that zein may consist of only two polypeptide chains and might then be the product of only two structural genes. Misra [2] demonstrated that there are two discrete bands when zein is separated electrophoretically on the basis of molecular weight on SDS-polyacrylamide gels. This is in contrast to an enormous heterogeneity which is observed when zein is separated on the basis of net charge [3]. However, this apparent heterogeneity could be in part artifactual and in part generated post-synthetically:

1) Zein, being a very hydrophobic protein, aggregates in aqueous buffers and is difficult to dissociate, even in the presence of urea.
2) Both zein chains contain free sulfhydryls which form intermolecular disulfide bridges if electrophoresis is not done under strict reducing conditions or if the cysteine side chains are not blocked by alkylation.
3) Zein has an unusually high number of amidated side chains [4]. These residues normally become randomly deaminated during the life of the protein. During this process the protein gains net negative charge.
4) Newly translated zein polypeptides undergo proteolytic cleavage and glucosylation. Both of these modifications alter electrophoretic mobility and any variation in this process would lead to apparent heterogeneity.

The only apparent resolution to the problem of whether electrophoretic heterogeneity is a postsynthetic phenomenon is to examine the amino acid sequence of the polypeptides. We are currently attempting to determine if there is any sequence heterogeneity at the amino terminus of the zein light (19,000 dalton) chain.

We had previously reported that zein mRNA can be purified to apparent homogeneity [5]. The preparation contains at least two mRNAs coding for light or heavy chains of zein. These messages are 1100 to 1200 nucleotides long of which approximately 100 nucleotides is the 3' polyadenylate terminus (F.A. Burr, in preparation).

Techniques have been developed for synthesizing double-stranded DNA copies from a single-stranded mRNA [6]. Briefly, this involves making a complementary copy of the mRNA using the enzyme reverse transcriptase and completing the second strand with DNA polymerase. This DNA can be replicated by inserting it into circular bacterial plasmid DNA which is then

used to infect bacterial cells. This cloning process purifies each synthetic molecule to homogeneity since each bacterial colony is derived from a single transfection event. Tom St. John and Ron W. Davis of Stanford University have made double-stranded DNA copies of zein mRNA and inserted them into the plasmid pMB9 by adding poly(dA) tails to the cDNA and poly-(dT) tails to linearlized plasmid DNA. These recombinant molecules were used to transfect Eschericia coli. The plasmid carries resistance to tetracycline and bacteria infected with the plasmid can be selected in the presence of the antibiotic.

Twenty clones were selected for study. Their inserts ranged in size from 100 to 1300 base pairs including the synthetic poly(AT) tails. We used two methods to show that these recombinant plasmids contained sequences complementary to zein mRNA:

1) The most direct method was to bind linearized recombinant plasmid DNA to a cellulose matrix [7]. This was exposed to zein mRNA. Non-specifically bound RNA was washed off and the specifically bound RNA was eluted and added to the wheat germ protein synthesizing system. The results of this approach indicated that 15 out of 19 plasmid DNAs bound zein mRNA for the light or heavy zein polypeptide chains under stringent hybridization conditions.
2) Plasmid DNA was hybridized with an excess of RNA under conditions which favor RNA-DNA hybrid formation over DNA-DNA interactions [8]. Nonhybridized material was digested with single strand nuclease and the hybrids were displayed on agarose gels [9]. Only two of the twenty plasmids failed to form hybrids with zein mRNA. Plasmids contain sequences from 200-900 nucleotides long which appear to be complementary to the zein mRNA.

DNA from selected plasmids will be radioactively labeled and used in quantitative hybridization experiments to measure the number of copies of zein structural genes in the maize genome. The probes can probably also be used to attempt to localize the zein genes to chromosome. Our approach is to prepare DNA from plants which are monosomic for each of the ten chromosomes of maize [10] and to use these in quantitative hybridization experiments. DNA from a plant lacking one of the chromosomes encoding zein is expected to hybridize half as much probe as DNA from a diploid plant.

Assuming that there are only two zein structural genes, it will be necessary to make multiple amino acid substitutions in any one gene before there can be an appreciable effect on total kernel amino acid composition. We estimate that it will be necessary to add two lysines and two tryptophans in each of the two proteins to double the overall level in the kernel. This could be done if the zein gene chromosomal location were known so that recombination between close outside markers could be used to monitor potential intragenic recombinants.

ACKNOWLEDGEMENTS

Research was performed at Brookhaven National Laboratory under the auspices of the U.S. Department of Energy and was supported in part by Grant GM 240057 from the National Institutes of Health.

REFERENCES

[1] NELSON, O.E., Fed. Proc. 25 (1966) 1676.

[2] MISRA, P.S., MERTZ, E.T., GLOVER, D.V., "Characteristics of proteins in single and double endosperm mutants of maize", High Quality Protein Maize (BAUMAN, L.F., et al., Eds), Dowden, Hutchinson and Ross, Inc., Stroudsburg, Pennsylvania (1975) 291.

[3] SOAVE, C. et al., Maydica 21 (1976) 61.

[4] WALL, J.S., "Cereal Proteins", Proteins and Their Reactions (SCHULTZ, H.W., ANGLEMIER, A.F., Eds), Avi Publishing Co., Westport, Connecticut (1964) 315.

[5] BURR, B., et al., Proc. Nat Acad. Sci., U.S. 75 (1978) 696.

[6] MANIATIS, T., et al., Cell 8 (1976) 163.

[7] NOYES, B.E., STARK, G.R., Cell 5 (1975) 301.

[8] CASEY, J., DAVIDSON, N., Nucleic Acids Res. 4 (1977) 1539.

[9] BERK, A.J., SHARP, P.A., Cell 12 (1977) 721.

[10] PLEWA, M.J., WEBER, D.F., Genetics 81 (1975) 277.