Genomic Nucleotide Sequence of a Brassica napus 20-Kilodalton Oleosin Gene

Keunmyoung Lee and Anthony H. C. Huang*
Department of Botany and Plant Sciences, University of California, Riverside, California 92521

The storage triacylglycerols in seeds are confined to discrete spherical organelles called oil bodies (10). Each oil body of 1 μm diameter contains a matrix of triacylglycerols surrounded by a layer of phospholipids embedded with abundant and unique proteins called oleosins (2, 6). Oleosins are hydrophobic proteins of low M, ranging from 16 to 26 kD, depending on the isoforms and plant species. They are present in the seeds of both dicotyledons and monocotyledons. Their possible functions include stabilizing the oil bodies in an aqueous environment and providing signal recognition for the specific binding of lipase during germination. The partial amino acid sequences of the oleosins from maize (16 kD) (9) and Brassica (20 kD) (4) and the complete amino acid sequences of the oleosins from maize (18 kD) (5) and carrot (19 kD) (1) have been deduced from their cDNA and genomic sequences, respectively.

Here we report the complete nucleotide sequence of a Brassica napus oleosin gene (Fig. 1). The Brassica oleosin gene possesses one intron, whereas all the previously reported oleosin genes (two from maize and one from carrot) do not contain introns. The intron in the Brassica oleosin gene occurs between the two exons encoding the central hydrophobic domain and the C-terminal domain, respectively.

The deduced amino acid sequence of the Brassica oleosin shows the same three basic structural domains that are common to the four sequenced oleosins. These three domains are an N-terminal hydrophilic domain, a central hydrophobic domain, and a C-terminal amphipathic domain. The Brassica oleosin shares significant homology in amino acid sequences at the central hydrophobic domain, but not at the other two domains, with the other reported oleosins. The Brassica oleosin does not possess appreciable cleavable signal sequence at the N-terminus, and this is in agreement with its M, of 20,000 estimated by SDS-PAGE being similar to its M, of 20,682 derived from the amino acid sequence.

The N-terminal amino acid sequence of the Brassica oleosin bears little similarity with those of maize 18-kD oleosin and carrot 19-kD oleosin. This is intriguing, because the maize 18-kD oleosin is correctly targeted to the seed oil bodies in Brassica transformed with the maize oleosin gene (3). Presumably, the intracellular targeting signal for oleosin resides in

Table I. Characteristics of a 20-kD Oleosin Gene from Brassica napus

<table>
<thead>
<tr>
<th>Organism:</th>
<th>Brassica napus, L., var Bridger</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location on Chromosome:</td>
<td>Unknown</td>
</tr>
<tr>
<td>Function:</td>
<td>Oleosin on the surface of seed oil bodies serving to stabilize the organelles and possibly acting as a receptor for lipase binding during germination.</td>
</tr>
<tr>
<td>Gene designation:</td>
<td>BNO20</td>
</tr>
<tr>
<td>Source:</td>
<td>Obtained from a λ EMBL-3 genomic library (from Clontech, Palo Alto, CA) using maize 16-kD oleosin cDNA as a probe (9). Fragments subcloned in pUC118 and pUC119 and sequenced by dideoxy sequencing of both strands.</td>
</tr>
<tr>
<td>Methods of Identification:</td>
<td>Sequence identical with a segment of the amino acid sequence of an isolated oleosin (8).</td>
</tr>
<tr>
<td>Expression and Regulation:</td>
<td>Oleosin synthesized only in maturing seeds. Representing a few percent of the total seed proteins. Oleosin genes in maize (2), carrot (1), and Brassica (7) positively regulated by ABA. Putative ABA regulatory elements present at the 5’ upstream (Fig. 1).</td>
</tr>
<tr>
<td>Features of gene structure:</td>
<td>An open reading frame of 561 bp interrupted by an intron of 462 bp. Intron occurring at the junction of reading frame encoding the central hydrophobic domain and the C-terminal domain. Putative regulatory sequences underlined (Fig. 1).</td>
</tr>
</tbody>
</table>

Codon Usage:

- 51.3% G + C in the third nucleotide.
- (G + C) Content: 51.5% in the coding region.

Structural Features of Protein:

- ORF of 561 bp encoding a polypeptide of 187 amino acid residues of M, 20,682, a value similar to the M, of the isolated oleosin. Three structural domains present in the protein (N-terminal hydrophilic domain, central hydrophobic domain, and C-terminal amphipathic domain). Belonging to the "low M," oleosin isoform.

Antibodies:

- Prepared earlier (8).

Subcellular Location:

- Localized on the surface of oil bodies.

EMBL Accession No.

- M 63985

1 Support was received from National Science Foundation grant DMB 88-05618.

2 Abbreviation: bp, base pair.
Figure 1. A 2640-bp sequence of a B. napus 20-kD oleosin gene. The transcription initiation codon ATG is numbered as the first three nucleotides. The putative regulatory sequences are underlined, including TATA box, CAAT boxes, and the polyadenylation signal. The intron nucleotides. The putative regulatory sequences are underlined, including TATA box, CAAT boxes, and the maize 16-kD oleosin apparently belong to the same class of proteins associated with lipid bodies.  

The highly conserved central hydrophobic domain or in the yet-unidentified secondary structures in the N- or C-terminus. The seeds from different, and within the same, species (8). The predicted amino acid sequence is shown. The amino acid sequence, which was determined earlier by microsequencing (8), is underlined. The predicted amino acid sequence is shown. The amino acid sequence, which was determined earlier by peptide microsequencing (8), is underlined.  

LITERATURE CITED  


