Identification and characterization of carrot HAP factors that form a complex with the embryo-specific transcription factor C-LEC1

Katsumi Yazawa* and Hiroshi Kamada

Gene Research Center, Graduate School of Life and Environmental Sciences, University of Tsukuba, 1-1-1 Ten-noudai, Tsukuba, Ibaraki 305-8572, Japan

Received 17 June 2007; Revised 22 August 2007; Accepted 5 September 2007

Abstract

C-LEC1, an orthologue of Arabidopsis LEC1, is thought to be an essential transcriptional activator required for normal development during the early and late phases of embryogenesis. C-LEC1 is similar in sequence to the HAP3 subunits of other organisms. To understand C-LEC1 function better, a cDNA library of carrot somatic embryos was screened for factors that form complexes with C-LEC1. Two carrot HAP5 homologues and two carrot HAP2 homologues were identified; these factors have significant sequence similarity to the conserved regions of HAP5 and HAP2, respectively. Some of these proteins form heterotrimERIC complexes that bind specifically to DNA fragments containing a CCAAT sequence in vitro. The results suggest that C-LEC1 is a component of the CCAAT-box-binding factor and forms a complex with C-HAP2B and C-HAP5A or C-HAP5B that regulates gene expression during carrot embryo development.

Key words: Carrot, CCAAT-box, C-HAP, C-LEC1, Daucus carota, DNA binding, embryogenesis.

Introduction

Plant embryo development is divided into three processes: the establishment of embryo morphology; the accumulation of storage compounds; and the induction of desiccation tolerance and seed dormancy. In Arabidopsis thaliana, these processes are controlled by a network of transcription factors, including LEAFY COTYLEDON1 (LEC1), LEC1-LIKE (LIL), LEAFY COTYLEDON2 (LEC2), FUSCA3 (FUS3), and ABSCISIC ACID INSENSITIVE3 (ABI3) (To et al., 2006; Suzuki et al., 2007). Mutations in these genes produce various seed maturation defects, including defects in cotyledon identity (lec1, lec2, and fus3), abnormal chlorophyll degradation (abi3), abnormal anthocyanin accumulation (lec1, fus3, and lec2), desiccation intolerance (lec1, fus3, and abi3), and defects in oil and protein deposition (lec1, lec2, fus3, and abi3) (Giraudat et al., 1992; Lotan et al., 1998; Luerßen et al., 1998; Stone et al., 2001; Kwong et al., 2003).

LEC2, FUS3, and ABI3 encode plant-specific transcription factors that contain the B3 DNA-binding motif (Giraudat et al., 1992; Luerßen et al., 1998; Stone et al., 2001). These factors are thought to be involved in the accumulation of storage compounds and lipids, and in the transition from embryo morphogenesis to embryo maturation and the acquisition of desiccation tolerance (Delseny et al., 2001; Kagaya et al., 2005; Wang et al., 2007). In contrast, LEC1 and LIL encode homologues of the HAP3 subunit of the CCAAT-box-binding factor (CBF). LEC1 is thought to be a key regulator of embryonic development (Lotan et al., 1998; Kagaya et al., 2005). Although both LEC1 and LIL are essential for embryo development, they have distinct endogenous functions (Kwong et al., 2003). Recently, it was reported that LEC1 regulates the expression of ABI3 and FUS3 (Kagaya et al., 2005; Wang et al., 2007).

The CCAAT-box is a cis-acting element that is often found in the gene promoters of higher eukaryotes. This motif is recognized by a highly specific heteromeric factor.
variously called HAP, CBF, or NF-Y (Mantovani, 1999). This heteromeric complex contains at least three subunits: HAP2 (CBF-B, NF-YA), HAP3 (CBF-A, NF-YB), and HAP5 (CBF-C, NF-YC).

The HAP3 and HAP5 subunits form a tight heterodimer via protein–protein interactions similar to those of the histone-fold motif (HFM), a conserved protein–protein and DNA-binding interaction module (Luger et al., 1997). The formation of the HAP3/HAP5 heterodimer creates a surface for HAP2 association, resulting in the formation of the HAP2/HAP3/HAP5 heterotrimer, which can bind to CCAAT sequences with high specificity and affinity. In yeast, the HAP complex recruits a fourth subunit, HAP4, which contains a transcriptional activation domain, to the complex. In vertebrates, this domain is incorporated into other subunits (Forsburg et al., 1989).

Compared with yeast and vertebrates, little is known about the functional mechanisms of plant CBFs. CCAAT-like motifs are found in several plant promoters, and binding activity to CCAAT sequences has been identified in plant nuclear extracts. Unlike yeast and vertebrates, which have one gene that encodes each HAP subunit, plants have multiple genes for each of the HAP2, HAP3, and HAP5 proteins (Edwards et al., 1998; Gusmaroli et al., 2001, 2002). However, it is unclear whether the plant HAP2, 3, and 5 proteins can form functional trimeric complexes and bind to CCAAT sequences. The Arabidopsis (At) HAP2, HAP3, and HAP5 gene families (known as AtNF-YA1-10, -YB1-10, and -YC1-9, respectively) have some members that are ubiquitous and others that are tissue specific and induced only after the switch to reproductive growth in flowers and siliques (Gusmaroli et al., 2001, 2002). In rice (Oryza sativa; Os), 11 HAP3 genes have been identified through cDNA cloning and database searching. Four OsHAP3 genes (OsHAP3-A–C and OsNF-YB1) have been characterized, and at least one of these genes regulates chloroplast development (Miyoshi et al., 2003). Multiple HAP subunits in higher plants may associate with each other in various combinations to create CBFs that regulate the expression of specific gene sets. However, the large number of possible combinations has hindered the analysis of plant HAP complexes.

The characterization of the HAP2 and HAP5 subunits that may form functional complexes with LEC1 during embryogenesis is necessary to understand the functional mechanisms of LEC1. However, a precise analysis of these genes during embryogenesis is hindered by the very small areas in which embryogenesis occurs within flowers and immature fruits in the early stages of development and by the enclosure of the embryo in endosperm. Somatic embryogenesis is a simple experimental model system with which to analyse plant embryogenesis. Since the first protocols for the induction of somatic embryogenesis were described in 1958 (Reinert, 1958; Steward et al., 1958), carrot embryogenesis has become one of the most widely used model systems. In Arabidopsis, several methods of somatic embryogenesis have been developed, including those involving induction from immature zygotic embryos, protoplasts of leaf-derived cells, and stress-treated shoot apical tip explants (O’Neill et al., 1993; Luo et al., 1997; Ikeda-Iwai et al., 2002, 2003). Compared with the Arabidopsis system, the carrot system offers greater simplicity and can be used to obtain large numbers of developmentally synchronized somatic embryos. Several factors related to carrot embryogenesis have been identified, including carrot-ABI3 (C-ABI3) and carrot embryogenic cell proteins (DeECPs) (Zhu et al., 1997; Shiota et al., 1998). Recently, a carrot LEC1 orthologue (C-LEC1) (Yazawa et al., 2004) has been isolated, which encodes the HAP3 subunit and can complement the Arabidopsis lec1 mutant.

Carrot somatic embryogenesis as a model system was used to identify gene expression programmes that are associated with the initiation of embryo development in plants. A screen was carried out for carrot HAP2 and HAP5 factors that are expressed in developing somatic embryos. These factors were able to interact with C-LEC1 to form heterotrimERIC complexes that bind specifically to DNA fragments containing the sequence CCAAT.

Materials and methods

Plant materials

Carrot (Daucus carota L. cv. US-Harumakigosun) seedlings were grown at 25 °C under a 16 h light/8 h dark cycle of white light at 50 μmol photons m⁻² s⁻¹. Carrot embryogenic cells (ECs), non-embryogenic cells (NCs), and somatic embryos (SEs) were isolated as described by Satoh et al. (1986). Mature leaves were collected from plants that had been grown for 2 months at 25 °C under the above light conditions. Seeds were collected from carrot (D. carota L. cv. Yohimegosun) plants grown in the field at Takii Seed Co., Ltd (Ryugasaki, Ibaraki, Japan).

Yeast strains, media, and transformation

The Saccharomyces cerevisiae strain AH109 used for yeast two-hybrid screening was purchased from Clontech. The hap2 knockout strain was purchased from Open Biosystems. All yeast strains were grown at 30 °C on YPD medium (1% yeast extract, 1% peptone, and 2% glucose, pH 5.8). Transformants were selected on SD medium [0.76% yeast nitrogen base without amino acids (Difco, Detroit, MI, USA) and 2% glucose, pH 5.8], which was supplemented with amino acids when required. Functional complementation was selected by growth on lactate medium (1% yeast extract, 1% peptone, and 2% lactate, pH 4.8). Standard yeast genetic manipulations were performed according to the Yeast Protocols Handbook (Clontech).

Construction of the carrot cDNA library and two-hybrid screening

A carrot SE cDNA library was constructed 7 d after induction using the HybriZAP-2.1 XR Library Construction and HybriZAP-2.1 XR cDNA Synthesis Kits (Stratagene) and cloned into the phagemid vector pAD-GAL4-2.1. Screening, selection of putative positive
clones, and β-galactosidase assays were performed according to the Yeast Protocols Handbook (Clontech). The bait, i.e. C-LEC1 1-137 (C-LEC1 amino acids 1–137), was cloned into the pGADT7 vector from the Matchmaker GAL4 Two-Hybrid System 3 (Clontech) and transformed into AH109 using lithium acetate.

**Molecular cloning of cDNA for C-HAP2B, C-HAP5A, and C-HAP5B**

A partial C-HAP2B fragment was obtained by 5′-RACE (rapid amplification of cDNA ends) using the primers 5′-TWGTTGSAATGCDAAARCTAC-3′ and 5′-ACTACATAGGGCTCGAG-GGC-3′, and cDNA generated from carrot SES using the Marathon cDNA Amplification Kit (Clontech) as a template. Partial fragments of C-HAP5A and C-HAP5B were obtained from screened clones by digestion of the plasmids with EcoRI and XhoI. Using these fragments as probes, ~3.0×10⁶ phages from the carrot cDNA library were screened by plaque hybridization. The plasmids from positive plaques were rescued into Escherichia coli XLOLR, and the isolated plasmids were sequenced. Primary sequence comparisons were performed using Clustal X (Genetics Computer Group).

**RNA isolation and RT-PCR analysis**

Carrot cDNA samples were prepared from ECs, NCs, SEs, seedlings, mature leaves, and flowers as described previously (Yazawa et al., 2004). PCR was performed with 30 or 40 cycles using the following sets of gene-specific primers (shown 5′ → 3′): GGAATTCATGGATCAGTGTGAGGAATC-HAP2B TAACC and CCTCGAGATGATGCACCCAAAACCTAGATC for C-HAP5B, and AGACAAGGAGCTGTACGAGATCGAATCTCCATTTTCCTCCTGATC for To construct the expression vectors, the coding regions of the desired cDNA fragments had been amplified. PCR products were subcloned and sequenced to confirm that the TAA, for C-HAP2A AAACCGATTCAGACTTAGGA and CAAGTTCAGTGGTAGCA-3′.

**Table 1. Double-stranded oligonucleotides used**

<table>
<thead>
<tr>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>UAS2UP1 5′-GGCGAGCGTTGATGGTGATGCAAG-3′</td>
</tr>
<tr>
<td>mUAS2UP1 5′-GGCGAGCGTTGATGGTGATGCAAG-3′</td>
</tr>
<tr>
<td>Ez 5′-GGATTTTCTGATGGTCATAAAAGT-3′</td>
</tr>
<tr>
<td>mEz 5′-GGATTTTCTGATGGTCATAAAAGT-3′</td>
</tr>
</tbody>
</table>

Bold letters indicate the CCAAT/CATCA sequence.

**DNA purification and end-labeling**

The double-stranded DNA was purified and end-labelled using [γ-32P]dCTP and a Megaprime DNA Labeling system (GE Healthcare-Bio-Sciences).

**Binding reaction mixtures containing 20 ng of purified recombinant protein were incubated with the radioactive probe at room temperature for 20 min.** For EMSAs, the protein extract was incubated with unlabelled UAS2UP1 or mUAS2UP1, respectively, prior to the addition of the radioactive probe. The protein–DNA complexes were resolved by 5% PAGE in 0.5× TBE at 4 °C. After electrophoresis, the gels were dried and visualized by autoradiography using a Phosphor-Imager (FUJIX BAS5000).

**Immunoprecipitation assay**

Haemagglutinin (HA)-tagged proteins were prepared using a TNT-T7 Quick Coupled Transcription/Translation system (Promega) from cDNA cloned into pGADT7. A 5 µl aliquot of the reaction mixture and 250 ng of purified GST fusion protein were mixed and incubated at room temperature for 20 min. For EMSAs, the protein extract was incubated with unlabelled UAS2UP1 or mUAS2UP1, respectively, prior to the addition of the radioactive probe. The protein–DNA complexes were resolved by 5% PAGE in 0.5× TBE at 4 °C. After electrophoresis, the gels were dried and visualized by autoradiography using a Phosphor-Imager (FUJIX BAS5000).

**Results**

**Cloning and sequencing of HAP5 homologues from carrot**

A yeast two-hybrid system was used to isolate the counterparts of C-LEC1 protein. First a bait protein comprising full-length C-LEC1 fused with the GAL4 DNA-binding domain (BD) was created, but this bait protein activated the reporter genes in all transformants. To avoid this autonomous activation, a bait protein consisting of the GAL4 DNA-BD fused to amino acids 1–137 of C-LEC1, a segment that contains the putative subunit interaction and DNA-binding domains, was used instead. The prey proteins were produced by fusing the GAL4 activation domain (AD) to a carrot cDNA library constructed from somatic...
embryo mRNA. The mRNA was taken 7 d after induction because the amount of C-LEC1 transcripts increases primarily during this period (Yazawa et al., 2004).

Yeast strain AH109 containing lacZ and the GAL4-HIS and GAL4-ADE reporter genes was co-transformed with the bait and prey plasmids (pGBK-C-LEC11–137 and pGAD-carrot cDNA, respectively). Screening of 3 × 10⁶ independent transformants yielded six blue clones with HIS²ADE² phenotypes. The plasmids that were recovered from these positive clones were classified into three groups based on their sequences. One clone contained a segment that had 53.9% identity to a portion of the yeast HAP5 gene and was designated C-HAP5A. Three clones contained a segment that had 56.6% identity to a portion of the yeast HAP5 gene and was designated C-HAP5B. Two clones showed 56.2% identity to a portion of the Arabidopsis ethylene-insensitive 3-like 1 (EIL1) gene (Chao et al., 1997), but exhibited no homology to the yeast HAP5 gene.

To isolate the full-length C-HAP5A and C-HAP5B genes, a carrot somatic embryo cDNA library was screened by plaque hybridization using the partial C-HAP5A and C-HAP5B fragments as probes. The resulting 926 nucleotide C-HAP5A cDNA contained an open reading frame (ORF) encoding a 229 amino acid polypeptide with a predicted mol. wt. of 26.2 kDa. The 1167 nucleotide C-HAP5B cDNA contained an ORF encoding a 249 amino acid polypeptide with a predicted mol. wt. of 28.2 kDa. The predicted C-HAP5A and C-HAP5B amino acid sequences were 69.3% identical and each contained an ~80 residue fragment with strong similarity to the conserved domain of HAP5 subunits in other organisms (Fig. 1A). Of these 80 residues, 45 were absolutely conserved among all HAP5 subunits. The N- and C-terminal regions of the HAP5 subunits differed substantially in their lengths and amino acid sequences, but were relatively rich in P and Q. A P- and Q-rich region was also described for the Arabidopsis

Fig. 1. Alignment of the HAP5 and HAP2 sequences from different organisms. (A) Alignment of HAP5 sequences. Amino acid sequences were aligned using CLUSTAL X. The sequences and accession numbers included in the alignment are: AtNF-YC1 (Y13726), AtNF-YC2 (Y13725), ScHAP5 (U19932), and RtCBF-C (U17607). The bars indicate the positions of α-helices in the histone folds; L1 and L2 are the loop regions. (B) Alignment of HAP2 sequences. The C-HAP2A sequence is clearly incomplete because full-length C-HAP2A was not obtained from carrot somatic embryos. The sequences and accession numbers included in the alignment are: AtNF-YA1 (Y13720), AtNF-YA2 (Y13721), ScHAP2 (M15243), and RtCBF-B (M34238). Conserved amino acid residues are shown as black boxes. Dashes indicate gaps in the sequence. The subunit interaction (NF-YA1) and DNA-binding (NF-YA2) domains are underlined. Arrowheads indicate functionally critical amino acids. At, Arabidopsis thaliana; Sc, Saccharomyces cerevisiae; Rt, Rattus norvegicus.
HAP5 proteins (AtNF-YC1-9) and may function in transcriptional activation (Gusmaroli et al., 2001, 2002).

Cloning and sequencing of HAP2 homologues from carrot

HAP2 homologues in various organisms contain a highly conserved, 64 residue segment with DNA-binding and subunit interaction domains (Olsen et al., 1990; Xing et al., 1993). Seven well-conserved amino acids from this region were selected and the corresponding degenerate primer was synthesized. Using this degenerate primer for 5'-RACE of carrot somatic embryo cDNA, two carrot HAP2 cDNA fragments of approximately 600 bp were obtained. These fragments were >74% identical to the HAP2 conserved region and were designated C-HAP2A and C-HAP2B.

To isolate full-length C-HAP2A and C-HAP2B, a carrot somatic embryo cDNA library was screened by plaque hybridization using the C-HAP2A and C-HAP2B fragments as probes. No positive plaques containing the C-HAP2A fragment were obtained from this library, but six positive plaques containing C-HAP2B were obtained. The 1434 nucleotide C-HAP2B cDNA contained an ORF encoding a 303 amino acid polypeptide with a calculated mol. wt of 32.6 kDa.

Amino acid residues 162–213 of C-HAP2B comprise a segment with 63% identity to the HAP2 regions that are conserved from yeast to vertebrates. This domain is closely followed by an 18 amino acid segment (NF-YA1 domain) that is required for association with the HAP3/5 heterodimer, a short linker region, and a 21 amino acid segment (NF-YA2 domain) that is required for binding CCAAT sequences (Fig. 1B) (Olsen et al., 1990; Xing et al., 1993; Romier et al., 2003). An arginine in the NF-YA1 domain and three histidine, three arginine, one tyrosine, one phenylalanine, and one alanine in the NF-YA2 domain are key residues for yeast HAP2 function and are completely conserved in C-HAP2B and Arabidopsis HAP2 (AtNF-YA1–10) (Fig. 1B).

To determine whether C-HAP2B functions as a HAP2 subunit, a complementation test was performed using a yeast hap2 mutant. The mutant, which is unable to grow on non-fermentable carbon sources such as lactate (Pinkham and Guarente, 1985), was transformed with recombinant pNV11 containing yeast HAP2, AtNF-YA1, and C-HAP2B. The growth of the yeast hap2 mutant on lactate was rescued by each gene (see Supplementary Fig. S1 at JXB online).

Expression of C-HAP genes in carrot

C-LEC1 was specifically expressed in embryonic tissues, including ECs, SEs, and seeds containing developing embryos. In developing SEs and developing seeds, C-LEC1 expression increased gradually, reaching a maximum at the torpedo stage. It then decreased gradually as the embryos matured (Fig. 2; Yazawa et al., 2004). C-HAP5A expression was strongly detected in ECs, SEs, and developing seeds, and weakly detected in seedlings and leaves (Fig. 2). C-HAP5A expression was higher in the reproductive tissues than in the vegetative tissues. Although C-HAP5B was also detected in all tissues, its expression was weak in developing seeds (40 cycles) compared with other tissues (30 cycles). C-HAP2A was strongly detected in developing seeds (30 cycles) and weakly detected in all other tissues (40 cycles). In contrast, C-HAP2B was detected in all tissues at essentially the same level.

Fig. 2. Expression analysis of C-HAP genes in carrot. The expression of HAP family genes was examined in various carrot tissues and in carrot somatic embryos at various developmental stages. RNA extracted from the indicated materials was analysed by RT-PCR. The number of PCR cycles is indicated to the right of the panel. The analysed tissues were: 14-d-old seedlings (seedlings); leaves; flowers; non-embryogenic cells (NCs); embryogenic cells 38–63 μm in diameter (ECs); 1–14-d-old somatic embryos after induction (SEs); and developing seeds 5–62 d after flowering. DsActin was amplified as an internal control.
In vitro reconstitution of CCAAT-binding complexes

To determine whether C-HAP2B, C-LEC1, C-HAP5A, and C-HAP5B are able to form CCAAT-box-binding complexes, EMSAs were performed using purified recombinant GST fusions of these proteins. In plants, the genes that are under direct regulation by LEC1 and the cis-element that LEC1 acts on have not yet been identified. Thus, UAS2UP1, an oligonucleotide that was designed based on a yeast CYC1 promoter fragment, was used as a probe (Table 1); this fragment binds to yeast and mouse CBF complexes (Forsburg et al., 1989; Xing et al., 1993). As expected, the individual fusion proteins were unable to bind to the DNA fragments (Fig. 3A, B; lanes 3–5). In addition, no DNA binding occurred for any heterodimeric combination of the three recombinant proteins (Fig. 3A, B; lanes 6–8). CCAAT-binding activity was detected only when GST–C-HAP2B, GST–C-LEC1, and either GST–C-HAP5A or GST–C-HAP5B [hereafter, GST–C-HAP5(A/B)] were present in the DNA binding reactions (Fig. 3A, B; lane 9).

Protein–DNA complexes were not detected when a mutated UAS2UP1 (mUAS2UP1) probe containing a CCAAT→CATCA substitution was used (Fig. 3A, B; lane 10; Table 1). To confirm the specificity of the CCAAT-binding activity, competition experiments were performed using unlabelled UAS2UP1 and mUAS2UP1 oligonucleotides. UAS2UP1 successfully competed for binding to the labelled DNA (Fig. 3C, D; lanes 2–5), whereas mUAS2UP1 did not (Fig. 3C, D; lanes 6 and 7).

Fig. 3. In vitro binding of recombinant GST–C-HAP2B, GST–C-LEC1, and GST–C-HAP5A or GST–C-HAP5B to the CCAAT-box of the oligonucleotide UAS2UP1. Electrophoretic mobility shift assays (EMSAs) were performed with DNA-binding reactions containing the radiolabelled oligonucleotide UAS2UP1 incubated with purified recombinant GST, GST–C-HAP2B, GST–C-LEC1, and (A) GST–C-HAP5A or (B) GST–C-HAP5B in the indicated combinations (lanes 1–9). Lane 10, radiolabelled oligonucleotide mUAS2UP1 containing mutations in the CCAAT motif (CCAAT→CATCA) was used. Arrows indicate the position of specifically bound C-HAP2B–C-LEC1–C-HAP5(A/B) complexes. The free probe is located at the bottom of the gel. (C, D) Competition assays with increasing amounts of unlabelled UAS2UP1 (lanes 2–5) and mUAS2UP1 (lanes 6 and 7). As a control, the binding reaction in lane 1 contains unbound probe.
To establish the binding specificity further, an Ea oligonucleotide that was designed from a fragment of the mouse major histocompatibility complex (MHC) class II Eα promoter was also used as a probe (Table 1). This fragment binds to the mouse CBF complex (Sinha et al., 1995; Liang et al., 1998). The C-HAP2B/C-LEC1/C-HAP5(A/B) complex was able to bind to the Eα fragment, but not to the mutated Eα fragment (mEα; CCAAT→CATCA), indicating that the C-HAP2B/C-LEC1/C-HAP5(A/B) complex specifically recognizes the sequence CCAAT (data not shown).

Interactions among the four proteins

The interactions of the three CBF subunits with each other were next examined (Fig. 4) using immunoprecipitation assays to detect specific interactions between HA- and GST-tagged C-HAP proteins. Each HA-C-HAP protein was incubated separately with each of the GST–C-HAP pro-

![A](https://academic.oup.com/jxb/article-abstract/58/13/3819/497537)

![B](https://academic.oup.com/jxb/article-abstract/58/13/3819/497537)

![C](https://academic.oup.com/jxb/article-abstract/58/13/3819/497537)

![D](https://academic.oup.com/jxb/article-abstract/58/13/3819/497537)

Fig. 4. Immunoprecipitation analysis of interactions between C-HAP2B, C-LEC1, C-HAP5A, and C-HAP5B. (A) In vitro translated HA-C-LEC1 (5 μl) was incubated in separate reaction mixtures with 250 ng of each GST–C-HAP protein or 250 ng of GST as indicated. Similar experiments were performed using each HA-C-HAP protein, respectively (B, C, D). (A–D) Lane 1, in vitro translated HA-C-HAP protein (5 μl) was loaded as a positive control.

Discussion

It was shown that C-LEC1 interacts with C-HAP2B and C-HAP5(A/B) to form a complex capable of binding the sequence CCAAT in vitro. A speculative model was then developed for the formation of the C-HAP2B/C-LEC1/C-HAP5(A/B) complex (Fig. 5). In this model, C-LEC1 and C-HAP5(A/B) associate to form a tight heterodimer, which creates a surface for interaction with C-HAP2B. The resulting heterotrimer then binds to DNA containing CCAAT fragments. The reaction mixtures were then incubated with glutathione–agarose resin to precipitate the GST–C-HAP proteins and their complexes. The precipitates were analysed using SDS–PAGE and western blotting with anti-HA antibodies.

In this assay, HA-C-LEC1 was precipitated by C-HAP5 (A/B) (Fig. 4A; lanes 5 and 6), but not by GST alone, GST–C-HAP2B, or GST–C-LEC1 (Fig. 4A; lanes 2–4), and HA-C-HAP5A was precipitated by GST–C-LEC1 and GST–C-HAP5(A/B) (Fig. 4C; lanes 4–6). Similarly, HA-C-HAP5B was precipitated by GST–C-LEC1 and GST–C-HAP5(A/B) (Fig. 4D; lanes 4–6). Although HA-C-HAP5A and HA-C-HAP5B exhibited weak interactions with GST–C-HAP2B (Fig. 4C, D; lane 3), HA-C-HAP2B was not precipitated by C-HAP5(A/B) or GST–C-LEC1 (Fig. 4B; lanes 4–6). It was not possible to confirm whether C-HAP2 interacts with C-LEC1 and C-HAP5(A/B). GST–C-LEC1 and GST–C-HAP2B did not form homodimers, but both GST–C-HAP5A and GST–C-HAP5B could form homodimers and they could associate to form a heterodimer. These results suggest a model for the formation of C-HAP2B/C-LEC1/C-HAP5(A/B) complexes (Fig. 5).
the sequence CCAAT. Many of the results that were expected based on previous studies in yeast and mice were obtained, but there were a number of unexpected findings. First, both C-HAP5A and C-HAP5B can form homodimers and can associate to form a heterodimer in vitro (Fig. 4C, D). In previous studies of vertebrates and yeast, none of the HAP subunits was able to form a homodimer. C-HAP5(A/B) homodimers appeared to have formed, but no DNA binding was detected (Fig. 3A, B; lane 5), indicating that C-HAP5(A/B) homodimers are not able to bind to the DNA fragment. Whether these dimers are formed in vivo and have plant-specific functions remains to be determined. Secondly, although the vertebrate and yeast HAP2 subunits did not interact with HAP3 or HAP5, C-HAP2B may interact directly with C-HAP5(A/B) (Fig. 4C, D; lane 3). The fact that C-HAP2B and AtNF-YA1 could complement the yeast hap2 mutant (see Supplementary Fig. S1 at JXB online; Edwards et al., 1998) suggests that the functions of some plant HAP2 subunits are similar to those of yeast HAP2 and that these proteins can form functional complexes with HAP3/HAP5 in vivo. Still, the mechanism of binding between plant HAP2 and HAP3 or HAP5 may be slightly different from that in vertebrates and yeast.

The process by which the target promoters of the C-HAP2B/C-LEC1/C-HAP5(A/B) complexes are recognized is of great interest. Although C-LEC1-containing complexes are thought to regulate the expression of embryogenesis-regulated genes that have CCAAT-boxes in their promoters, genes that are directly regulated by C-LEC1 or Arabidopsis LEC1 have yet to be identified. The present EMSA results using UAS2UP1 and Ea indicate that the sequences flanking the CCAAT-box are not important for the binding of C-HAP2B/C-LEC1/C-HAP5(A/B) (Fig. 3; Table 1; data not shown). Furthermore, the specific selection of target promoters for CBF cannot be directed solely by the CCAAT-box in the transcription activation process. However, interactions of C-LEC1-containing complexes with other factors may confer specific functionality on the complexes.

In various higher eukaryotic promoters, the functional CBF-binding sites are relatively close in location to the TATA motif (Bucher, 1990) and are invariably flanked by at least one functionally important promoter element. Several reports have shown that various factors, including transcription factors, co-activators, and TATA-binding proteins, interact with CBF or its subunits (Mantovani, 1999). To understand the function of CBF during embryogenesis, it is necessary to identify the genes that are directly regulated by LEC1/C-LEC1 and to analyse their cis-acting elements in detail.

The sequences located outside of the conserved domains vary considerably between C-HAP2A and C-HAP2B, and between C-HAP5A and C-HAP5B (Fig. 1A, B). In the Arabidopsis and rice HAP subunits, including HAP3, these regions are rather variable in composition and length (Gusmaroli et al., 2001, 2002; Miyoshi et al., 2003). These facts suggest that C-HAP2B–C-LEC1–C-HAP5(A/B) complexes may interact both with a variety of flanking sequence-specific DNA-binding transcription factors and with some of the proteins involved in the formation of pre-initiation transcription complexes. The carrot EIL1-like protein, which has been isolated by yeast two-hybrid screening and is homologous to AtEIL1, may be one of these proteins. Whether this protein can bind to C-HAP2B/C-LEC1/C-HAP5(A/B) complexes remains to be clarified.

If the variant region in each subunit is related to the recognition of a specific promoter, it is important that C-LEC1 forms a complex with the necessary subunit at the appropriate place and time. Although C-HAP2(A/B) and C-HAP5(A/B) were expressed nearly ubiquitously and at essentially the same level in all of the tissues and organs tested (Fig. 2), it is not clear whether they are expressed in the peripheral region of the embryo where C-LEC1 is specifically expressed (Yazawa et al., 2004). In Arabidopsis, 29 HAP genes have been identified and their expression patterns have been surveyed (Gusmaroli et al., 2001, 2002). Twenty-three of these genes are expressed in siliques, but little is known about their localization. LEC1-type HAP3 genes, LEC1, and LIL were mainly expressed in the outer cell layers of the embryo, similar to C-LEC1 (Lotan et al., 1998; Kwong et al., 2003). In Linum usitatissimum (flax), a HAP3 transcript (Lu11146) was detected only in the procambium of the embryo axis, in the root cap, and in the upper three layers of the apical meristem (Gutierrez et al., 2006). To identify the combinations of C-HAP proteins that form a complex, it is important to investigate the localization of each protein in detail.

Additional homologues of C-HAP2 and C-HAP5 may be expressed during embryogenesis, but it is thought that C-HAP2B, C-HAP5A, and C-HAP5B are the most likely of the C-HAP factors to form complexes with C-LEC1. It is planned to test whether these C-HAP complexes are formed in vivo during embryogenesis. Although all four C-HAP proteins were isolated from somatic embryos, it is unknown whether preferential binding occurs between different subunits. The plant HAP system might be more intricate than that in vertebrates and yeast, which have only one gene that encodes each HAP subunit. The genes that are directly regulated by C-LEC1 also remain to be identified. Analyses of the promoter regions of these genes will be extremely important in understanding the functions of C-HAP2/C-LEC1/C-HAP5 complexes.

**Supplementary material**

Supplementary material can be found at JXB online.
Fig. S1. Complementation of the yeast hap2 mutant with AtNF-YA1 or C-HAP2B cDNA.

Acknowledgements

This work was supported in part by a Grant-in-Aid from the Ministry of Education, Science, Culture and Sports, Japan (No. 16370017), and a Grant-in-Aid for special Research on Priority Areas (No. 19043007) to HK. We thank the staff of the Ibaraki Experimental Field of the Takii Seed Co., Ltd, for collecting carrot materials.

References


