2168–2177 Nucleic Acids Research, 1995, Vol. 23, No. 12

ZEMa, a member of a novel group of MADS box genes, is alternatively spliced in maize endosperm

Katharina Montag*, Francesco Salamini and Richard D. Thompson

Max-Planck-Institut für Züchtungsforschung, Carl-von-Linné Weg 10, 50829 Köln, Germany

Received March 6, 1995; Revised and Accepted May 5, 1995

EMBL accession nos X85334–X85336 (incl.)

ABSTRACT

The identification of a number of *cis*-elements which direct gene expression in maize endosperm, and the characterization of corresponding DNA binding proteins, point to the interaction of different classes of transcription factors in this tissue. To assess whether MADS box genes are also involved in maize endosperm development, cDNA and genomic MADS box clones have been isolated. The three cDNA clones ZEM1, ZEM2 and ZEM3 were cloned from a maize endosperm cDNA library using a probe based on sequences conserved in plant MADS box genes. Further transcripts were cloned by RT–PCR experiments and designated ZEM4 and ZEM5. Analysis of the corresponding genomic clones led to the identification of the ZEM2 MADS box gene family, three members of which were characterized sharing 97% sequence identity in corresponding domains. 100% sequence identities between cDNA and one of the genomic clones, conserved exon–intron boundaries and the demonstration of *in vivo* splicing in a maize endosperm transient expression system, show that the transcripts ZEM1–5 are derived by alternative splicing of *ZEMa*, one ZEM2 member. The *ZEMa* transcripts are present in almost all maize tissues, but specific differentially spliced forms accumulate preferentially in maturing endosperm and leaf. The function of the *ZEMa* gene is discussed in the light of similarities in the expression pattern with members of the human MEF2/RSRF gene family.

INTRODUCTION

Maize endosperm development has been extensively studied as a model system to investigate plant developmental processes. In this tissue complex genetic controls determine the final morphology of the grain (1–3). Moreover, *trans*-acting regulatory proteins have been identified in seed-specific developmental pathways leading to storage protein synthesis (4,5), seed pigmentation (6,7) and seed dormancy (8). All these are processes taking place during the last maturation phase of endosperm development from 12 days after pollination (dap) to maturity. Based on their sequence and interaction with target promoters, different transcription factor families have been shown to be involved including basic region/leucine-zippers (*OPAQUE-2*), helix–loop–helix proteins (*R*) and *Myb*-like factors (*C1*).

The MADS box genes represent a further family of regulatory factors active in plants. Members of this gene family share a highly conserved motif, designated the MADS box (9), a sequence-specific DNA-binding and dimerization domain (10–12). In plants, the MADS proteins best described to date are expressed during floral organogenesis, where they play a key role in regulating organ identity (reviewed in 13). However, a number of plant MADS box genes are also expressed in vegetative tissues such as leaf (14–16), root (17) and shoot apical meristem (17,18), indicating that plant MADS box gene function is not restricted to floral development.

The functions of MADS box-homologous genes so far identified in yeast and animals are quite different. The differentiation of mating type in MATa and MATα yeast cells is a combined function of the product of the MCM1 protein interacting with the gene products of a- or α-specific genes (19). In mammals, the regulation of cell growth is controlled by *SRF* (10), while genes of the MEF2/RSRF family are involved in muscle development in man (20–24). However, most MADS box-homologous genes characterized to date share an involvement in the control of cell differentiation and cell-type specificity. This regulation is effected through binding of the encoded proteins as homo- or heterodimers to target sequences, which conform to the consensus CArG (20,25).

The presence of canonical CArG-boxes in the promoters of endosperm-specific genes, like *O2* and *C1*, suggested a possible role of MADS box transcription factors in the regulation of endosperm-specific processes. An RT–PCR MADS box probe was therefore used to isolate a series of cDNA clones from a 10 dap endosperm library, termed ZEM (*Zea* endosperm MADS box gene). Analysis of the corresponding genomic clones indicated that this group of MADS box genes encodes a series of transcripts via alternative splicing.

MATERIALS AND METHODS

Polymerase chain reaction

Single strand cDNA was synthesized from 8 μg poly(A)$^+$ RNA isolated from 12 dap maize seed (var. Lorena) using the primer: 5′-GAATTCGGATCCAAGC(T)$_{20}$-3′. PCR was performed using the degenerate second primers: 5′-A(G/C)ATCAA(G/A)-(C/A)GIAT(A/T/C)GA(G/A)AAT-3′ (T-nucleotide end primer;

* To whom correspondence should be addressed

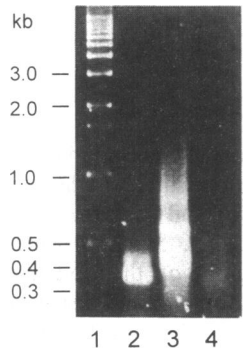

Figure 1. RT–PCR with degenerate MADS box primers. First strand cDNA was reverse-transcribed from poly(A)$^+$ RNA extracted from 12 dap seed. For the PCR, two degenerate 5′ primers, the T-nucleotide end primers (lane 2) and the C-nucleotide end primers (lane 3), were used. In lane 4 a sample incubated without primer and in lane 1 a DNA size marker were loaded. The probe used for screening cDNA and genomic libraries (pB11 clone, Fig. 3A) was a subclone of the PCR products generated with the C-nucleotide end primers (lane 3).

see Fig. 1, lane 2) and 5′-A(G/C)ATCAA(G/A)(C/A)GIAT(A/-T/C)GA(G/A)AAC-3′ (C-nucleotide end primer; see Fig 1, lane 3) based on the conserved amino acid sequence IKRIEN of the DEFA and AG MADS box. The amplification reaction was carried out in 80 µl *Taq*-buffer (Amersham) supplemented with 25 pmol 5′ and 3′ primer, 400 µM of each deoxynucleotide triphosphate, 1 mM $MgCl_2$ and 1 U *Taq* DNA polymerase (Amersham). After 25 cycles with an annealing temperature of 40°C, fresh *Taq* polymerase was added for a second amplification round using the same cycling program. The gel-purified amplification products were subcloned blunt-end into the *Sma*I restriction site of the vector pGEM3Zf(+) (Promega) for sequencing (pB clones).

For the RT–PCR analysis using sequence-specific primers, cDNA from 10 ng poly(A)$^+$ RNA template was amplified for three cycles with an annealing temperature of 50°C followed by 32 cycles with annealing at 65°C. The sequences of the ZEM domain-specific oligonucleotides are:

5′-TGGAAGATCTAGGGGCAAGATCGAGATC-3′ (MADS box 5′ primer),
5′-TGGAGATCTAGCCACCGACCTTGAAGT-3′ (ZEM1/3 domain 3′ primer) and
5′-ATCCCCCGGGGCAATTCTAGATCTATCTTG-3′ (ZEM1 domain 3′ primer).

The underlined sequences correspond to the cDNA sequences shown in Figure 2. The amplification products were digested at *Bgl*II restriction sites in the 5′ ends of the primers and cloned into the *Bam*HI site of pUC18.

Screening of cDNA and genomic libraries

A maize cDNA library made from poly(A)$^+$ RNA of wild-type (A69Y$^+$) 10 dap endosperm using a λZAPII cDNA cloning kit (Stratagene) was screened with a ^{32}P-labeled fragment of the PCR clone pB11 (probe a+b; Fig. 3). Plaques (~6 × 10^4) were screened by hybridization at 62°C in 5 × SSPE (1 × SSPE is 0.15 M NaCl, 10 mM sodium phosphate, 1 mM EDTA, pH 7.4), 5 × Denhardt's solution (1 × Denhardt's solution is 0.02% Ficoll, 0.02% PVP, 0.02% BSA), 0.5% SDS, 100 µg/ml salmon sperm DNA, for 20 h, followed by washes in 2 × SSPE at 55°C. The probe was pre-hybridized to a pUC-filter for 2.5 h at 68°C to reduce plasmid background signal.

The genomic library (kindly provided by A. Gierl, Institute of Applied Plant Molecular Biology, Technical University of München) consists of *Mbo*I-partial digest fragments of the maize line *wx-844::En* (26) cloned into the *Bam*HI site of λEMBL4. Plaques (6 × 10^5) were screened with the same probe and under the same conditions as for the cDNA library screening. Genomic inserts were analysed by DNA gel blot analysis of plaque-purified

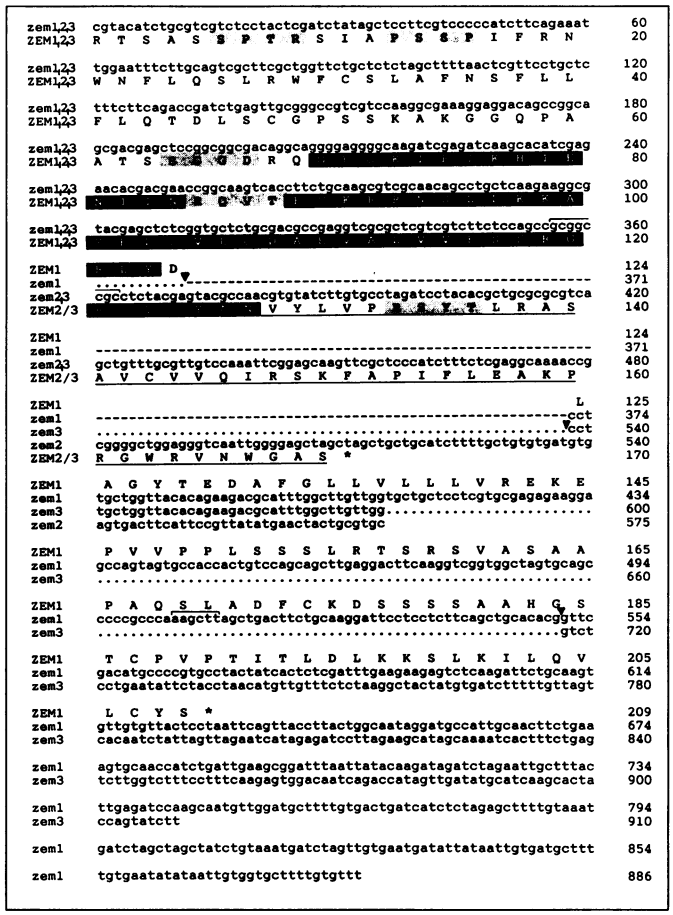

Figure 2. Nucleotide and deduced amino acid sequences of ZEM1, ZEM2 and ZEM3. Amino acid sequences are shown in capital and nucleotide sequences in small letters as indicated on the left side. The nucleotide (1–886) and deduced amino acid (1–209) sequences of ZEM1—which are identical to those of ZEM2 and ZEM3 up to nucleotide 371—are shown in the first two lines. After position 371, ZEM2 (nucleotide sequence 5–575) and ZEM3 (nucleotide sequence 8–910) continue being identical up to nucleotide 537 and contain the same open reading frame encoding for amino acids 3–170 (ZEM2/3). From nucleotide 538 up to 718 the nucleotide sequence of ZEM3 (middle) is identical to the corresponding ZEM1 nucleotide sequence (above). The dashes in nucleotide sequences indicate gaps and the dots identical nucleotides. The MADS box is shown in black and the ZEM2/3 domain, present in all isolated genomic ZEM clones (see Fig. 3B, light grey), is underlined. Putative phosphorylation sites (see Discussion) are grey shaded. The positions of introns are denoted by triangles in the corresponding cDNA sequences. The 3′ ends of the ZEM clones were defined by the presence of poly(A) tails in each case. The *Not*I and *Hin*dIII restriction sites used for constructing the GUS fusions (see Materials and Methods) are indicated by brackets above.

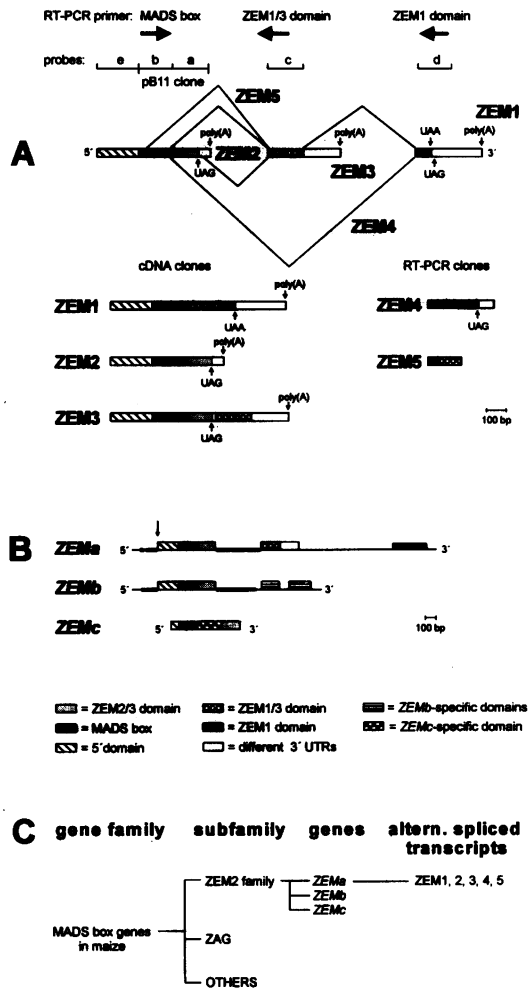

Figure 3. The ZEM2 gene family. (**A**) Scheme of alternative splicing events used to derive the transcripts ZEM1–5 from the *ZEMa* gene. The figure is based on data from several experiments as described in the text. The different splicing events are shown above (ZEM1, ZEM5) and below the splicing diagram (ZEM3, ZEM4). No splicing is necessary for ZEM2. 3′ untranslated regions (3′ UTRs) are marked by the stop codons and poly(A) tails found in the cDNA clones (open boxes) and are not related. The stop codon of ZEM2 and ZEM3 is the same (UAG) with the consequence that the ZEM1/3 domain (checkered) is untranslated in ZEM3. The open reading frames of ZEM1 and ZEM4 are different within the ZEM1 domain, which they share, as indicated by different stop codons, and those of ZEM1 and ZEM5 are the same within the ZEM1/3 domain. The positions corresponding to the PCR probe used for screening the maize libraries (a and b together) and to probes for RNA and DNA blot analysis (a–e) are shown, as well as those for the RT–PCR primers. The nucleotide sequences of all spliced *ZEMa* forms shown are 100% identical in corresponding domains, whereas those of the different *ZEM* genes (B) show 97% identity between. (**B**) Schematic representation of the genomic clones corresponding to *ZEMa*, *ZEMb* and *ZEMc*. The three genomic clones share the MADS box (black), the 5′ domain N-terminal of the MADS box (hatched) and a region C-terminal to the MADS box which is related to the ZEM2/3 domain of the cDNA clones. Exon regions, derived from comparisons with all isolated cDNA and PCR clones, are shown as boxes, and introns as lines. The *ZEMb*-specific domains (stripes) are related to the 3′ sequence of one of the pB clones (not shown) and are considered as two exons separated by a small intron. In *ZEMc* the MADS box and the ZEM2/3 domain are separated by a *ZEMc*-specific domain assumed to be an exon (circles), because of the continuous open reading frame throughout and of the missing intron consensus sequences at its 5′ and 3′ ends. Homologous regions are marked by corresponding patterns (exons) or by thickened lines (introns). The 5′ end of the cDNA clones ZEM1, ZEM2 and ZEM3 is marked by an arrow in *ZEMa*. (**C**) Arrangement of the ZEM2 subfamily within the MADS gene family of maize.

restricted phage DNA. MADS box containing DNA fragments were subcloned into the pUC18 vector for sequencing.

RNA and DNA gel blot analysis

Poly(A)$^+$ RNA was obtained from the whole detached seed, developing endosperm, pericarp at 14 dap, embryo at 20 dap, root of germinated seedling and leaf from 3-week-old plants as previously described (27). Samples were isolated from the developing seed by cutting off the top and one side of the seed coat, removing the endosperm and discarding the embryo. The remaining husks were taken as the pericarp sample. The unfertilized caryopsis was isolated 5 days after the time point of silk emergence from the husk leaf. Poly(A)$^+$ RNA samples (2–6 μg) were separated on 1.5% agarose gels containing 6% formaldehyde and 1 × MOPS buffer (20 mM 3-[*N*-morpholino] propanesulfonic acid, 5 mM sodium acetate, 1 mM EDTA, pH 7.0) and transferred onto Hybond-N membranes (Amersham). Filters were hybridized overnight at 42°C in 5 × SSPE, 50% deionized formamide, 5 × Denhardt's solution, 0.5% SDS and 200 μg/ml salmon sperm DNA, and washed in 2 × SSPE, 0.1% SDS at 42°C.

Genomic DNA was isolated from 4-week-old maize leaves (28). Samples (5 μg) were digested, fractionated by agarose gel elecrophoresis, and blotted onto N-Hybond membranes (Amersham). Hybridization and washing was performed at 64°C as for the library screenings.

DNA and RNA gel blots were hybridized with gel-purified, labeled PCR probes (a–e) that correspond to the different domains of the cDNA clones (see Fig. 3A). A parsley polyubiquitin probe (29) was used to control uniform quantities of RNA were loaded. Filters were exposed to Kodak XAR-5 film for 4–14 days.

GUS fusion constructs

For the GUS fusion constructs (Fig. 8) the *Not*I–*Hin*dIII fragment of the genomic clone of *ZEMa* (815 bp) as well as that of the cDNA clones ZEM1 (145 bp) and ZEM3 (310 bp) were subcloned into a *Not*I–*Hin*dIII digested M13mp18 vector mutagenesis cassette that carries the *Sac*I–*Hin*dIII polylinker fragment of pBluescript (Stratagene). *In vitro* mutagenesis was performed using the T7 Sculptor mutagenesis kit (Amersham). The suppression of the two stop codons UAG in the reading frame of the ZEM3 GUS fusions was obtained by point mutation of the adenines (see Fig. 2, nucleotides 512 and 653 of ZEM3) in both the genomic and cDNA fragment to give UUG (ZEM3SUP, ZEM3SUP^{AT} and cZEM3). The insertion of a thymidine one nucleotide before the *Hin*dIII restriction site of the genomic and ZEM1 fragment (see Fig. 2, between nucleotides 503 and 504 of ZEM1) results in the mutated sequence CCAT*AAGCTT present in ZEM1IN, ZEM1IN^{AT} and cZEM1. In addition, in the 5′ splice site mutant constructs ZEM1IN^{AT} and ZEM3SUP^{AT}, the 5′ splice site consensus GT in the intron was changed to AT. The *Nco*I–*Not*I fragment, comprising the N-terminal part of the ZEM clones fused to an artificial start codon was subcloned together with the mutated *Not*I–*Hin*dIII fragment into the *Nco*I restriction site of the 35S GUS vector pRT103GUS (30).

Particle bombardment

Maize wild-type ears (A69Y$^+$), harvested between 13 and 16 days post-pollination, were surface-sterilized for 20 min in 7% NaOCl

containing 0.1% SDS and rinsed in sterile water for 2 min. Radial endosperm cross-sections were dissected directly from the ear and seven discs were placed on a filter paper (Whatman No. 4, 2.5 cm diameter) in Petri dishes with solid medium. The medium (31,32) was supplemented with 0.6% agar, 3% sucrose, 8.2 g/l total amino acid mix and 0.5 g/l Claforan (Hoechst, Frankfurt, D) as antibiotic agent, adjusted to pH 5.8 with potassium hydroxide. Qiagen-purified plasmid DNA (Qiagen, Düsseldorf) (10 µg) was coated onto 3 mg gold particles of 1.6 µm diameter according to the manufacturer's instructions (Dupont, Wilmington, USA). Seven endosperm slices were bombarded using the PDS-1000/Helium gun (Biorad) and rupture discs of 1100 p.s.i. (Biorad). After incubation at room temperature for 2 days in the dark, the endosperm was stained by incubation overnight with the GUS substrate X-Glc (5-bromo 4-chloro 3-indolyl β-D-glucuronic acid) (Biomol) in 100 mM sodium phosphate pH 7, 0.1% Triton X-100.

RESULTS

Structure of the ZEM cDNA clones isolated from maize endosperm

An RT–PCR was performed on poly(A)$^+$ RNA isolated from maize seed using degenerate MADS box primers, based on sequences conserved between the plant genes *DEFICIENS* and *AGAMOUS* (Fig. 1). The two primers differ only in their 3-terminal nucleotide reflecting two possible codon usages for asparagine in the peptide IKRIEN that is conserved between the two known sequences. The largest of four cloned PCR products (pB clones) containing a MADS box sequence (pB11, Fig. 3) was chosen to screen a maize cDNA library made from 10 dap endosperm. Three cDNA clones, designated ZEM1, ZEM2 and ZEM3 were isolated. As shown in Figure 2, each contains an open reading frame encoding a putative protein with a MADS box domain. The open reading frames of ZEM2 and ZEM3 are the same. Sequence comparison at the amino acid level showed this conserved region was most homologous to the AGAMOUS and AGL1/AGL5 MADS domains of *Arabidopsis* (14,33) containing only four conservative amino acid substitutions through a stretch of 56 amino acids. At the nucleotide level it most resembled the *ZAG1* MADS box of maize (34) which is 92% identical in sequence. Outside the MADS box, no other homologies to DNA sequences stored in the GenBank or EMBL data bases were found. An N-terminal domain upstream of the MADS box was present in the three ZEM encoded proteins, as has also been reported for AGAMOUS. No longer cDNA containing a putative ATG translation start codon has so far been identified. In contrast to most plant MADS box genes, the ZEM clones do not contain a K-box. This is a leucine-rich conserved domain of the coiled-coil keratin type positioned ~30–50 amino acids C-terminal of the MADS box, probably mediating protein–protein interaction (14,15). All the ZEM clones possess, instead, a very short C-terminal region, comprising 86 amino acids for ZEM1 and 43 amino acids for ZEM2 and ZEM3.

The ZEM2 gene family

The ZEM cDNA clones ZEM1–3 have identical nucleotide sequences in their shared domains (Fig. 2). It was therefore hypothesized that all three ZEM clones might have arisen from the same gene by alternative splicing. This hypothesis was

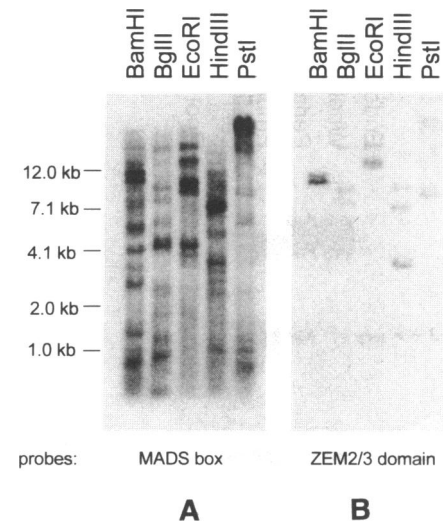

Figure 4. Existence of a subfamily of MADS box genes in maize. Genomic DNA of the wild-type maize line A69Y$^+$ were digested with different restriction enzymes, indicated in each lane, and hybridized with the MADS box probe (**A**) or the ZEM2/3 domain probe (**B**) under moderately stringent conditions (64°C, 5 × SSPE).

subsequently supported by their extensive nucleotide sequence identity with the genomic clone of *ZEMa*, the presence of conserved exon–intron boundaries for each cDNA clone in the *ZEMa* sequence (Fig. 3B) and transient expression studies (see below). In Figure 3A, the type and position of splicing events needed to generate the transcripts ZEM1–5 from *ZEMa* are indicated. (The origin of ZEM4 and ZEM5 will be described later.)

The hybridization of a maize genomic library (kindly provided by A. Gierl, IAPMB, München) under moderately stringent conditions with the same probe as that used for screening the cDNA bank, comprising the MADS box [probe (b), Fig. 3A] and the DNA sequence present in both ZEM2 and ZEM3 clones [ZEM2/3 domain probe (a), Fig. 3A] resulted in the isolation of seven MADS box-containing genomic clones. Further analysis by PCR and sequencing showed that they represented at least three related genes, *ZEMa*, *ZEMb* and *ZEMc* (Fig. 3B), all containing the ZEM2/3 domain C-terminal and the 5′ domain N-terminal to the MADS box. The nucleotide sequences of *ZEMa*, *ZEMb* and *ZEMc* are 97% identical throughout corresponding exons and throughout those introns marked by a thickened line in Figure 3B. *ZEMa* contains all the domains present in the three cDNA clones and its sequence shows 100% identity in each exon.

Southern blot analysis of genomic maize DNA with the MADS box probe (b) revealed a minimum of 13 strong hybridizing bands (Fig. 4A), indicating the existence of a large number of genes in maize containing the MADS box sequence. The same filter was hybridized with probe (a) corresponding to the ZEM2/3 domain. Hybridizations with this probe still resulted in four to five bands (Fig. 4B), suggesting the existence of a family of MADS box genes characterized by the presence of the ZEM2/3 domain, designated here as the ZEM2 gene family.

Figure 3C shows the relationship of the ZEM2 subfamily within the MADS gene family and gives the designations used in the following text.

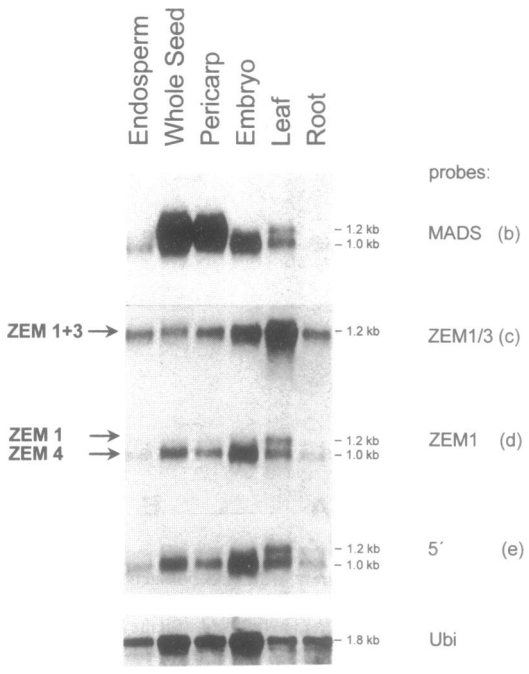

Figure 5. Analysis of *ZEM* expression in different maize tissues. Poly(A)$^+$ RNA from endosperm and whole seed harvested at 10 dap, pericarp, embryo, leaf and root were loaded on an RNA gel as indicated in each lane. The same RNA gel blot was successively hybridized, after removing the previous probe, with the different ZEM specific probes (b)–(e) (see Fig. 3A). A parsley polyubiquitin probe (29) was used as an internal control to compare the amount of poly(A)$^+$ RNA loaded. The lengths of the transcripts are estimated at 1.2 kb for ZEM1 and ZEM3 and 1.0 kb for ZEM4.

Expression of *ZEMa* encoded transcripts

Gel blots of poly(A)$^+$ RNA from a series of maize tissues were hybridized with different probes to investigate the sites of expression of the various *ZEMa* spliced forms (Fig. 5). Probes (a)–(e) correspond to the different ZEM domains and their positions are indicated in Figure 3A. Transcripts were found in all tissues analysed with the exception of pollen.

Hybridization with the ZEM1/3 probe (c) revealed the presence of one band in each tissue, a result consistent with the similar lengths of ZEM1 and ZEM3 cDNA clones (0.89 and 0.90 kb). The signal presumably represented the superimposed expression of these two transcripts. The highest level of expression was present in leaf.

Using the ZEM1 probe (d), the ZEM1 message was strongly expressed in leaf and only very weakly in the endosperm (10 dap) and in other tissues. A further transcript, termed ZEM4, also hybridized to probe (d), but was expressed more uniformly in different tissues if compared with the ubiquitin control. The size of this mRNA, ~200 bp shorter than that of ZEM1, and its absence in the ZEM1/3 probe (c) hybridization, indicate that it might have arisen by splicing out ('skipping') of the second exon of ZEM1 (Fig. 3A). This origin of ZEM4 was verified by cloning and analysis of the corresponding PCR product obtained with a MADS box and a ZEM1 domain primer pair positioned as in Figure 3 (described in next section: RT–PCR).

The RNA blot hybridization using probe (e) (5′ end of ZEM1–3), showed the same expression pattern as that of the ZEM1 probe (d), strongly indicating that the RT–PCR clone ZEM4 also contains the 5 domain of the three cDNA clones ZEM1–3.

Besides the *ZEMa* gene, *ZEMb* and *ZEMc* also contain the 5′ domain, but not the ZEM1 domain sequence (see Fig. 3B). However, the similar strength and distribution of the signals in the two hybridizations (e) and (d) suggest that those in (e) are mostly due to expression of the *ZEMa* gene.

The hybridization (e), however, was different from the one obtained with the ZEM1/3 probe (c). As mentioned above, the signal seen with the ZEM1/3 probe covers both ZEM1 and ZEM3 mRNAs. It should represent mostly ZEM3 expression, if the ZEM1 contribution to this signal is very small, as is evident from the ZEM1 probe hybridization. Because the cDNA sequence of ZEM3 also contains the 5′ domain, all strong signals ascribed to ZEM3 in panel (c) should be visible in the corresponding RNA blot revealed by the 5′ end probe (e). Apparently this is not the case, indicating the existence of a further 1.2 kb transcript, which must contain the ZEM1/3 domain, but not the 5′- and the MADS box-related domains.

Hybridization of the RNA blot with the MADS box probe (b) gives similar signals to those observed for the ZEM1 and the 5′ domain probes, with the addition of strong hybridization in the whole seed and pericarp samples. This is attributable to cross-hybridization to the maize flower-specific *ZAG* transcripts (34) as the MADS box sequences of *ZAG1* and *ZEMa* are 92% identical.

To determine whether specific *ZEM* mRNAs were expressed during maize endosperm development, RNA gel blots of poly(A)$^+$ RNA from endosperm and whole seed harvested between 5 and 20 dap were hybridized with probe (d) (Fig. 6). This detected ZEM1 mRNA in endosperm from 14 dap which was elevated in 20 dap endosperm, taking into account the different amounts of RNA loaded. The signal seen in the whole seed sample could be accounted for entirely by the endosperm component. Although ZEM1 signal was not visible in 10 dap endosperm in the RNA blot analysis, its isolation by RT–PCR, and from an endosperm cDNA from this stage showed a basal expression there. The level of ZEM1 mRNA was nevertheless similar to that of ZEM4 transcript at 14 and 20 dap. The strong ZEM4 signal in the young seed (5 dap) may be derived from the nucellar tissue, as high ZEM4 expression was also observed in the unfertilized caryopsis (data not shown). Its level in the whole seed then decreases during development, especially between 5 and 12 dap. In this differentiation phase of endosperm development the nucellus is gradually replaced by the very rapidly growing endosperm body and the remaining nucellar cells are compressed to the outer edge of the kernel cavity (35). Thus, the decrease of the ZEM4 signal might be due to dilution of maternal tissue in the developing whole seed. In addition, however, a lower level of ZEM4 transcript was also present in the endosperm itself, as seen in the 14 and 20 dap samples.

Transcript distribution by RT–PCR analysis

The characterization of the genomic clones corresponding to *ZEMa*, *ZEMb* and *ZEMc* indicates that a small family of closely related genes exists in maize. Therefore, in an RNA blot hybridization the possibility of co-hybridization of their transcripts arises. To demonstrate that *ZEMa*-derived transcripts of sizes predicted from the RNA blot analysis (Figs 5 and 6) were

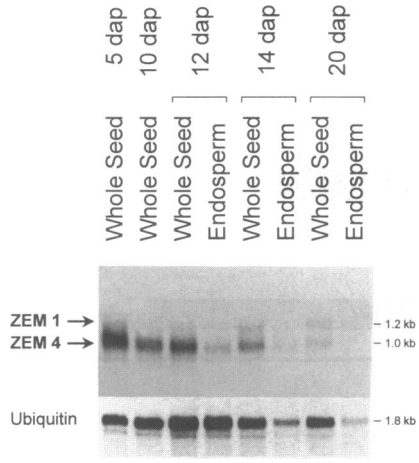

Figure 6. Analysis of ZEM1 and ZEM4 expression in developing endosperm. Poly(A)$^+$ RNA from endosperm and whole seed harvested between 5 and 20 dap were loaded. The blot was hybridized with the ZEM1 specific probe (d) and with the ubiquitin probe.

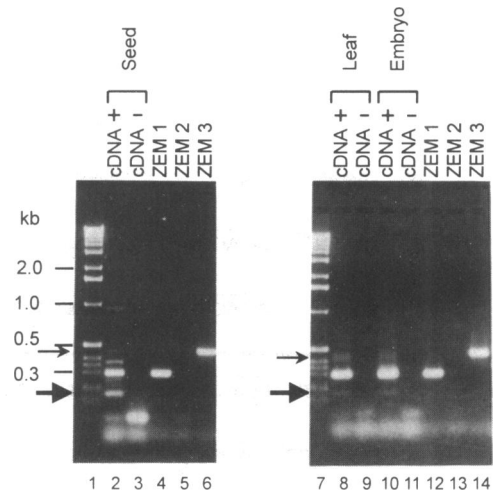

Figure 7. RT–PCR experiments. Using an oligo(dT)-primer, first strand cDNAs were reverse-transcribed with M-MuLV reverse transcriptase from poly(A)$^+$ RNA isolated from seed (10 dap), leaf and embryo. Amplified cDNAs (cDNA +) were loaded in lanes 2, 8 and 10. Corresponding samples incubated without reverse transcriptase (cDNA –) are shown in lanes 3, 9 and 11. For the PCR, cDNA synthesized from 10 ng poly(A)$^+$ RNA and 1 ng plasmid DNA of ZEM1 (lanes 4 and 12), ZEM2 (lanes 5 and 13) and ZEM3 (lanes 6 and 14) were amplified using a MADS box primer together with a ZEM1/3 domain primer (see Fig. 3A). In lanes 1 and 12 a DNA size marker was loaded. The low level amplification products of the ZEM3 transcript are indicated by small arrow, while those of the ZEM5 signal are marked with a large arrow (lanes 2, 8 and 10).

present and to check for the possible existence of further ZEM transcripts, a series of RT–PCR experiments were carried out (for primers see Fig. 3A).

The amplification products of cDNA reverse-transcribed from poly(A)$^+$ RNA of maize seed, leaf and embryo using a MADS box-specific and a ZEM1/3 domain primer (lanes 2, 8 and 10) are shown in Figure 7. Products corresponding in size to that of ZEM1 (290 bp, compare with ZEM1 control in lanes 4 and 12) are clearly present in all three tissues. This confirms the presence of ZEM1 transcript also in young endosperm and embryo which is only faintly visible on RNA blot hybridization with probe (d) (Fig. 5). In the same PCR (lanes 2, 8 and 10), low level amplification products corresponding to ZEM3 were noted in all three tissues (450 bp, compare with ZEM3 control in lanes 6 and 14), indicated by a small arrow in Figure 7.

One further novel transcript, ZEM5 (see Fig. 3A), was identified by cloning a 170 bp RT–PCR product obtained with the above-mentioned primers from 17 dap maize endosperm. The corresponding RT–PCR product from 10 dap seed is indicated by the large arrow in lane 2, and its structure is given in Figure 3A. This product was also found in leaf and embryo (large arrow in lanes 8 and 10).

In a corresponding RT–PCR with the same MADS box primer and a ZEM1 domain primer, the ZEM4 transcript (Figs 3A and 5), was amplified in seed, leaf and embryo (data not shown). Its product size of 350 was 180 bp shorter than that of the co-amplified ZEM1 product (530 bp), which is consistent with the size difference seen between ZEM1 and ZEM4 transcripts detected by probe (d) (Fig. 5). Cloning and sequence analysis of the PCR products confirmed the structure of ZEM4 shown in Figure 3A. Nucleotide sequences of the ZEM1- and ZEM4-derived PCR products from seed, leaf and embryo (two were analysed from each mRNA and each tissue) clearly supported the conclusion that in the different maize tissues, identical transcripts were derived from the gene *ZEMa*, rather than from related genes of the same family.

Together, the RT–PCR experiments showed that besides the cDNA clones ZEM1, ZEM2 and ZEM3, the alternatively spliced forms ZEM4 and ZEM5 of the *ZEMa* gene exist, as indicated by nucleotide sequences and conserved exon–intron boundaries. Splice sites are used in common, for example those of ZEM4 and ZEM1 or the 3′ splice sites of ZEM5, ZEM3 and the first intron of ZEM1, and exon skipping occurs (see Fig. 3A), both characteristics of alternative splicing mechanisms. Moreover, sequence analysis supported the conclusions (i) that corresponding amplification products exist in different maize tissues, (ii) that all the isolated transcripts are due to the expression of one gene, *ZEMa*, and (iii) that all the *ZEMa* transcripts isolated so far are expressed in the endosperm.

Splicing of alternative transcripts in maize endosperm

The RNA gel blot analysis indicated that certain spliced *ZEMa* forms were preferentially expressed in specific tissues, for example, the predominance of the ZEM1 spliced form in leaf (Fig. 5) and in 20 dap endosperm (Fig. 6). As our interest was mainly centered on the role of MADS box genes in endosperm development, an experimental model was created to investigate alternative splicing of *ZEMa* transcripts specifically in this tissue. This approach was necessitated by the very high homologies found between *ZEMa*, *ZEMb* and *ZEMc* (97% sequence identity in corresponding domains). In these circumstances, RNase protection experiments would fail to distinguish between splicing of the same transcript and expression of different but related genes.

The experimental model consisted of endosperm cells that were transfected via particle bombardment and used as an *in vivo* system monitoring differential splicing events in this tissue. The putative 5′ alternative splicing of the first intron of ZEM1 and

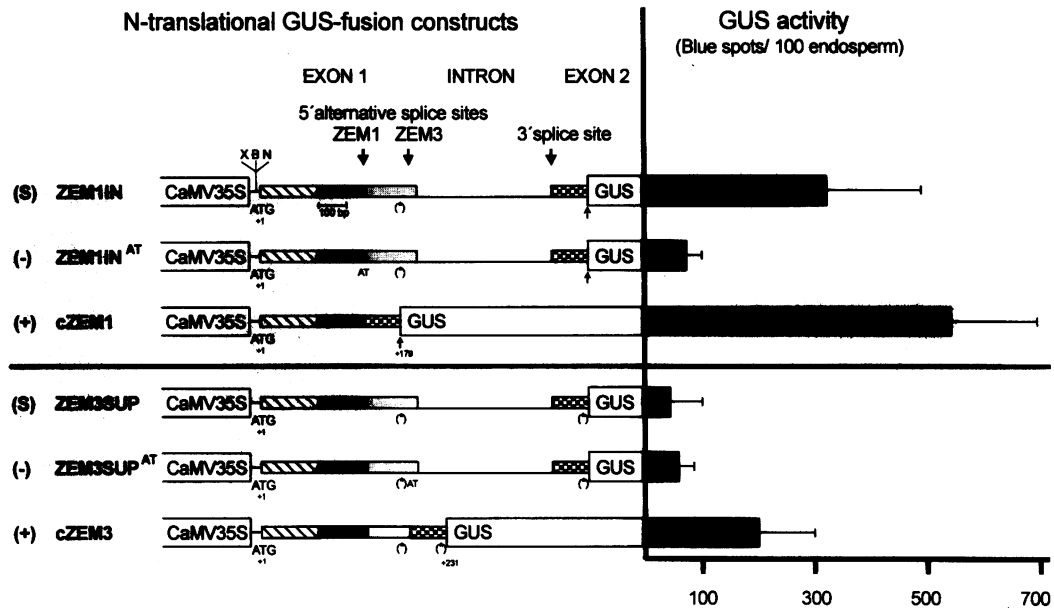

Figure 8. Splicing of alternative transcripts in maize endosperm. Schematic representation of the GUS fusion constructs and corresponding GUS activities in 12 dap maize endosperm. The splicing efficiency for ZEM1 is estimated by comparing GUS activities of the three upper constructs and that for ZEM3 by the three lower ones. (S) symbolizes the corresponding splicing constructs, (+) the 100% processed constructs and (–) the splice mutant constructs. The insertion of one nucleotide for the ZEM1-GUS in-frame reading is indicated by an arrow and the mutation of stop codons for the ZEM3-GUS in-frame reading by (*)-symbols. The mutations of the ZEM1 5′ splice site present in the ZEM1IN^{AT} and the ZEM3 5′ splice site in ZEM3SUP^{AT} are marked by an AT-symbol. The N-terminal GUS fusions are 179 amino acids for ZEM1 and 231 amino acids for ZEM3. The expression of the constructs is driven by the constitutive CaMV35S promoter. The resulting GUS activity of each construct is given in blue spots per 100 endosperm, with the standard deviations beside.

ZEM3 was chosen for analysis (see Fig. 3A). N-translational GUS fusions with an exon 1–intron–exon 2-fragment arrangement of the genomic clone of *ZEMa* were constructed (Fig. 8). The fusions were performed in such a way that the resulting spliced transcripts will only preserve the GUS reading frame, depending on which splicing event takes place. To achieve this, the following point mutations were introduced.

(i) The construct ZEM1IN contained an insertion of one nucleotide in the exon fused to GUS (as indicated by a small arrow in Fig. 8) preserving the GUS open reading frame after splicing from the ZEM1 5 splice site.

(ii) In the ZEM3SUP construct, two stop codons were suppressed, both indicated by asterisks in brackets, enabling a readthrough into the GUS gene only if splicing from the ZEM3 5′ splice site occured.

(iii) The natural stop codon of ZEM3 in the ZEM2/3 domain was also mutated in ZEM1IN (asterisks in brackets) to provide most equivalent constructs.

Thus, β-glucuronidase (GUS) activity obtained after expression of the construct ZEM1IN represents the splicing using the ZEM1 5′ splice site, whereas GUS activity from the construct ZEM3SUP represents splicing using the ZEM3 5′ splice site. The GUS activities of the two splicing constructs ZEM1IN and ZEM3SUP, (S), were compared with the activities of the corresponding cDNA GUS fusions cZEM1 and cZEM3, (+) in Figure 8. cZEM1 and cZEM3 were also modified by the corresponding point mutations present in the genomic constructs (arrow in cZEM1 and asterisks in brackets in cZEM3) to obtain an in-frame-reading of the GUS sequence. The comparisons provide an estimate of the splicing efficiency for ZEM1 and ZEM3.

In addition to the point mutations mentioned above, the GT at the 5′ intron junction of the two splicing constructs (S) was converted to an AT to abolish in each case a functional splice site (36) and generate the splice site mutant controls ZEM1IN^{AT} and ZEM3SUP^{AT}, (–) in Figure 8.

Particle bombardments of 12 dap maize endosperm showed GUS activities of ~60% for the use of the ZEM1 5′ splice site when the (S) and (+) constructs are compared (Fig. 8). The remaining GUS activity of 13% in the ZEM1 5′ splice mutant ZEM1IN^{AT} (–) might be due to splicing from the ZEM5 5′ splice site which is located further upstream (see Fig. 3A), but would also result in the same reading frame as splicing from the ZEM1 5′ splice site, or to the activation of cryptic splice sites (37) generating the same reading frame. This result showed, nevertheless that most of the GUS activity obtained from ZEM1IN is due to a splicing event from the ZEM1 5′ splice site and ruled out a mechanism other than splicing, such as the possibility of GUS arising from usage of internal start codons.

The GUS activity of ZEM3SUP (S) was ~21% of cZEM3 activity (+), but the ZEM3 5′ splicing mutant (ZEM3SUP^{AT}) showed similar activity. Thus, most of the GUS activity cannot be attributed to a fusion protein arising from the use of the ZEM3 5′ splice site. Possibly other 'cryptic' splice sites are used when this site is mutated; alternatively, an internal ATG could be responsible for the residual activity seen in both ZEM3SUP and ZEM3SUP^{AT}.

DISCUSSION

In the course of characterizing MADS box-related transcripts in maize endosperm, we have isolated *ZEMa*, a representative of a

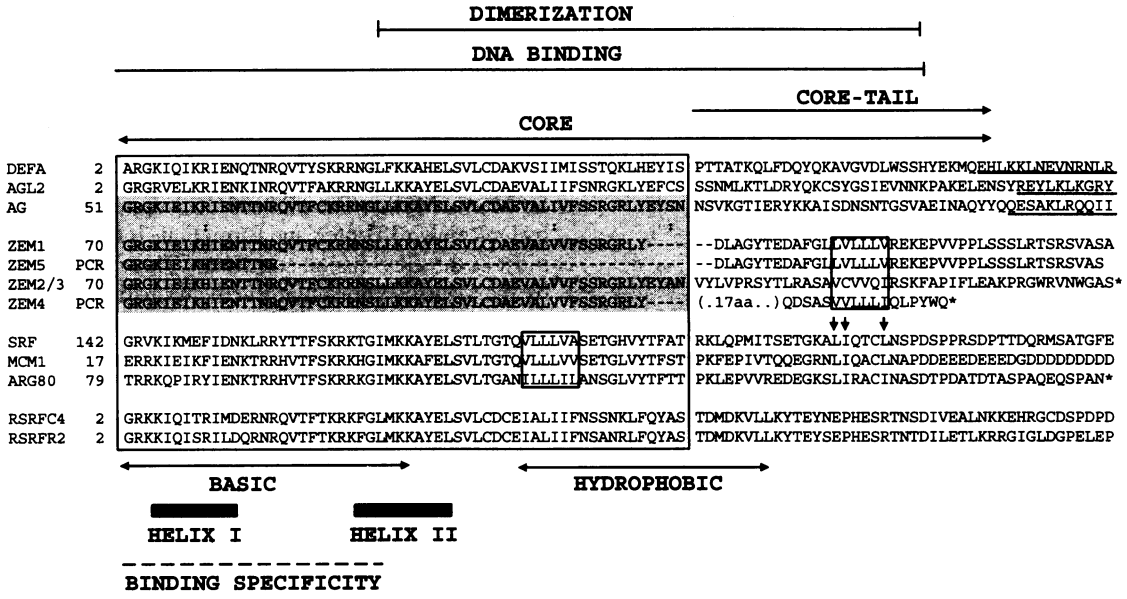

Figure 9. Alignment of MADS box cores. Partial sequences of the DEFA (12), AGL2 (14), AG (33), SRF (10), MCM1 (57), ARG80 (58), RSRFC4 (*MEF2A*), RSRFR2 (*MEF2B*) (20,21) and the predicted ZEM proteins derived from *ZEMa* are reported. The MADS box region is shown within the large box and the homology of the ZEM and the AG proteins is indicated by shading. The extent of the basic and hydrophobic regions, including two predicted regions of α-helix (exception: ARG80), the specificity-determining region and sequences involved in DNA-binding and dimerization for the SRF-like proteins are taken from (20). The positions of the conserved hydrophobic amino acids in the 'core-tails' of SRF, MCM1, ARG80 and the predicted ZEM proteins are indicated by black arrows. The highly hydrophobic motif in the ZEM 'core-tail' and in the MADS box of SRF, MCM1 and ARG80 is boxed. The N-termini of the K-boxes in DEFA, AGL2 and AG are underlined. The dashes in the protein sequences indicate gaps and the asterisks stop codons.

novel MADS gene family in plants. As this gene was expressed as a series of transcripts related by alternative splicing, we devised an assay to estimate the frequency of a given splicing event in transfected endosperm cells.

Novel domain structure of *ZEMa* transcripts

The *ZEMa* cDNA clones ZEM1–3, isolated from maize endosperm, possess an unusually short domain C-terminal of the MADS box ('core-tail') which differs from that of most other plant MADS box genes in several aspects. First, it lacks the K-box, a leucine-rich region with similarity to the coiled-coil domain of human type II keratin (38). The importance of the K-box in certain plant genes is indicated by the existence of temperature-sensitive mutants of *Antirrhinum* (*defA-101*) and *Arabidopsis* (*ap3-1*), in which a single lysine residue is affected within this domain (12,39). However, neither human SRF and members of the RSRF/MEF2 family, nor the yeast MCM1 and ARG80 (ARGRI) MADS box proteins contain such a motif.

The ZEM cDNA 'core-tail' (Fig. 9) is rich in non-polar amino acids, and thus extends the hydrophobic region of the MADS box. This extended hydrophobic stretch may provide the same function as a similar region within the MADS box of the SRF/MCM1/ARG80 group of MADS box genes (boxed motif), in which it is involved in dimerization of subunits, and interacting with accessory proteins (10,11,20). Intriguingly, this block is not found in the plant MADS box genes *AG, AGL2* or *DEFA*. The three isolated members of the ZEM2 family, *ZEMa*, *ZEMb* and *ZEMc*, are characterized by the highly conserved ZEM2/3 domain (43 amino acids) C-terminal to the MADS box (see Fig. 3B). By analogy, proteins encoded by the MEF2/RSRF family are characterized by a 29 amino acid 'core-tail', the MEF2 domain (21).

The absence of a K-box, the presence of hydrophobic motifs in the 'core-tail' of ZEM proteins similar to regions involved in dimerization of the SRF group proteins, and the existence of highly conserved 'core-tails' in the *ZEM* genes structurally relate them to animal MADS box genes rather than to other plant MADS genes. The variability in the C-terminal amino acid sequence between ZEM1, 2, 3, 4 and 5 might contribute to specificity of dimerization.

A second feature distinguishing the ZEM2 family from most other plant MADS box genes, is the relatively long region N-terminal to the MADS box. With the exception of the AGAMOUS (*Arabidopsis*) and ZAG1 proteins (maize), all other plant MADS box genes isolated so far encode only a few amino acids N-terminal to the MADS box, or start directly with the MADS box at the N-terminus. The N-terminal sequence is identical in ZEM1–3 and conserved between all *ZEM* genes (see Fig. 3B). It contains several putative recognition sites for phosphorylation by protein kinases (Fig. 2, shaded motifs) including for CDC2 protein kinase (consensus S*-P-x-R), the MAP2-related ERT kinases (P-x-S*/T*-P) and casein kinase II (S/T-$x_{(0-2)}$-E/D) (40–42). *In vitro* phosphorylation of similar recognition sequences present in the N-terminus of SRF, affected its DNA binding properties (43–46), and may play a role for ZEM protein function. Potential phosphorylation sites for a calmodulin-dependent protein kinase (consensus RxxS*/T*) are present within the conserved MADS box (RQVT) and in the ZEM2/3 C-terminal domain of all *ZEM* genes (RSYT) (47). The Ca^{2+}-dependence of calmodulin-governed phosphorylation may link environmental and hormonal influences to the proteins via Ca^{2+}-concentration and calmodulin activity in plants (48,49).

The MADS box of the three ZEM cDNA clones is nearly identical to that of the AGAMOUS protein in *Arabidopsis* with

the exception of four conservative amino acid exchanges (Fig. 9, shading). Because the specificity of DNA-binding of SRF-like proteins has been shown to be governed by the N-terminal part of the basic MADS box region, including helix I in Figure 9 (19,20,50), the putative target DNA-binding sequence of ZEM1–4 may resemble that of the AGAMOUS protein (TTT/A/GCC(A/T)$_6$GGA/T/CAA) (51).

Expression pattern of the *ZEMa* transcripts

The comparison of transcripts found in maize leaf, embryo and endosperm indicates that the gene *ZEMa* is expressed in all three tissues but also that certain spliced forms are generated preferentially in a particular tissue (for example, ZEM1 in the leaf and maturing endosperm). Other transcripts, such as ZEM4, are more uniformly present. A preferential expression pattern in certain tissues of alternatively spliced forms derived from MADS box genes was also observed for the human MEF2 gene family (21). Alternatively spliced forms of the *MEF2* (*MEF2A*) gene were preferentially expressed in muscle-related and neuronal tissues, although found in all cell-types investigated. The presence of a short exon in the mRNA of one form (aMEF2) increased tissue-selective expression compared with the same form without that exon (a*MEF2 = RSRFC9). Furthermore, tissue-specific alternatively spliced MEF forms existed together with ubiquitously distributed transcripts. The splicing pattern of the *ZEMa* gene shares several features with that of the *MEF2* (*MEF2A*) gene. The exon skipping of exon 2 of ZEM1 resulting in ZEM4 is similar to the aMEF2/a*MEF2 exon exclusion (21). As in the case of aMEF2, ZEM1 is preferentially, but not exclusively, expressed in specific tissues (leaf and maturing endosperm), whereas ZEM4, lacking the exon, is ubiquitously distributed, as was observed for a*MEF2.

Selection and function of splice sites in *ZEMa*

To date, few examples of alternative splicing in plants have been reported (52–54), and the *ZEMa* gene described here is the first example in which tissue-specific splicing occurs. The transient assay used in this paper is based on GUS activity as a criterion for the gene fusion being in or out of frame depending on the splicing event taken place. The validity of quantitating GUS expression levels by counting blue spots has been demonstrated by others (55,56) who have shown that the number of spots correlates well with promoter strength.

The transient expression studies in endosperm show that the use of the ZEM1 5′ splice site is clearly preferred to that of the ZEM3 or ZEM5 5′ splice sites in this tissue (Fig. 8). The expression pattern of the ZEM transcripts also suggests the preferential formation of the ZEM1 transcript in maturing endosperm and in leaf (Figs 5 and 6). This implies the existence of either tissue-specific splicing components that recognize certain splice sites, or sequence-specific effects of the general splicing machinery.

Alternative splicing further increases the repertoire of protein variants available from the *ZEMa* gene over and above those produced by heterodimerization with other MADS box proteins and by post-translational modifications. The significance of this process in maize development must await functional tests in transgenic plants.

ACKNOWLEDGEMENTS

We thank Professor Dr Alfons Gierl and Dr Goyo Hueros for the gift of the genomic and cDNA libraries. We are grateful to Professor Dr Heinz Saedler and Dr Günther Theißen for supplying genomic mapping data and useful discussions. Helpful advice on the particle bombardment technique was provided by Dr Martin Müller. The pRT103GUS vector was kindly provided by Dr Reinhard Töpfer, and Birgit Fiala is thanked for sequencing most of the GUS fusion constructs. This research was supported by the Deutsche Forschungsgemeinschaft (SFB 274).

REFERENCES

1 Nelson, O.E., Jr (1980) *Adv. Cereal Sci. Tech.*, **3**, 41–71.
2 Fedoroff, N.V. (1983) In Shapiro, J.A. (ed.), *Mobile Genetic Elements*. Academic Press, NY, pp 1–64.
3 Motto, M., Di Fonzo, N., Hartings, H., Maddaloni, M., Salamini, F., Soave, C. and Thompson, R.D. (1989) *Oxf. Surv. Plant Mol. Cell Biol.*, **6**, 87–114.
4 Schmidt, R.J., Burr, F.A., Aukerman, M.J. and Burr, B. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 46–50.
5 Lohmer, S., Maddaloni, M., Motto, M., Di Fonzo, N., Hartings, H., Salamini, F. and Thompson, R.D. (1991) *EMBO J.*, **10**, 617–624.
6 Ludwig, S.R., Habera, L.F., Dellaporta, S.L. and Wessler, S.R. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 7092–7096.
7 Goff, S.A., Cone, K.C. and Fromm, M.E. (1991) *Genes Dev.*, **5**, 298–309.
8 McCarty, D.R., Hattori, T., Carson, C.B., Vasil, V., Lazar, M. and Vasil, K. (1991) *Cell*, **66**, 895–905.
9 Schwarz-Sommer, Z., Huijser, P., Nacken, W., Saedler, H. and Sommer, H. (1990) *Science*, **250**, 931–936.
10 Norman, C., Runswick, M., Pollock, R. and Treisman, R. (1988) *Cell*, **55**, 989–1003.
11 Mueller, C.G.F. and Nordheim, A. (1991) *EMBO J.*, **10**, 4219–4229.
12 Schwarz-Sommer, Z., Hue, I., Huijser, P., Flor, P.J., Hansen, R., Tetens, F., Lönnig, W.-E., Saedler, H. and Sommer, H. (1992) *EMBO J.*, **11**, 251–263.
13 Ma, H. (1994) *Genes Dev.*, **8**, 745–756.
14 Ma, H., Yanofsky, M.F. and Meyerowitz, E.M. (1991) *Genes Dev.*, **5**, 484–495.
15 Pnueli, L., Abu-Abeid, M., Zamir, D., Nacken, W., Schwarz-Sommer, Z. and Lifschitz, E. (1991) *Plant J.*, **1**, 255–266.
16 van der Krol, A.R., Brunelle, A., Tsuchimoto, S. and Chua, N.-H. (1993) *Genes Dev.*, **7**, 1214–1228.
17 Mandel, T., Lutzinger, I. and Kuhlemeier, C. (1994) *Plant Mol. Biol.*, **25**, 319–321.
18 Fleming, A.J., Mandel, T., Roth, I. and Kuhlemeier, C. (1993) *Plant Cell*, **5**, 297–309.
19 Passmore, S., Elble, R. and Tye, B.-K. (1989) *Genes Dev.*, **3**, 921–935.
20 Pollock, R. and Treisman, R. (1991) *Genes Dev.*, **5**, 2327–2341.
21 Yu, Y.-T., Breitbart, R.E., Smoot, L.B., Lee, Y., Mahdavi, V. and Nadal-Ginard, B. (1992) *Genes Dev.*, **6**, 1783–1798.
22 Breitbart, R.E., Liang, C., Smoot, L.B., Laheru, D.A., Mahdavi, V. and Nadal-Ginard, B. (1993) *Development*, **118**, 1095–1106.
23 Leifer, D., Krainc, D., Yu, Y.-T., McDermott, J., Breitbart, R., Heng, J., Neve, R.L., Kosofsky, B. and Nadal-Ginard, B. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 1546–1550.
24 Kaushal, S., Schneider, J.W., Nadal-Ginard, B. and Mahdavi, V. (1994) *Science*, **260**, 1236–1240.
25 Minty, A. and Kedes, L.J. (1986) *Mol. Cell. Biol.*, **6**, 2125–2136.
26 Pereira, A., Schwarz-Sommer, Z., Gierl, A., Peteson, P.A. and Saedler, H. (1985) *EMBO J.*, **4**, 17–23.
27 Bartels, D. and Thompson, R.D. (1983) *Nucleic Acids Res.*, **11**, 2961–2977.
28 Dellaporta, S.L., Wood, J. and Hicks, J.B. (1983) *Plant Mol. Biol. Rep.*, **1**, 19.
29 Kawallek, P., Somssich, I.E., Feldbrügge, M., Hahlbrock, K. and Weisshaar, B. (1993) *Plant Mol. Biol.*, **21**, 673–684.
30 Töpfer, R., Pröls, M., Schell, J. and Steinbiß, H.-H. (1988) *Plant Cell Rep.*, **7**, 225–228.
31 Donovan, G.R. and Lee, J.W. (1977) *Austr. J. Plant Physiol.*, **5**, 81–87.
32 McCully, D.E., Gengenbach, B.G., Smith, J.A., Rubenstein, I., Connelly, J.A. and Park, W.D. (1984) *Plant Physiol.*, **74**, 389–394.

33 Yanofsky, M.F., Ma, H., Bowman, J.L., Drews, G.N., Feldmann, K.A. and Meyerowitz, E.M. (1990) *Nature*, **346**, 35–39.
34 Schmidt, R.J., Veit, B., Mandel, M.A., Mena, M., Hake, S. and Yanofsky, M.F. (1993) *Plant Cell*, **5**, 729–737.
35 Kowles, R.V. and Phillips, R.L. (1988) *Int. Rev. Cytol.*, **112**, 97–136.
36 Aebi, M., Hornig, H., Padgett, R.A., Reiser, J. and Weissmann, C. (1986) *Cell*, **47**, 555–565.
37 Green, M.R. (1991) *Annu. Rev. Cell Biol.*, **7**, 559–599.
38 Tyner, A.L., Eichman, M.J. and Fuchs, E. (1985) *Proc. Natl. Acad. Sci.* USA, **82**, 4683–4687.
39 Jack, T., Brockman, L.L. and Meyerowitz, E.M. (1992) *Cell*, **68**, 683–697.
40 Alvarez, E., Northwood, I.C., Gonzalez, F.A., Latour, D.A., Seth, A., Abate, C., Curran, T. and Davis, R.J. (1991) *J. Biol. Chem.*, **266**, 15277–15285.
41 Kenelly, P.J. and Krebs, E.G. (1991) *J. Biol. Chem.*, **266**, 15555–15558.
42 Pearson, R.B. and Kemp, B.E. (1991) *Methods Enzymol.*, **200**, 62–81.
43 Manak, J.R., de Bisschop, N., Kris, R.M. and Prywes, R. (1990) *Genes Dev.*, **4**, 955–967.
44 Manak, J.R. and Prywes, R. (1991) *Mol. Cell. Biol.*, **7**, 3652–3659.
45 Janknecht, R., Hipskind, R.A., Houthaeve, T., Nordheim, A. and Stunnenberg, H.G. (1992) *EMBO J.*, **11**, 1045–1054.
46 Marais, R.M., Hsuan, J.J., McGuigan, C., Wynne, J. and Treisman, R. (1992) *EMBO J.*, **11**, 97–105.
47 Cohen, P. (1988) *Proc. R. Soc. Lond. Ser. B.*, **234**, 115–144.
48 Kelly, G.J. (1984) *Trends Biochem. Sci.*, **9**, 4–5.
49 Gilroy, S., Hughes, W.A. and Trewavas, A.J. (1987) *Development*, **100**, 181–184.
50 Hayes, T.E., Sengupta, P. and Cochran, B.H. (1988) *Genes Dev.*, **2**, 1713–1722.
51 Shiraishi, H., Okada, K. and Shimura, Y. (1993) *Plant J.*, **4**, 385–398.
52 Masson, P., Rutherford, G., Banks, J.A. and Fedoroff, N.V. (1989) *Cell*, **58**, 755–765.
53 Grotewold, E., Athma, P. and Peterson, T. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 4587–4591.
54 Luehrsen, K.R., Taha, S. and Walbot, V. (1994) *Prog. Nucleic Acid Res. Mol. Biol.*, **47**, 149–193.
55 Ellis, D.D., McCabe, D., Russell, D., Martinell, B. and McCown, B.H. (1991) *Plant Mol. Biol.* **17**, 19–27.
56 Luan, S. and Bogorad, L. (1992) *Plant Cell* **4**, 971–981.
57 Passmore, S., Maine, G.T., Elble, R., Christ, C. and Tye, B.-K. (1988) *J. Mol. Biol.*, **204**, 593–606.
58 Dubois, E., Bercy, J. and Messenguy, F. (1987) *Mol. Gen. Genet.*, **207**, 142–148.

NOTE ADDED IN PROOF

ZEMa is located on λEMBL4-20a of Fischer *et al.* [(1995) *Nucleic Acids Res.* **11**, 1901–1911], and contains segments A, B and C of the TMZ1 family described by these authors.