In plant seeds, the storage triacylglycerol is packaged and stored in organelles called lipid bodies. These spherical lipid bodies are approximately 0.5 μm in diameter and are composed of a central core of triacylglycerol surrounded by a "half-unit" membrane consisting of a single layer of phospholipid and a small set of tightly associated proteins (1-3). The storage of seed lipids in many small subcellular organelles rather than in a single large lipid droplet/cell as seen in mammalian white adipose tissue may be related to the need for rapid mobilization of these reserves during seed germination and seedling growth. Partitioning a single large lipid sphere into many small ones greatly increases the surface area/volume of lipid. This increased surface area may render the seed lipid more accessible to degradative enzymes and facilitate its rapid mobilization. The proteins associated with lipid bodies and their role in lipid packaging and metabolism have not been well-studied.

We have been using the maize seed as a model system to study lipid bodies and their associated proteins (4-7). The lipid bodies in maize are primarily confined to a specialized embryonic tissue called the scutellum, where they comprise 50% of the cell dry weight at seed maturity. Only four major polypeptides are associated with maize lipid bodies, one of higher M_r (45,000) called H protein and three of lower M_r (15,500, 18,000, and 19,500) called L3, L2, and L1 proteins, respectively. The smallest of the L proteins (L3) is quite abundant, constituting nearly 30% of the total lipid body protein and approximately 1% of total protein in the scutella.

The structure of the proteins and the nature of their association with one another and the lipid and phospholipid components of the lipid body is unknown.

We have cloned and sequenced a nearly full-length cDNA for L3, the major protein associated with the lipid bodies of maize. The deduced amino acid sequence was characterized according to local hydrophathy and its secondary structure predicted by computer analysis. These analyses suggest that there are three structural domains in L3, and allow us to propose a simple model for the topology for L3 on the lipid body. The structure of the protein is unique but has features in common with both mammalian apolipoproteins and intrinsic membrane proteins.

**EXPERIMENTAL PROCEDURES**

**RESULTS**

**Identification of L3-cDNA**

A maize cDNA library was constructed from maturing maize embryo RNA, and putative L3 clones were selected as described under "Experimental Procedures." The identity of the clones was determined by hybrid-selected translation analysis. Several clones specifically selected an mRNA which was translated into a polypeptide with the approximate M_r of L3 (data shown for one of the clones, Fig. 1). The hybrid-selected translation product was immunoprecipitated with polyclonal antiserum specific to the L3 protein (Fig. 1). Little or no translatable mRNA was selected by nonrecombinant vector.

The nucleotide sequence of an L3-cDNA clone was determined by the dideoxynucleotide chain termination method. The sequencing strategy is shown in Fig. 2 and the nucleotide sequence and deduced amino acid sequence are shown in Fig.
3. The cDNA insert is 757 bp long excluding the poly(dA) tract and hybridizes to a maize scutellar mRNA approximately 1000 nucleotides long (Fig. 4A). These data are consistent with primer extension analysis which indicates that the cDNA is 86% complete, missing approximately 125 nucleotides at the 5' end (Fig. 4B). The GC content of the cDNA sequence is quite high, 62.6% overall and 71.6% in the coding region. The clone terminates in a poly(dA) tract preceded by a perfect polyadenylation signal (AATAAA) (20) located 13 bp upstream. Analysis of the period constraints (Fickett's statistic) of the cDNA was used to predict coding and non-coding regions independent of reading frame (21). This analysis suggests with 95% confidence that the first 400 bp of the cDNA clone are protein coding while the last 300 bp are noncoding. The L3-cDNA contains two long open reading frames, both of which contain this presumptive protein coding region. ORF-1 begins at nucleotide 1 and encodes 147 amino acids, while ORF-2 begins at nucleotide 3 and encodes 161 amino acids.

Two major lines of evidence indicate that ORF-1 encodes the L3 protein.

Codon Usage—1) The codon usage in ORF-1 is dramatically skewed (95.1%) in favor of C or G in the third or "wobble" position, a phenomenon which has been positively correlated with protein coding regions in prokaryotic and eukaryotic organisms (22). The codon usage for ORF-2 is not skewed. 2) Computer analyses indicate that ORF-1 uses preferred codons and does not use rare codons as determined from the combined codon usage tables of genes from related tissues in maize and other monocots (Ref. 23, data not shown). In contrast, ORF-2 uses rare codons and does not use preferred codons.

Protein Characteristics—1) The calculated amino acid composition of the ORF-1 deduced protein fits well with the previously reported (5), experimentally determined composition of L3 (Table 1). The amino acid composition of the ORF-2 deduced protein is inconsistent with the L3 values, especially agreement with the proposed initiator codon consensus sequence of L3 (Table 1). 2) ORF-1 encodes a hydrophobic protein, in agreement with published experimentally determined hydrophobicity of L3 (5). ORF-2 encodes a hydrophilic protein. 3) The proteinase (trypsin, Staphylococcal V8 protease) digestion patterns of L3 are consistent with those predicted by the ORF-1 deduced amino acid sequence (data not shown). The digestion patterns predicted by ORF-2 deduced amino acid sequence are quite different and do not fit the experimentally determined L3 data.

We were unable to determine directly if the clone contains the entire protein coding region since the amino-terminal sequence of L3 is not known. Direct sequence analysis indicates that the protein is blocked at the amino terminus. However, the first ATG in ORF-1 is located within a very GC-rich area, the sequence context of which is not in good agreement with the proposed initiator codon consensus sequences of eukaryotes (24). Furthermore, the skewed codon usage pattern noted in the coding region is continued in the region upstream from the ATG. These data suggest the clone may not contain the entire coding region for the L3 protein. An incomplete deduced amino acid sequence could account for minor discrepancies in the comparison of deduced and experimental amino acid composition.

**Analysis of Deduced L3 Amino Acid Sequence**

Local hydrophathy of the deduced L3 amino acid sequence was analyzed by the quantitative method of Kyte and Doolittle (25), which plots the average hydrophathy over a moving window of 7 amino acids. This analysis shows that the L3 protein may be divided into three domains based on the hydrophilic characteristics of the molecule (Fig. 5). Both the amino- and carboxyl-terminal domains are largely hydrophilic. An internal region of 72 amino acids consists entirely of hydrophobic, weakly hydrophobic, or neutral amino acids and constitutes a large hydrophobic domain.

Chou and Fasman (26) rules for empirical prediction of protein secondary structure were used to determine the potential of L3 to form α-helical, β-sheet, and β-turn conformations. This analysis suggests that the amino acids at or near the amino terminus (amino acids 15–38) of the protein and those at the carboxyl terminus (amino acids 115–147) will form α-helices. Empirical prediction of secondary structure in strongly hydrophobic regions is unreliable due to the absence of an appropriate data base of known structures of membrane-associated proteins (27, 28).

When the carboxyl-terminal 33 amino acids are plotted according to their individual hydrophathy, it can be seen that hydrophilic amino acids alternate with hydrophobic or neutral amino acids in a regular manner. This pattern of hydrophobicity was examined with respect to the α-helical structure of the region. The helix was visualized on a "helical wheel" (29) which shows the view looking down an α-helix with amino acids at the correct angular displacement from one another, and in a cylindrical plot (30) which shows the longitudinal arrangement of amino acids using a diagram of a cylinder split parallel to its axis and laid flat. The helical wheel representation of the carboxyl-terminal 18 amino acids is shown in Fig. 6A and the cylindrical plot visualization of the same region is shown in Fig. 6B. These analyses suggest that the L3 carboxyl-terminal α-helix is amphipathic, since hydrophobic amino acids are separated from hydrophilic amino acids on opposing sides of the helix.

**Model for L3 Topology on Lipid Body**

We suggest the following simple model for the topology of L3 on the lipid body based on predicted secondary structure and hydrophobicity analyses (Fig. 7). L3 is assumed to be a monomer since there is no data available on protein-protein interactions on the lipid body. The sequence at or near the amino terminus is predominantly hydrophilic and is predicted to form an α-helix starting at amino acid 15. This hydrophilic domain is expected to protrude from the lipid body surface.
The carboxyl-terminal 18 amino acids of the projection of all side chain positions along the helix axis onto a plane. wheel or cylindrical plot. The polar and nonpolar faces are indicated. The basic residues at the interface and negatively charged residues along the center acidic residues at the center polar faces with positively charged amino acids occurring at the phosphate group of the lipid body, its nonpolar face interacting with hydrophobic acyl tails of surrounding phospholipid molecules and its polar face orienting toward the cytosol. The specific distribution of charged amino acids in the helix allows interaction of positively charged amino acids with the phosphate group of the phospholipids and of negatively charged amino acids with positively charged groups on the phospholipids (e.g. choline of phosphatidylcholine).

FIG. 6. The carboxyl-terminal \( \alpha \)-helix of L3 was examined for amphipathic characteristics by representation on a helical wheel or cylindrical plot. A, helical wheel representation of the carboxyl-terminal 18 amino acids of the L3 \( \alpha \)-helix. The wheel is the projection of all side chain positions along the helix axis onto a plane. The polar and nonpolar faces are indicated. The basic residues separating the opposing polar and nonpolar faces are circled. The acidic residues at the center of the polar face are boxed. B, the cylindrical plot visualization of the same region. The hydrophobic face is outlined.

FIG. 7. Proposed model for the topology of L3 protein on the lipid body. Cylinders are used to depict \( \alpha \)-helices. The phospholipid and triacylglycerol moieties of the lipid body are demarcated into the cytosol. The central domain of the protein consists of 72 contiguous hydrophobic or neutral amino acids and is expected to penetrate the triacylglycerol core of the lipid body due to hydrophobic interactions. This hydrophobic domain is represented in a helical hairpin conformation for several reasons. The helical conformation (either \( \alpha \) or \( \beta \)) satisfies hydrogen-bonding requirements of the peptide backbone and represents the most energetically feasible conformation in a nonpolar environment (31). The division of the hydrophobic region into a pair of antiparallel helices connected by a bend is predicted in view of the presence of a central region of 12 amino acids containing 3 prolines which serve to interrupt the helical conformation and facilitate the turn. The carboxyl-terminal domain of L3 has certain features which render it suitable for interaction with surface phospholipids according to the amphipathic \( \alpha \)-helix hypothesis (32). It is predicted to be an amphipathic \( \alpha \)-helix containing exposing polar and nonpolar faces with positively charged amino acids occurring at the interface and negatively charged residues along the center of the polar face (Fig. 6). The helix may sit at the surface of the lipid body, its nonpolar face interacting with hydrophobic acyl tails of surrounding phospholipid molecules and its polar face orienting toward the cytosol. The specific distribution of charged amino acids in the helix allows interaction of positively charged amino acids with the phosphate group of the phospholipids and of negatively charged amino acids with positively charged groups on the phospholipids (e.g. choline of phosphatidylcholine).

DISCUSSION

Comparison with Other Lipid-associated Proteins—The lipid body of maize is composed of triacylglycerol surrounded by a layer of phospholipid and a set of specific proteins. Similar lipid-phospholipid-protein complexes termed lipoproteins are present in the plasma of mammals. The lipoproteins play a number of roles, including lipid transport, direction of lipids to certain tissues, mediation of lipid accessibility to various enzymes, and regulation of enzyme activities (33–35). The protein components of lipoproteins, called apolipoproteins, are a group of related proteins which have been studied extensively. Many apolipoproteins are composed of tandem amino acid repeats which form the dominant structure of the proteins, an amphipathic \( \alpha \)-helix with charge distributions suitable for interaction with phospholipids (36).

Since L3-protein may play a similar structural and/or functional role with the apolipoproteins, its structure was compared to that of the known apolipoproteins at both the amino acid and cDNA level. As a unique feature, the structure of L3 is dominated by a completely hydrophobic domain, which is unparalleled in the apolipoproteins, and does not show the internally repeated sequences which are so common in the apolipoproteins. On the other hand, L3 does contain an amphipathic \( \alpha \)-helix at its carboxyl terminus, which is composed of 33 amino acids. If homology comparison is extended to include related amino acids, the L3 amphipathic \( \alpha \)-helix shares 24.2% homology with a conserved common block of 33 amino acids encoded by a portion of exon 3 of most apolipoproteins (37). The homology at the above-mentioned regions of the polypeptides extends to some degree to the nucleic acid level. The L3 mRNA shares 33–46% homology with the conserved region of various apolipoprotein genes, as determined using the local homology algorithm of Smith et al. (38). However, these homologies may simply reflect the common structural constraints involved in producing an amphipathic helix.

The intrinsic membrane proteins form another class of lipid-associated proteins. Many of these proteins contain simple lipophilic transmembrane segments which serve as membrane anchors (31, 39, 40). The transmembrane domains are characteristically composed of 20-33 uncharged and largely nonpolar amino acids flanked on both sides by charged amino acids. The internal domain of L3 is quite similar to this characteristic simple transmembrane domain in that it is a large hydrophobic region bounded on either side by a charged amino acid. The major difference is the length of the hydrophobic region, which is unusually long in L3. Transmembrane domains of intrinsic membrane proteins are tailored to the width of the membrane. In contrast, proteins associated with the lipid body need only penetrate a single layer of phospholipid into a totally hydrophobic environment.

Possible Function of L3-polypeptide—Although the exact function of L3 is unknown, the unusual structure of the protein and its abundance on the lipid body suggest that it may serve a "structural" role to stabilize the lipid body. Thus, the phospholipid coat and associated proteins serve to package the stored lipid in small discrete particles, providing maximal surface area such that the lipid can be rapidly degraded during germination.

It is also possible that L3 itself, or in association with one or more of the other lipid body proteins, serves functions other than or in addition to this structural "packaging" func-
tion. The lipids are hydrolyzed during germination by the enzyme lipase, which is synthesized de novo on free polysomes and attaches itself specifically to the surface of the lipid body (6). The lipid body proteins, singly or in association, may act as the recognition signal for the specific binding of lipase to the lipid bodies.

Acknowledgments—We thank Dr. Lewis Bowman for excellent technical advice and critical reading of the manuscript, and Dr. Lucas Lebioda for helpful discussions of protein structures.

REFERENCES
1. Huang, A. H. C. (1985) in Modern Methods of Plant Analysis (Linskens, H. F., and Jackson, J. F., eds) Vol. 1, pp. 145–151, Springer-Verlag, Berlin.
2. Yatsu, L. Y., and Jacks, T. L. (1972) Plant Physiol. 49, 937–943.
3. Gurr, M. I. (1980) in The Biochemistry of Plants (Stumpf, P. K., and Conn, E. E., eds) Vol. 4, pp. 205–248, Academic Press, Orlando, FL.
4. Stymne, S., and Stobert, A. K. (1987) in The Biochemistry of Plants (Stumpf, P. K., and Conn, E. E., eds) Vol. 10, Academic Press, Orlando, FL, in press.
5. Qu, R., Wang, S., Lin, Y., Vance, V. B., and Huang, A. H. C. (1986) Biochem. J. 235, 57–65.
6. Wang, S. M., and Huang, A. H. C. (1987) J. Biol. Chem. 262, 2270–2274.
7. Huang, A. H. C. (1987) in The Biochemistry of Plants (Stumpf, P. K., and Conn, E. E., eds) Vol. 10, Academic Press, Orlando, FL, in press.
8. Wickena, M. P., Buell, G. N., and Schimke, R. T. (1978) J. Biol. Chem. 253, 2483–2495.
9. Vance, V. B., and Beachy, R. N. (1984) Virology 138, 26–36.
10. Young, R. A., and Davis, R. W. (1983) Proc. Natl. Acad. Sci. U. S. A. 80, 1194–1198.
11. Young, R. A., and Davis, R. W. (1983) Science 222, 778–782.
12. Dougherty, W. G., and Hiebert, E. (1980) Virology 101, 464–474.
13. Yamamoto, K. R., Alberts, B. M., Benzing, R., Lawhorne, L., and Treiber, G. (1970) Virology 40, 734–744.
14. Ricciardi, R. P., Miller, J. S., and Roberts, B. E. (1979) Proc. Natl. Acad. Sci. U. S. A. 76, 4927–4931.
15. Weber, K., and Osborn, M. (1969) J. Biol. Chem. 244, 4406–4412.
16. Sanger, F., Nicklen, S., and Colson, A. R. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 5463–5467.
17. McMasters, G. K., and Carmichael, G. G. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 4835–4838.
18. Thomas, P. S. (1980) Methods Enzymol. 100, 255–266.
19. Vance, V. B., Thompson, E. A., and Bowman, L. H. (1985) Nucleic Acids Res. 13, 7499–7513.
20. Proudfoot, N. J., and Brownlee, G. G. (1976) Nature 263, 211–214.
21. Fickett, J. W. (1982) Nucleic Acids Res. 10, 5303–5318.
22. Bibb, M. J., Findlay, P. R., and Johnson, M. W. (1984) Gene (Amst.) 30, 157–166.
23. Grishkow, M., Devereux, J., and Burgess, R. R. (1984) Nucleic Acids Res. 12, 539–548.
24. Heidecker, G., and Messing, J. (1986) Annu. Rev. Plant Physiol. 37, 439–466.
25. Kyte, J., and Doolittle, R. F. (1982) J. Mol. Biol. 157, 105–132.
26. Chou, P. Y., and Fasman, G. D. (1978) Annu. Rev. Biochem. 47, 251–276.
27. MacLennan, D. H., Brandl, C. J., Korczak, B., and Green, N. M. (1986) Nature 316, 696–700.
28. Green, N. M., and Flanagan, M. T. (1976) Biochem. J. 153, 729–732.
29. Schiffer, M., and Edmundson, A. B. (1976) Biophys. J. 7, 121–125.
30. Taylor, J. W., Miller, R. J., and Kaiser, E. T. (1982) Mol. Pharmacol. 22, 657–666.
31. Engleman, D. M., and Steltz, T. A. (1981) Cell 23, 411–422.
32. Sekrest, J. P., Jackson, R. L., Morrisett, J. D., and Gotto, A. M. (1974) FEBS Lett. 38, 247–253.
33. Kostner, G. M. (1983) Adv. Lipid Res. 20, 1–43.
34. Breslow, J. L. (1985) Annu. Rev. Biochem. 54, 699–727.
35. Goldstein, J. L., and Brown, M. S. (1983) in The Metabolic Basis of Inherited Disease (Stanbury, J. B., Wyngaarden, J. B., Fredrickson, D. S., Goldstein, J. L., and Brown, M. S., eds) 5th Ed, pp. 672–712, McGraw-Hill, New York.
36. Rajavashisth, T. B., Kaptein, J. S., Reue, K. L., and Luisi, A. J. (1985) Proc. Natl. Acad. Sci. U. S. A. 82, 8085–8089.
37. Luo, C.-C., Li, W.-H., Moore, M. N., and Chan, L. (1986) J. Mol. Biol. 187, 325–340.
38. Smith, T. F., Waterman, M. S., and Sadler, J. R. (1981) Adv. Appl. Math. 2, 482–489.
39. Rice, C. M., Bell, J. R., Hunkapiller, M. W., Strauss, E. G., and Strauss, J. H. (1982) J. Mol. Biol. 154, 355–378.
40. Eisenberg, D. (1984) Annu. Rev. Biochem. 53, 595–623.
SUPPLEMENTAL MATERIALS TO:

The Major Protein from Lipid Bodies of Maize: Characterization and structure based on cDNA Cloning

Vicki Bowman Vance and Anthony H.C. Huang

EXPERIMENTAL PROCEDURES

Construction of a hgtll maize embryo cDNA library

Poly A RNA was isolated from mature maize embryos (25–30 days after pollination as previously described (4)). The RNA (15 μg) was reverse transcribed in a 100 μl reaction medium containing 50 μM Tris-HCl, pH 8.3, 8 μg oligo(dT) 15, 0.4 mg/ml RNase-free bovine serum albumin and primed with a synthetic double-stranded primer, 5'-CACCGACGGGAGCTTTCTTGACCG-3' (large fragment) and packaged in vitro into λ phage (Gigapack, Vector Cloning System), and the library amplified on E. coli strain Y109 (12,13). The resulting cDNA library contained 850,000 phage, approximately 25% of which were recombinant.

Screening of maize embryo cDNA library

The library was screened by hybridization with 32P-labeled cDNA made from poly A RNA enriched in L3-mRNA. Poly A RNA was isolated from mature maize embryos (3) and size fractionated by sucrose gradient centrifugation. RNA fractions were translated in the wheat germ in vivo translation system (11) and the transcription products analyzed by SDS-PAGE both before and after immunoprecipitation (12) with nonspecific polyclonal L3 antibodies (3). The enriched L3-mRNA fraction used to construct the λ2P-λDNA probe was estimated to contain L3-mRNA as approximately 40% of the total translatable RNA.

Hybrid-selected translation

Phage 121E GM125 (14) from a tentative L3-CDNA clones (13) was bound to nitrocellulose paper. Poly A RNA (30 μg/ml) from mature maize embryos was hybridized to these immobilized RNAs at 42°C for 16 hrs (14). The filters were washed at 60°C in 2X SSC and hybridized RNA eluted in a solution of 0.5 μM RNA, 10 μg/ml yeast tRNAs at 45°C for 2 hrs. Hybrid-selected clones were used to construct the in vitro translation system. In vitro translation products were analyzed by SDS-PAGE (15) and the gels impregnated with PPO, dried and fluorographed.

Sequencing of λ2P-λDNA

Restriction fragments were subcloned into various M13 vectors using E. coli strains M13K07 or 2350K as host. Each sequencing was performed by the dideoxy nucleotide chain termination method (8). Clones were extracted, ethanol precipitated, dissolved in formamide at 90°C for 3 minutes and annealed 12 hr at 4°C. The minihairpin loop was formed by a 20-fold molar excess of reverse transcriptase. The hairpin loop was then isolated from the in vitro translation system (4) and purified by digestion with 0.3 M NaCl. cDNAs larger than 300 bp were ligated to EcoRI Cut ligation solution of 2 μM tris-HCl, pH 8.3, 0.5 μM MgAc. The synthetic 19 nucleotide single-stranded primer was end-labeled primer, heated at 70°C for 3 minutes and annealed 12 hr at 4°C. The primer was 5'-GGAGCCTCGAGGAGATGACG-3' (large fragment) and hybridized to these immobilized DNA at 42°C for 16 hrs (14). The derived DNA was bound to the respective nucleic acids. Lines are drawn above and below the polyadenylation sequence showing denaturing gel electrophoresis analysis of extended DNA derived by reverse transcription of a synthetic primer on poly A+ RNA from maturing maize embryos. The primer was a synthetic single-stranded oligonucleotide (19-mer) complementary to the region shown in Figure 3 (nucleotide positions 81-99). Lane 1. yeast tRNA added to the reaction mix. Lane 2. maize poly A+RNA added to the reaction mix. The migration of molecular weight markers (in bp) is indicated to the right of the figure. The major band at the L3 position in lane 2 is caused by the reverse transcriptase to stop extending. However, it could represent full extension of a minor and slightly shorter 3.5-kb L3-CDNA. The very faint band at the L3 position in lane 1 may be background caused by the reverse transcriptase to stop extending. However, it could represent full extension of a minor and slightly shorter 3.5-kb L3-CDNA.

Computer analyses of nucleotide and amino acid sequences

All sequence analyses were performed using the Sequence Analysis Software Package of the Wisconsin Genetics Computer Group (Version 6).

Figure 1. Fluorograph of SDS-PAGE analysis of in vitro translation products obtained from hybrid-selected λ2P-λDNA clones. Lane 1. In vitro translation products of RNA selected by hybrid-selected λ2P-λDNA clones. The translation products were immunoprecipitated with polyclonal L3 antibodies specific to the L3 protein. Lane 2. λ2P-λDNA clones of L3-CDNA. Lane 3. In vitro translation products of RNA selected by nonhybridizing λ2P-λDNA clones. Lane 4. In vitro translation products of partially purified L3-CDNA used to screen the λ2P-λDNA library. Lane 3, endogenous in vivo translation products of the wheat germ system used in this experiment. (The very faint band at the L3 position in lane 3 may be background or a minor contaminant from other lanes.)

Figure 2. Partial restriction map and sequencing strategy of λ2P-λDNA clones. Arrow denotes direction of sequencing. The left and right borders represent the E. coli sites used to insert the phage λ DNA into the vector. Lane A; Lane B, bp, base pairs.

Figure 3. Characterization of the λ2P-λDNA by northern and primer extension analyses. A. Northern blot of total RNA (5 μg) from maturing maize embryos (lane 2) hybridized to nick-translated isolated insert of λ2P-λDNA clone. Lane 1 shows the migration of radiolabeled RNA from a hgtll digest of the replicative form of λ phage. The top four bands, from top to bottom, represent the major bands in the hgtll banding pattern. Lane 2, derived by reverse transcription of a synthetic primer on poly A+ RNA from maturing maize embryos. The primer was a synthetic single-stranded oligonucleotide (19-mer) complementary to the region shown in Figure 3 (nucleotide positions 81-99). Lane 1. yeast tRNA added to the reaction mix. Lane 2. maize poly A+RNA added to the reaction mix. The migration of molecular weight markers (in bp) is indicated to the right of the figure. The major band at the L3 position in lane 2 is caused by the reverse transcriptase to stop extending. However, it could represent full extension of a minor and slightly shorter 3.5-kb L3-CDNA.

Table 1. Comparison of cDNA derived amino acid sequence of L3 with experimental data.

| Amino acid | Experimentally Determined | cDNA Derived | Experimental Data |
|------------|--------------------------|--------------|------------------|
| ala        | 12.4                     | 17.6         | 15.5             |
| asp        | 9.1                      | 4.1          | 6.3              |
| gln        | 5.7                      | 2.7          | 5.0              |
| glu        | 12.0                     | 14.2         | 17.4             |
| his        | 4.0                      | 2.7          | 6.8              |
| leu        | 3.2                      | 4.1          | 3.7              |
| lys        | 4.8                      | 4.7          | 6.1              |
| met        | 9.0                      | 10.8         | 8.6              |
| pro        | 2.8                      | 3.4          | 10.6             |
| ser        | 3.7                      | 3.3          | 21.7             |
| thr        | 7.3                      | 5.4          | 2.4              |
| tyr        | 4.8                      | 5.0          | 0.7              |
| val        | 5.6                      | 6.8          | 5.6              |
| tyr        | 0.7                      | 2.7          | 0.7              |
| gln        | 9.9                      | 7.4          | 6.8              |

The mole percent of each amino acid was calculated from the derived amino acid sequence of 120 bp of cDNA. The experimental values shown for L3 are from the work of Qu et al. (5).