

A Seed-Specific Heat Shock Transcription Factor Involved in Developmental Regulation during Embryogenesis in Sunflower

Concepción Almoguera, Anabel Rojas[‡], Juan Díaz-Martín, Pilar Prieto-Dapena, Raúl Carranco[§], and Juan Jordano^{*}.

Instituto de Recursos Naturales y Agrobiología, C.S.I.C. Apartado 1052, 41080 Sevilla. SPAIN.

[‡]Current address: Cardiovascular Research Institute, University of California, San Francisco, 505 Parnassus Ave., San Francisco, California 94143-0130, U.S.A.

[§]Current address: Institute of Developmental and Molecular Biology and Department of Biology, Texas A&M University, College Station, TX 77843-3155, U.S.A.

^{*}Corresponding author; phone 34-954-624711 (ext. 145) fax 34-954-624002; email fraga@cica.es

Running title: Embryo-specific HSF in sunflower

SUMMARY

We report the cloning and functional characterization of the first heat-shock transcription factor that is specifically expressed during embryogenesis in the absence of environmental stress. In sunflower embryos this factor, HaHSFA9, *trans*-activated promoters with poor consensus heat-shock *cis*-elements (HSEs), including that of the seed-specific *Hahsp17.6G1* gene. Mutations that improved the HSE consensus at the *Hahsp17.7G4* promoter impaired transient activation by HaHSFA9 in sunflower embryos. The same mutations did not affect heat-shock induced gene expression of this promoter in transgenic tobacco plants; but reduced the developmental activation by endogenous HSFs in seeds. Sunflower, and perhaps other plants as tobacco, differs from the vertebrate animal systems in having at least one specialized HSF with expression and (or) activation patterns strictly restricted to embryos. Our results strongly indicate that HaHSFA9 is a transcription factor critically involved in the developmental activation of *Hahsp17.6G1* and in that of similar target genes as *Hahsp17.7G4*.

INTRODUCTION

In eukaryotes, the heat-shock response and some developmental processes are under the control of a family of conserved DNA-binding proteins known as the heat-shock transcription factors (HSFs)¹. Whereas in some systems, as in *Drosophila melanogaster* this regulation involves a single HSF (1), multigenic families of HSFs participate in vertebrate and in plant systems. These families have different size, which, together with particular gene expression and activation patterns for the HSFs, might have consequence in the degree of overlapping of regulatory functions mediated by these factors. The specific role of the different HSFs is mostly unknown, particularly for the plant HSFs; and for involvement in developmental processes, as the regulation of gene expression during embryogenesis (See for example, Ref. 2 and the reviews in Refs. 3 and 4).

In vertebrate systems, three different HSFs (HSF1, HSF2, and HSF3) have ubiquitous expression patterns (for example, Refs. 5, 6; and the review in Ref. 3). A fourth HSF found in humans displays tissue-specific expression patterns, which suggested specialized functions but not related to embryogenesis (HSF4, Ref. 7). Plants contain the highest number of HSF genes in eukaryotes. This is inferred from *in silico* analyses from the fully sequenced *Arabidopsis thaliana* model; and from functional analyses of different cloned HSFs in tomato, Arabidopsis, and other plants (reviewed in Refs. 4 and 8, see references therein). Plant HSFs share unique structural and phylogenetic relationships compared to the vertebrate HSFs (9). Fifteen of the 21 putative HSFs from *Arabidopsis thaliana* thus contain an insertion of 21 amino acid residues in the oligomerization domain (characteristic of the plant class A HSFs), whereas class B HSFs have no such insertion. Gene expression studies for plant HSFs are very scarce, with only fragmentary data at the mRNA level and even scarcer reports for protein accumulation (also reviewed in Refs. 4 and 8, see references therein). The potential higher specialization of plant HSFs, compared to vertebrate animal systems, is thus mostly unexplored.

Observations from our lab indicated the involvement of peculiar HSF(s) and HSE(s) in developmental regulation, of small heat-shock protein (sHSP)

gene expression, during zygotic embryogenesis in sunflower and other plant systems (10-12). Here, we exploited these observations to clone by one hybrid interaction in yeast HaHSFA9, a class A HSF with a unique embryo-specific gene expression pattern. We show that HaHSFA9 *trans*-activates two sHSP gene promoters (*Hahsp17.7G4* and *Hahsp17.6G1*), the later of them embryo-specific; and that transcriptional regulation by HaHSFA9 in sunflower embryos depends on particular characteristics of the DNA sequences used for cloning (imperfect consensus HSE sequences). We extend similar observations to other endogenous HSFs in transgenic tobacco; thus, we demonstrate the involvement of at least a specialized HSF for sHSP gene regulation during sunflower (and perhaps other plant) embryogenesis.

EXPERIMENTAL PROCEDURES

Construction of the sunflower embryo cDNA library

mRNA from 14 dpa sunflower embryos (*Helianthus annuus*, cv P113HS, seeds from ARLESA Semillas S.A., Spain) was isolated from total RNA using the Poly (A) Quik^R mRNA Isolation Kit (Stratagene). We used 5 μ g Poly (A) for cDNA synthesis performed with the ZAP-cDNA^R Kit (Stratagene). The cDNA, which included 5' *EcoRI* adapters, was digested with *XhoI*. This generated the 3' cloning site. The digested cDNA was size-fractionated with SizeSepTM 400 Sepharose CL-4B spun columns (Amersham Biosciences), and directionally inserted within the *EcoRI* and *SalI* sites in the polylinker of the pGAD424 vector (Clontech). The primary cDNA library contained 830,000 individual transformants, all with a cDNA insert of average size of 1.3 Kb.

One-hybrid cloning in yeast

General methods for one-hybrid cloning and related experimental manipulations in yeast were as described (13). We constructed a yeast strain derivative of YM4271 (Clontech). In this strain, we integrated a *HIS3* reporter gene construct [(G4HSE I)_{x3}::*HIS3*] containing a trimer of the proximal HSE sequences from *Ha hsp17.7 G4* (from positions -89 to -57, see Ref. 11). This construct was obtained from plasmid pHISi (Clontech) by insertion, between the *XbaI* (end-filled with klenow DNA-polymerase) and *EcoRI* sites, of the DNA fragment generated by annealing of the following two complementary oligonucleotides:

5'aattcTTCTTCAAGCTTCAAGACAATCCTAGAAATTACTTCTTCAAGC
TTCAAGACAATCCTAGAAATTACTTCTTCAAGCTTCAAGACAATCCTAGAAA
TTAC3' (top strand) and,

5'GTAATTTCTAGGATTGTCTTGAAGCTTGAAGAAGTAATTTCTAGGA
TTGTCTTGAAGCTTGAAGAAGTAATTTCTAGGATTGTCTTGAAGCTTGAAGA
Ag3' (bottom strand).

For one-hybrid screening, the (G4HSE I)_{x3}::*HIS3* reporter yeast strain was transformed with DNA prepared from the embryo cDNA library, after amplification of 1,660,000 primary clones. Five million yeast transformants were

plated on SD-HIS-LEU+15 mM 3-AT. After 4-8 days of growth at 30°C, twenty-four putative positive yeast clones were selected for further analyses. Four cDNAs encoded the same HSF and are described in this work. The cDNA for one of these clones, HSFA9-36, was subcloned in pBluescript SK+ (Stratagene); thus obtaining pSKHSFA9-36.

RNA ligase-mediated amplification of the 5' cDNA end of HaHSFA9 and assembly of the full-length clone

The full-length 5'-end of HaHSFA9 cDNA was obtained by rapid amplification mediated by selective ligation to decapped mRNA of an RNA oligonucleotide using T4 RNA ligase (RLM-RACE). We employed the general materials and conditions included in the GeneRacer™ Kit (Invitrogen). Specific conditions in our case included the use of 1 µg total RNA from 14 dpa sunflower embryos as starting material. Reverse transcription was performed at 42°C using random primers. The nested amplification of the full-length cDNA 5'-end was performed at annealing temperature of 60°C. We used the following oligonucleotide primer pairs: GeneRacer™ 5' Primer (for the 1st PCR) or GeneRacer™ 5' Nested Primer (for the 2nd PCR), and HSFA9-RACE primer, which spans between positions 265 and 241 in the noncoding strand of HaHSFA9 cDNA. The PCR-amplified cDNA sequences were cloned in vector pCR^R4-TOPO^R (plasmid pCR^R4-TOPO^R::HSFA9-5'), after addition of 3' A-overhangs in the conditions of the TOPO TA Cloning^R Kit (Invitrogen). The full-length HaHSFA9 cDNA clone (plasmid pSKHSFA9-F) was assembled in vector pBluescript SK+ by ligation of a 120 bp DNA fragment from pCR^R4-TOPO^R::HSFA9-5' with the rest of the cDNA sequences in plasmid pSKHSFA9-36. The 120 bp fragment was obtained by PCR amplification (using the primers TOPO-1, 5'-atcATATTCTCCTTCAAAA-3', and HSFA9-RACE), followed by digestion with *StyI* (the joining site, at position 122 in the HaHSFA9 cDNA). The *EcoRV* half site (lower case nucleotides) in primer TOPO-1 provided the 5'-insertion point in the vector polylinker sequence.

Ribonuclease protection assays

In vitro transcription, probe purification, and ribonuclease protection assays were performed as described (10). To obtain the HaHSFA9 riboprobe, we cloned (in pBluescript SK+) the *EcoRI* fragment from plasmid pSKHSFA9-36, and generated pSKHSFA9-RI. The orientation of the insert in pSKHSFA9-RI was such that after linearization with *XhoI*, transcription from the T3 promoter originated the, 507-nucleotide, antisense probe used in our experiments.

Antibody production and purification for western immunodetection

The antigen was expressed in BL21 *E. coli* cells from plasmid pRSET-HSFA9 Δ *BglII*. This plasmid included the HaHSFA9 cDNA sequences, between *BglII* (position 683) and *PstI* (position 1247), inserted in the vector pRSET-A (Invitrogen). For rabbit immunizations, we purified the antigen by TALON^R metal affinity chromatography (Clontech). The whole serum was affinity purified using the antigen bound to ImmobilonTM-P (Millipore) (14). The purified antibodies were used at final dilution of 1:4000. For hybridization with the Rabbit anti-HSC70 antibodies (diluted to 1/2000, StressGen Biotechnologies Corp.), blots were stripped using the *Re-Blot Plus* recycling kit (Chemicon International).

Functional assays in yeast and transient expression analyses in sunflower embryos

The expression vector used for the functional assays in yeast contained the full-length HaHSFA9 cDNA, which was excised from plasmid pSKHSFA9-F as a *SalI-SacI* DNA fragment. This fragment was used to replace the LpHSFA2 *SalI-SacI* DNA sequences in a previously described plasmid derived from pAD5 Δ (15). The RSY4 yeast (*Saccharomyces cerevisiae*) strain and the protocols for β -galactosidase activity assays were also described (15). To substitute ScHSF1 for HaHSFA9 in RSY4, the HSF1 mutant strain with the ScHSF1 plasmid was transformed with the pAD5 Δ -HSFA9 plasmid. The ScHSF1 plasmid was then eliminated by growth in SD-LEU medium followed by plating at low density in SD+FOA. Final HSF plasmid contents was verified by PCR.

We previously described the conditions used for transient expression in sunflower embryos and the employed GUS reporter plasmids (16). Double digestion of the full-length cDNA plasmid (pSKHSFA9-F) with *EcoRV* and *SacI* allowed the substitution of the GUS sequences in plasmid pBI221 for those of HaHSFA9. That way, we obtained the effector plasmid (p35S::HSFA9) used in the transient activation assays.

Chimeric gene expression in tobacco transgenic plants

By cloning in pBIN19 the *Sall-Sacl* insert of a previously described pBluescript SK-derived plasmid, [-1132::GUS(mut P), see Ref. 16], we constructed the binary version of the P mutant chimeric gene. We then generated different primary (T0) tobacco transgenic plant using published procedures (17). The progeny (T1) of plants with the P mutant chimeric gene was also obtained. In parallel, we obtained transgenic plants (T0 and T1) for the previously described binary plasmid with wild type chimeric gene [-1132::GUS (WT), see Ref. 17]. We compared the developmental (using the T0 plants), and the heat stress induced expression patterns (using the T1 plants) of different independent transformants (between 4 and 10 per construct), with a similar number (1-3) of integration events. The fluorometric GUS assays during seed maturation, and the statistical analyses were as previously reported (11). Heat stress treatments were for 2.5 h at 42°C, followed by recovery for 3h at 25°C prior to GUS activity determinations.

RESULTS

Yeast one-hybrid cloning and sequence of HaHSFA9

To isolate *trans*-acting factors involved in the developmental activation of sHSP gene promoters during the embryogenesis of sunflower, we used the yeast one-hybrid cloning approach (18). Our previous characterization of two such promoters, *Hahsp17.6G1* and *Hahsp17.7G4* pointed to a similar and peculiar sequence arrangement in their functionally defined HSEs (10, 11). In the case of the *Hahsp17.7G4* promoter we found that the proximal HSE I region would contain overlapping binding sites for HSFs and distinct factors involved in activation during earlier stages of embryogenesis (11). Thus, we decided to use that region as bait that could allow cloning several factors at a time, including the HSF(s) involved in the late developmental activation of both promoters. We directly trimerized these sequences without adding nucleotides that would extend the arrangement of GAA-like and TTC-like HSE core elements with the natural two-nucleotide separation: nnGAAnnTTCnn, etc. We obtained 24 positive colonies. These clones were found to represent different groups. The nucleotide and deduced amino acid sequences showed that one group, with four independent cDNA isolates, encoded a putative transcription factor belonging to the HSF family. Interestingly, in none of them was the *GAL4* activation domain cloned in frame with the deduced amino acid sequences. This indicated that the HSF-encoded sequences were able to activate transcription by themselves, at least in the promoter context of the yeast reporter strain containing the multimerized bait.

We assembled a complete cDNA from the nucleotide sequences determined from the four independent HSF clones² (Figure 1). The four clones included identical nucleotide sequences predicted to encode a protein with DBD and oligomerization domains (with HR-A/B) that identify HSFs (see for example the reviews in Refs. 3 and 19). On the other hand, the nucleotide sequences of these clones maintained the same reading frame to the cloned 5'-ends. Thus, we could not exclude that the predicted HSF protein was N-terminally truncated. Full-length 5'-end cDNA sequences were cloned by RLM-RACE. This added 40

nucleotides that maintained the reading frame without including supplementary initiation codons (Figure 1).

We named the predicted protein HaHSFA9 because its amino acid sequence clearly showed the characteristics of a class A plant HSF (Figure 1, Ref. 4), and because it was most similar to the previously described AtHSFA9 (Ref. 4). Thus, the DBD and HR-A/B region of HaHSFA9 was 54.5% identical to AtHSFA9 (data not shown). HaHSFA9 showed the distinctive insertion of 21 amino acid residues between the repeat regions HR-A/B in the oligomerization domain (Figure 1, Ref. 9). HaHSFA9 also had a NLS located adjacent to the HR-A/B region; as well as a C-terminal HR-C region and AHA motifs, which all are classification criteria for class A HSFs (9). However, HaHSFA9 showed some peculiar characteristics compared to other HSFs. This included the very unusual presence in the DBD of an arginine residue in an invariant position encoding glycine in plant and non-plant HSFs (see R at position 131, in Figure 1). Another unusual amino acid change in HaHSFA9 replaced a Tyrosine residue conserved in the majority of class A plant HSFs (4), including AtHSFA9 (see R at position 114, in Figure 1). We indicate other unusual structural features (see Figure 1): an acidic region N-terminal of the DBD domain (amino acid positions 7-66); and peculiar, putative, AHA motifs. These AHA motifs would be related to those found in the activation domains of other class-A HSFs (reviewed in Ref. 4). The total number of AHAs and their molecular position (in particular that of C-terminal motif) is different for HaHSFA9.

Embryo-specific expression of HaHSFA9

We next investigated the expression patterns of HaHSFA9. We started by determining HaHSFA9 mRNA accumulation patterns using an RNase A protection approach. We performed hybridizations using a collection of total RNA samples from sunflower embryos and from different organs from control and stressed plants. We previously used the same RNA samples for other analyses of developmental mRNA accumulation, and to analyze the heat- and water-stress response of other sunflower genes by RNase A protection. This provided positive controls for the different stress treatments and for normal

developmental expression (14, 17, 20). For the hybridizations analyzed here, we employed a riboprobe with sequences from the non-coding strand of HaHSFA9. We expected that, after digestion with RNase A, full protection of this probe would produce a protected RNA fragment of 395 nucleotides. Results presented in the top of Figure 2 show that HaHSFA9 mRNAs accumulated from early zygotic embryogenesis in developing seeds, and disappeared very early during germination after seed imbibition. The mRNAs were detected at 8 dpa and their accumulation slightly increased up to 18 dpa and then started to decrease. These transcripts were barely detectable at 5 dpi, and were undetectable at control temperatures in older seedlings (14 dpi) and in different organs of adult plants, as in leaves or stems (Figure 2, bottom).

Because heat stress treatments have been shown to induce the mRNA accumulation of some plant HSFs (reviewed in Ref. 4), we investigated such behavior for HaHSFA9. In addition, we investigated the effect of water stress treatment, as the expression of possible target genes for HaHSFA9 is induced during the desiccation stages of embryogenesis (10, 17). Neither stress treatment, in particular heat, induced HaHSFA9 mRNA accumulation after embryogenesis. We only detected very low levels of water-stress induced transcripts in stems and leaves, for which we needed to substantially increase autoradiography times (Figure 2, bottom).

In these experiments we also detected at least two additional RNA protected fragments (arrowed with “2” in Figure 2). The most likely explanation for such fragments would be either the detection of unspliced mRNA or that of mRNAs that are shorter from the 5'-end (see the diagram at the bottom of Figure 2). The presence of an intron in the DBD region of the HaHSFA9 gene is supported by PCR amplification using primers from the cDNA flanking the conserved intron position (Ref. 4, data not shown).

We obtained specific antibodies against the HaHSFA9 protein (see Experimental Procedures) and analyzed expression at the protein level during zygotic embryogenesis and in control and stressed plants. The results of western immunodetection experiments showed in Figure 3 confirmed most of the results observed at the mRNA level (Figure 2): the detection of the

HaHSFA9 protein in embryos from 8 dpa, and its absence from leaves and stems of control and heat stressed plants. The protein persisted at low levels in germinated seeds but as observed for the mRNAs, only to 5 dpi. We also noticed some differences, as we could not detect HaHSFA9 protein accumulation in response to water stress in vegetative tissues. The detected protein has an apparent molecular weight of 53.7 kD.

Functional analysis of HaHSFA9 in yeast and in plant embryos

We confirmed the integrity and transcriptional activity of the HaHSFA9 protein(s) encoded by the nucleotide sequences assembled in the full-length cDNA². First, we constructed an appropriate yeast expression plasmid containing the complete coding sequence from the putative first methionine to position 1252 in the 3'-UTR. This plasmid was used for functional replacement analysis of ScHSF1, the sole HSF in the yeast *Saccharomyces cerevisiae*. The results in Figure 4A (top) show that this strain was as viable as the original ScHSF1 strain, at 28°C, and even at 35°C. In contrast only ScHSF1 allowed growth at 37°C. The two strains grew at 28°C in liquid YPD medium with a similar doubling time (2-2.5h, data not shown). These results demonstrated functional integrity of the encoded HaHSFA9 protein(s). This is comparable to results described for other class A HSFs from *Lycopersicon peruvianum* (15).

We also used the HaHSFA9 yeast strain to verify the transcriptional activation potential of HaHSFA9 and its dependence on HSE sequence complexity. We constructed LacZ reporter gene plasmids containing the natural HSE region of *Hahsp17.7G4* (WT), or a mutant version (m) of it with severe effects on chimeric HSF-dependent gene expression in transgenic plants (m=mutE in Ref. 11). These reporter plasmids contained two HSE repeats (promoter proximal and distal, with its natural spacing in between) in front of minimal yeast promoter, instead the oligomerized proximal HSE repeat used for one-hybrid cloning. HaHSFA9 strains transformed with the WT reporter showed high levels of β -galactosidase activity, and this transcriptional activation by HaHSFA9 was severely reduced in the mutant reporter strain (Figure 4A, bottom: G4HSE). Interestingly the activity measured in this strain was still

higher than background values reported for other plant HSFs and mutant HSE reporter plasmids in similar experiments performed without temperature stress (c.a. ≤ 1 , see Ref. 21). The m HSE array still has a small number of, gapped, GAA- and TTC-like core elements (Ref. 11, see also Figure 5A). In consequence, our results could reflect a singular capacity, of HaHSFA9, for binding or/(and) activation from these HSEs. We tested and confirmed this hypothesis by observing transcriptional activation of another LacZ reporter plasmid containing the HSE region from *Hahsp17.6G1* (Figure 4A, bottom: G1HSE). This HSE contains only a promoter-distal, shorter and gapped array of core elements (10). In this case, the HSE mutations reduced transcriptional activation by HsHSFA9 further than in the case of the G4HSE mutant plasmid, and to values more similar than those previously reported for background levels.

These results suggested that HaHSFA9 might activate transcription from the *Hahsp17.6G1* promoter in plant seeds. We performed transient expression assays in sunflower embryos. We used two chimeric genes that were previously analyzed in transgenic plants: -1486::GUS and -1486(m)::GUS. Such genes only differ in three crucial nucleotide substitutions, located in the HSE region of *Hahsp17.6G1*. These substitutions impaired transcriptional activation during desiccation stages of zygotic embryogenesis (20). The results shown in Figure 4B demonstrated that HaHSFA9, expressed from an appropriate plant effector plasmid, efficiently *trans*-activated the -1486::GUS gene. Mutations in the -1486(m)::GUS gene abolished transcriptional activation by HaHSFA9.

Embryo-specific effect of mutations that improve the HSE in the *Hahsp17.7G4* promoter

In sunflower, we previously described that there is another sHSP gene, *Hahsp18.6G2*, which is not transcriptionally activated in developing embryos, but it is responsive to heat stress. The HSE arrays of *Hahsp18.6G2* are more “perfect” than those in *Hahsp17.6G1* or *Hahsp17.7G4*, in the sense that they lack gaps between the core repeats, which are thus more similar to the aGAA/TTCt consensus sequences (10). These observations prompted us to study the

effect, on transcriptional activation by HaHSFA9 and related factors, of nucleotide substitutions that make HSEs more similar to those in *Hahsp18.6G2*.

We analyzed these mutations in the natural context of the *Hahsp17.7G4* promoter; as its proximal HSE sequences were used to clone HaHSFA9, and *Hahsp17.7G4* is both developmentally activated and heat stress-responsive. In Figure 5A, we depict the “perfect” (P) *Hahsp17.7G4* HSE mutant compared to the unaltered (WT) and strong negative mutant (m) HSE versions. We first investigated the effect of mutant HSEs on transient activation by HaHSFA9 in sunflower embryos (see Figure 5B). We confirmed that HaHSFA9 very efficiently activated transcription of the WT gene. In these assays, the m HSE completely abolished transcriptional activation by HaHSFA9, as the reporter activity observed with the HaHSFA9 effector plasmid did not differ from basal control levels ($F= 0.262$, $P= 0.61$). The P mutant HSE produced two contrasting effects: it significantly increased basal activity levels ($F= 40.01$, $P= 0.0001$), but also clearly reduced transcriptional activation by HaHSFA9 ($F= 9.82$, $P= 0.002$). Thus, HaHSFA9 differs from other HSFs present in sunflower embryos in its capacity of binding and (or) transcriptional activation from the P HSE.

We finally investigated the effect of the P mutant HSE on the developmental activation and heat stress response of the chimeric gene in tobacco plants, where we previously showed a faithful reproduction of transcriptional regulation of *Hahsp17.7G4* (11, 17). The results in Figure 5C show that, compared to the WT gene, mutant P was not affected in the heat stress response as determined by experiments performed with seedlings. Statistical analyses determined that both the control ($F= 0.06$, $P= 0.81$) and induced ($F= 0.002$, $P= 0.96$) activity levels of the two genes were similar. In contrast, the WT and P genes differed in their developmental activation during zygotic embryogenesis (Figure 5D). The P gene showed reduced activation during desiccation stages of late embryogenesis (28 dpa, $F= 36.5$, $P= 0.0001$), but not earlier (16 dpa, $F= 1.67$, $P= 0.2$). We conclude that the effect of mutant P is not only embryo-specific but also stage-specific. Furthermore, such decrease is analogous to what was observed in sunflower embryos for transcriptional activation of the same genes by HaHSFA9 (compare Figure 5B

and Figure 5D). The transgenic plant system (Figure 5D) showed a higher reduction in the activity of the P gene, most likely explained by differences between transient (Figure 5B) and stable (Figure 5D) expression conditions.

DISCUSSION

Embryo-specificity of HaHSFA9

The embryo-specific expression pattern observed for HaHSFA9 is supported by the results of mRNA (Figure 2), and protein (Figure 3) accumulation assays. RNase A protection (Figure 2) combines sensitivity comparable to RT-PCR with the highest specificity (22). In addition to the data presented in Figures 2 and 3, we have confirmed the results for mRNA using RT-PCR, and we were unable to detect the HaHSFA9 protein in roots from adult plants (data not shown). We thus explored most vegetative organs and tissues in young and adult plants, in the absence and presence of stress (heat- or drought-). We conclude that, within the detection limits in our experimental conditions, the HaHSFA9 protein is expressed only in embryos, and it disappears shortly after seed germination. The accumulation of the HaHSFA9 protein precedes and appears connected to the embryo desiccation phase, when potential target genes reach maximal expression levels (see for example, 10 and references therein).

The strict embryo-specificity of protein accumulation (Figure 3) restricts the function(s) of HaHSFA9, which appears as a regulator of developmental gene expression during zygotic embryogenesis in sunflower. This would be different from animal systems. In humans, only HSF4 has strict tissue-specific protein expression patterns in the lung and brain. However, the functions proposed for hHSF4 in these organs would not be related with embryogenesis (7). In vertebrates protein and mRNA accumulation assays showed that, HSF1, HSF2, and HSF3 have ubiquitous expression patterns (reviewed in Ref. 3). Even in the case of HSF2, for which functions related with embryo development have been proposed (2), expression is ubiquitous at low level (5). This suggests additional non-embryonic functions of HSF2 in most cell-types. HSF1 and HSF3 also have other functions in addition to their involvement in heat-shock tolerance. Specific interactions, with different transcription factors, could confer distinct functional specificity to the ubiquitous expression pattern of animal HSFs. An example of this is the interaction between c-Myb factor and HSF3 (3, 23).

Expression analyses for the more numerous plant HSFs are still scarce. For example, for the 21 HSFs predicted from the genomic sequence of *Arabidopsis thaliana* there are only fragmentary RNA studies (some of them limited to the description of different ESTs). Protein accumulation analyses using specific antibodies have been reported only for three HSFs from tomato: LpHSFA1, LpHSFA2 and LpHSFA3. Some peculiarities have been noted from the published information on plant HSFs, which has been recently reviewed in Refs. 4 and 8 (see other references therein). 1. - The existence not only of ubiquitous (as LpHSFA1), but also of heat-inducible HSFs such as LpHSFA2. 2. - The detection of LpHSFA3 in proliferating cell cultures (21), but not in unstressed leaves (unpublished observations mentioned in Ref. 4). Thus, the higher number of plant HSFs could potentially increase the complexity of interactions with other transcription factors, and provide additional functional diversification by evolving distinct specific expression patterns (as demonstrated in this work for HaHSFA9). Previous evidence supporting functional specialization of plant HSFs is scarce and connected only to thermo-tolerance (24, 25) or to cell death defense (26). This was deduced from the specific effects observed by gain of function (24) or, more recently, by loss of function of different HSFs (25, 26).

Functional characteristics and validation of HaHSFA9

Perhaps the most unusual characteristic of transcriptional activation by HaHSFA9 is the detrimental effect of HSE mutations that improve the consensus sequences; by replacing the natural non-consensus nucleotides ("gaps") located between the GAA repeats in the HSEs of *Hahsp17.7G4* (Figure 5). This observation is unprecedented for a HSF. Moreover, as in transgenic tobacco during late embryogenesis the mutant P gene also showed reduced activity, compared to WT, we infer that endogenous HSF(s), with similar transcriptional activation properties as HaHSFA9, are present in this system. Furthermore, the seed-specificity of the effect of mutant P in tobacco would be consistent with HSF(s) with functions mostly embryonic; as proposed above for HaHSFA9 based in its peculiar expression pattern.

The increased basal expression levels observed for the P mutant gene in sunflower embryos, compared to that of WT and m versions (Figure 5B), would indicate that additional HSFs are expressed in these embryos. This is consistent with reports (reviewed in Ref. 4), and our own unpublished observations, showing that cDNAs and ESTs for different plant HSFs can be cloned from samples containing embryos (seeds ± seed-pods). Contrary to HaHSFA9, these additional HSFs would *trans*-activate the P mutant gene more efficiently than the WT gene (Figure 5B). As another counter-example to HaHSFA9, we had shown that the same WT and P genes were equally *trans*-activated by LpHSFA1, in similar experiments performed in sunflower embryos (16). Thus, the DNA-binding (and/or) activation characteristics of HaHSFA9 would be exceptional among plant HSFs.

Previous work from our lab suggested the critical involvement, of atypical HSFs, in the seed-specific expression of *Hahsp17.6G1* (10,12,20), and in the developmental regulation of *Hahsp17.7G4* (11). The gapped HSEs in both promoters were necessary for their transcriptional activation during late embryogenesis in transgenic tobacco (11,20). In the case of *Hahsp17.7G4*, such involvement was inferred from the different effect of very specific HSE mutations, on the heat response, and on the developmental regulation (11). On the other hand, the *Hahsp17.6G1* promoter does not respond to heat shock (10, 20), and its HSE is very selective for activation by the heterologous HSFs LpHSFA1 and LpHSHA2 (12). These HSFs, when individually tested with the *Hahsp17.6G1* HSE, activated transcription only marginally either in yeast cells or in sunflower embryos (12). This contrasts the efficient *trans*-activation observed with HaHSFA9 in the analogous experiments respectively shown in Figures 4A (bottom) and 4B. LpHSFA1, or LpHSHA2, also efficiently *trans*-activated through different HSEs using the same yeast (21) and sunflower systems (16). Our precedent studies did not anticipate two properties of HSFs involved in developmental regulation in plants: 1. - The embryo-specificity. 2. - The detrimental transcriptional effect of HSE-consensus improvement. These two unique properties, together with the efficient activation through the *Hahsp17.6G1* HSE, have been crucial for functional identification of HaHSFA9

in sunflower, and to confirm the occurrence of equivalent HSFs in tobacco. Thus, HaHSFA9 is most likely involved in the developmental regulation of a peculiar subset of sHSP genes –including *Hahsp17.6G1* and *Hahsp17.7G4*–during zygotic embryogenesis. The findings reported here open the possibility of investigating functions that have been proposed for such sHSP genes (see, for example, Ref. 17). This might be achieved by gain-of-function, or loss-of-function, approaches in transgenic plants using HaHSFA9.

ACKNOWLEDGEMENTS

We thank support from Ph. D. fellowships from the Spanish “Ministerio de Educación y Cultura” (to A. R. and J. D. M.). We also thank Pilar Bazaga for excellent technical work.

REFERENCES

1. Jedlicka, P., Mortin, M. A., and Wu, C. (1997) *Embo J.* **16**, 2452-2462
2. Rallu, M., Loones, M., Lallemand, Y., Morimoto, R., Morange, M., and Mezger, V. (1997) *Proc Natl Acad Sci U S A* **94**, 2392-2397
3. Morimoto, R. I. (1998) *Genes Dev* **12**, 3788-3796
4. Nover, L., Bharti, K., Döring, P., Mishra, S. K., Ganguli, A., and Scharf, K. D. (2001) *Cell Stress Chaperones* **6**, 177-189
5. Fiorenza, M. T., Farkas, T., Dissing, M., Kolding, D., and Zimarino, V. (1995) *Nucleic Acids Res* **23**, 467-474
6. Mathew, A., Mathur, S. K., and Morimoto, R. I. (1998) *Mol Cell Biol* **18**, 5091-5098
7. Tanabe, M., Sasai, N., Nagata, K., Liu, X. D., Liu, P. C., Thiele, D. J., and Nakai, A. (1999) *J Biol Chem* **274**, 27845-27856
8. Schöffl, F., Prändl, R., and Reindl, A. (1998) *Plant Physiol* **117**, 1135-1141
9. Nover, L., Scharf, K. D., Gagliardi, D., Vergne, P., Czarnecka, V. E., and Gurley, W. B. (1996) *Cell Stress Chaperones* **1**, 215-223

10. Carranco, R., Almoguera, C., and Jordano, J. (1997) *J Biol Chem* **272**, 27470-27475
11. Almoguera, C., Prieto, D. P., and Jordano, J. (1998) *Plant J* **13**, 437-446
12. Rojas, A., Almoguera, C., Carranco, C., Scharf, K.D., and Jordano, J. (2002). *Plant Physiol* **129**, 1207-1215
13. Meijer, A.H., Schouten, J., Ouwerkerk, P.B.F., and Hoge, J.H.C. (2000) *Yeast as a versatile tool in transcription factor research* In *Plant Molecular Biology Manual* (Gelvin, S.B., Schilperoort, R.A., and Verma, D.P.S., eds) **E3**, pp. 1-28, Kluwer Academic Publishers, Dordrecht, The Netherlands.
14. Almoguera, C., Coca, M. A., and Jordano, J. (1993) *Plant J* **4**, 947-958
15. Boscheinen, O., Lyck, R., Queitsch, C., Treuter, E., Zimarino, V., and Scharf, K. D. (1997) *Mol Gen Genet* **255**, 322-331
16. Rojas, A., Almoguera, C., and Jordano, J. (1999) *Plant J* **20**, 601-610
17. Coca, M. A., Almoguera, C., Thomas, T. L., and Jordano, J. (1996) *Plant Mol Biol* **31**, 863-876.
18. Li, J. J., and Herskowitz, I. (1993) *Science* **262**, 1870-1874
19. Wu, C. (1995) *Annu Rev Cell Dev Biol* **11**, 441-469
20. Carranco, R., Almoguera, C., and Jordano, J. (1999) *Plant Physiol* **121**, 723-730
21. Bharti, K., Schmidt, E., Lyck, R., Heerklotz, D., Bublak, D., and Scharf, K. D. (2000) *Plant J* **22**, 355-365

22. Meisinger, C., and Grothe, C. (1996) *Mol Biotechnol* **5**, 289-291
23. Kanei, I. C., Tanikawa, J., Nakai, A., Morimoto, R. I., and Ishii, S. (1997) *Science* **277**, 246-248
24. Prändl, R., Hinderhofer, K., Eggers, S. G., and Schöffl, F. (1998) *Mol Gen Genet* **258**, 269-278
25. Mishra, S. K., Tripp, J., Winkelhaus, S., Tschiersch, B., Theres, K., Nover, L., and Scharf, K. D. (2002) *Genes Dev* **16**, 1555-1567
26. Yamanouchi, U., Yano, M., Lin, H., Ashikari, M., and Yamada, K. (2002) *Proc Natl Acad Sci U S A* **99**, 7530-7535

FOOTNOTES

This work was funded through grant BIO99-794 (to J.J., from the Spanish "Ministerio de Ciencia y Tecnología: Plan Nacional de I+D+I").

¹ The abbreviations used are: sHSP(s), small heat shock protein(s); HSE, heat shock element; HSF, heat shock factor; DBD, DNA-binding domain; HR, hydrophobic heptad repeats; NLS, nuclear localization signal; AHA, aromatic, large-hydrophobic and acidic amino acid motif; bp, base pair(s); dpi, days postimbibition; dpa, days postanthesis.

² The nucleotide, and predicted amino acid, sequence(s) reported in this paper has been submitted to the GenBankTM/ EBI Data Bank with accession number(s) AY099451.

FIGURE LEGENDS

Fig. 1: Full-length mRNA sequence and characteristics of the predicted HaHSFA9 protein. We assembled the nucleotide sequences of two cDNA

clones (HaHSFA9-36 and HFSFA9-38). These cDNAs showed incomplete 5'-UTRs with ends indicated by the vertical marks in the first (HFSFA9-38) and second (HaHSFA9-36) sequence lines. HaHFSFA9-38 and HaHSFA9-36 contained poly-A tails of 24 and 53 nucleotides, respectively inserted after the vertical mark in the last sequence line, or after the last nucleotide in this Figure. The arrow indicates the expected position for the conserved intron present in the genomic sequences. Below the nucleotide sequence, we show the conceptual translation of the HaHSFA9 protein. The stop codon is in boldface. We also indicate putative domains identified by sequence comparison with other HSFs: The DBD including the exceptional amino acid substitution (R, outlined). The oligomerization domain, with overlapping heptad repeats (HR-A/B) indicated by small circles and asterisks; a putative bipartite nuclear localization sequence (NLS); the HR-C repeat, the 21aa insertion between HR-A and HR-B (boldface), and finally the putative AHA motifs (underlined) containing tryptophan residues (w1-w3).

Fig. 2: Accumulation patterns of HaHSFA9 mRNAs. Top panel: expression of mRNAs during zygotic embryogenesis and disappearance upon seed germination. The developmental stage is indicated as days post-anthesis (dpa) or days post-imbibition (dpi). Bottom panel: absence of mRNAs in control (C) and heat-stressed (HS) vegetative organs. A marginal accumulation of mRNAs in response to drought stress (DS) could be detected only after much longer (72 h) autoradiography time than for the embryo samples (18 h). Molecular weight size markers (MW): *Hpa*II-digested, pBluescript (SK+), labeled-DNA. The quality and quantity of total RNA samples used in the Ribonuclease A protection assays were verified by electrophoresis in 1% agarose (gel pictures shown below each panel). The scheme at the bottom depicts the 5'-end of transcribed HaHSFA9 sequences (solid line: UTR, box: coding region with the predicted conserved intron position as an inserted dark triangle); and the undigested antisense riboprobe (Probe) used in these assays (dashed line). Thick lines below indicated our interpretation of the protected mRNA fragments 1 and 2

(arrowed in both panels). The question mark indicates a possible alternate mRNA initiation site .

Fig. 3: Seed-specific expression of the HaHSFA9 protein. Western immunodetection confirmed the expression of HaHSFA9 during embryogenesis (from 8 dpa) and mature seeds (≥ 22 dpa, top panel and data not shown). After seed germination, we detected very low levels of the HaHSFA9 protein (arrowed) only until 5dpi in control (C) conditions. Heat- (H) or Drought- (D) stress treatments did not re-induce accumulation of the HaHSFA9 protein in vegetative organs, even forcing film exposure to 30 min. In this case as a positive control, we included the 12-dpa-embryo sample (bottom panel, E). Molecular weight markers (left side) in kD. Sample quality and quantity was confirmed by Ponceau-S staining of total proteins (Top, P), or by immunodetection using antibodies against HSC70 (Bottom, 70).

Fig. 4: The HaHSFA9 cDNA encodes a functional transcription factor. Panel A: (top) functional replacement of the yeast SchSF1 by HaHSFA9. Aliquots ($3\mu\text{l}$) of the indicated RSY4 yeast strains were spotted on YPD, starting from mid-log phase cultures, at decreasing cell density (1:10 dilution steps, from left to right). Plates were incubated at the temperatures shown on the left and photographed after 4 days of growth. (Bottom) transcriptional activation assays, in yeast, using β -galactosidase reporter plasmids containing a minimal promoter with the natural (WT), or mutant (m), HSEs from the *Hahsp17.7G4* (G4HSE) and *Hahsp17.6G1* (G1HSE) genes. We used the HaHSFA9 yeast strain transformed with the indicated reporter plasmid. β -galactosidase assays were performed and the values (in duplicate) for three independent transformants per plasmid combination were averaged. The standard errors are represented with bars. Panel B: *trans*-activation of the *Hahsp17.6G1* promoter by the HaHSFA9 effector plasmid in bombarded sunflower embryos. Bars represent mean β -glucuronidase (GUS) activity, normalized to luciferase activity (LUC). Reporter plasmids (21) used with (+A9) and without (-A9) the HaHSFA9 effector plasmid are indicated. Standard errors of the means are indicated ($n \geq 25$).

Fig. 5: Transcriptional activation in plants of chimeric genes with the *Hahsp17.7G4* promoter: embryo-specific effect of HSE mutant P. The scheme on top (A) depicts the promoter context of all chimeric genes used. These were one with the natural HSEs WT (-1132::GUS, Ref. 16), and the previously described “quasi-null” (E, mutE), and “perfect” (P, mutP) HSE mutants (respectively, Refs. 10 and 15). Core HSE repeats are underlined, with nucleotide substitutions in lower case. Dots and ovals respectively indicate the crucial nucleotides and non-consensus gaps present in the WT HSE and discussed in the text. Panel B: *Trans*-activation by the HaHSFA9 effector plasmid in bombarded sunflower embryos. Bars represent mean β -glucuronidase (GUS) activity, normalized to luciferase activity (LUC). The reporter plasmid used with (+A9) and without (-A9) HaHSFA9 effector plasmid is indicated. Standard errors and size of samples as in Figure 4B. The reduced *trans*-activation observed with the P mutant plasmid is highlighted with a darker grey shade. Bottom: Chimeric gene expression in transgenic tobacco. Panel C (seedlings): lack of effect of mutant P in control (C) and heat stress (HS) conditions. Panel D (embryogenesis): reduced activity of the P mutant gene at 28 dpa. In this case, the statistical significance for the effect is indicated with the same shade as in panel B.

ATATTCTCCTTCAAAAAACACACACACAAAAACAACTTGTTTTTGATTTGTA¹CTTC
 |
 61 ATGGCAGGAGTAGTAAAAGAGTTTGACTCTAGTCATGATATGTTCAAAGTCAGAGAAGAA
 1 M A G V V K E F D S S H D M F K V R E E
 121 CCCAAGGTCGTTTTTAATAATAAAGATTGTCTGAAGAGATGATGAAGATTAAGAGGAA
 21 P K V V F N N K D L S E E M K I K E E
 181 CCCATGTTAGTGTGGATGAGCATGAAGATTTGTGGTGGTGAGGAGGCACCGAAGCCG
 41 P M L V F D E H E D F V G G E E A P K P
 241 ATTGAAGGCTGAGAGACGGCGGACTCCGCGTTTCTCAAGAAAACATTTGAAATGGTG
 61 I E G L R G D D G G P P P P F L K K T F E M V
 301 GATGACCCAACAACAGATTCGATTATTTTCATGGAGTAGTTCAAAAATAGTTTTATTCTT
 81 D D P T T D S I I S W S S S K N S F I L
 361 TGGGATCCTATAAGTTTTCTACTGATCTTTGCGCAAAGATTAAGCACAATAATTTCT
 101 W D P H K F S T D L L P Q R F K H N N F **DBD**
 421 TCTAGCTTTGTTTCGTCAACTAACACTTATAGGTTCAAGAAAATCGATCTCGACCGATGG
 121 S S F V R Q L N T Y **R** F K K I D P D R W
 481 GAATTCGCAAACGAGTTTTTCCAAAAGGTAAGAAGCATTGTAAAGAGACATAAAGAGG
 141 E F A N E F F Q C K G K K H L L R D I K R
 541 AGAACAACCACCCAAAAACAACAAAAACAAGGAAATCAGGAAACAAGAACAACA
 161 R T N Q P Q N T Q K Q E E I R K Q E Q Q
 601 CAATGTTGTGTCATCAACAAACTCTACAATGGAGACCGAGCTAAAGAATCTAAGAAAA
 181 Q C C G H Q T N S T M E T E L K N L R K **HR-A**
 * * * * *
 661 GAAAGAATCACACTTAAACAAGAGATCTTGAAGATGAAACAGCAACAAGAGAACCAGAA
 201 E R I T L K Q E I L K M K Q Q Q E N T E
 * * * * *
 721 AAGCATTGGAATGGTTGAAGAAGCTATGCTAAGAATGGAGTTCAGCAACAGCAGCTG
 221 K H L E M V E E R M L R M E F K Q Q Q L
 ° * ° * ° *
 781 TTAGTTTTCATGTCTAAAGCGTTTAAAGAAATCCGATTTTCGTTAAGCTTTTGCAACATCTG **HR-B**
 241 L V F M S K A F R N P I F V K L L Q H L
 ° *
 841 GTGCAAAAACAAAAACGGGAAGTGTGCAAAATGTGCAAAAACGGAAAACCTTGAGCAAATG
 261 V Q K Q K T G S V E M C K K R K L E Q M **NLS**
 901 CTAAACACTGATGATCTCGATCGTTTTCAAGAATGTGGAATATGATTGAACCTGATGCG
 281 L N T D D L D R F Q E M **W1 N M** I E P D A
 961 TATACTGTGTGTCATCAGATGGATCAGTAAGCCCGCCTGAAGATCAAAAACCGAGTGAC **HR-C**
 301 Y T V L S S D G S V S P P E D Q N T S D
 AAATCTGGATCCAACGGTTTCAGATTACAATTCGGAAAAGCTTTATATTGTGGGAGAACTA
 321 K S G S N G S D Y N S E S F I L **W2 E K L** **AHAs**
 1081 ATGGAGGACGAACTGATATTCCGGCGTGAACAATCGGGTAAAGATCAAAACAGAACTTAT
 341 M E D E L I F G G E Q S G K D Q T E T Y
 1141 CTCCAAGAATGGGAAGAGTTGATCCCAAAAGCT**TAA**GGAAGGGGGATCTCTGATGCAGTCAT
 361 L Q E **W3 E** G E L I P K A
 1201 GTGATCATTATCATCGTAAACATGATAAAGACATCTGGTAGGGTGTGCAGTGTCTCGGG
 1261 TAAGCTTTATCTCTGGTACTGGCGAGTTGATGTAGATGTAGAACAATGGGGTACAGT
 1321 ATAGTGTTTTGTGTGTGTTGTAATTTGTTTTGTTGATTTTGGTTGGGGTCTTGTAGA
 1381 AATAACGATGACAATAAACTCCGTAGTTTCTAATAGAATTTAGGGTTGTAATC







